Dropout Distillation
Supplementary Material

Samuel Rota Bulò ROTABULO@FBK.EU
FBK-irst, Trento, Italy

Lorenzo Porzi PORZI@FBK.EU
FBK-irst, Trento, Italy

Peter Kontschieder PKONTSCHIEDER@MAPILLARY.COM
Mapillary, Graz, Austria - Microsoft Research, Cambridge, UK

Abstract

This document provides the proof of Theorem 1 in the main paper (Rota Bulò et al., 2016).

Proof of Thm.1. The proof of *i)* follows by simple application of Jensen's inequality given the convexity assumption of ℓ:

$$\begin{aligned} J(q) = \mathbb{E}_{\boldsymbol{x}}\left[\ell(f_{\text{dropout}}(\boldsymbol{x}), q(\boldsymbol{x}))\right] \\ \leq \mathbb{E}_{\boldsymbol{x},\sigma}\left[\ell(f_{\Theta^*,\sigma}(\boldsymbol{x}), q(\boldsymbol{x}))\right] = J'(q)\,. \end{aligned}$$

As for *ii)*, by exploiting the expression for ℓ in terms of g_1, g_2, g_3 we have

$$\begin{aligned} J(q) &= \mathbb{E}_{\boldsymbol{x}}\left[g_1(f_{\text{dropout}}(\boldsymbol{x})) + g_2(q(\boldsymbol{x})) \right. \\ &\qquad \left. + f_{\text{dropout}}(\boldsymbol{x})^\top g_3(q(\boldsymbol{x}))\right] \\ &= \mathbb{E}_{\boldsymbol{x},\sigma}\left[g_1(f_{\Theta^*,\sigma}(\boldsymbol{x})) + g_2(q(\boldsymbol{x})) \right. \\ &\qquad \left. + f_{\Theta^*,\sigma}(\boldsymbol{x})^\top g_3(q(\boldsymbol{x}))\right] \\ &\qquad \underbrace{-\mathbb{E}_{\boldsymbol{x},\sigma}\left[g_1(f_{\Theta^*,\sigma}(\boldsymbol{x}))\right] + \mathbb{E}_{\boldsymbol{x}}\left[g_1(f_{\text{dropout}}(\boldsymbol{x}))\right]}_{\Delta} \\ &= \mathbb{E}_{\boldsymbol{x},\sigma}\left[\ell(f_{\Theta^*,\sigma}(\boldsymbol{x}), q(\boldsymbol{x}))\right] + \Delta \\ &= J'(q) + \Delta\,, \end{aligned}$$

where Δ is a constant that does not depend on q.

The same derivations hold true if we replace $\mathbb{E}_{\boldsymbol{x}}\left[\cdot\right]$ with an empirical average. □

References

Rota Bulò, S., Porzi, L., and Kontschieder, P. Dropout distillation. In *Int. Conf. on Machine Learning*, 2016.

Proceedings of the 33^{rd} *International Conference on Machine Learning*, New York, NY, USA, 2016. JMLR: W&CP volume 48.