

Conditional Neural Processes (ICML 2018)

Marta Garnelo, Dan Rosenbaum, Chris J. Maddison, Tiago Ramalho,
David Saxton, Murray Shanahan, Yee Whye Teh, Danilo J. Rezende, S.
M. Ali Eslami
DeepMind London

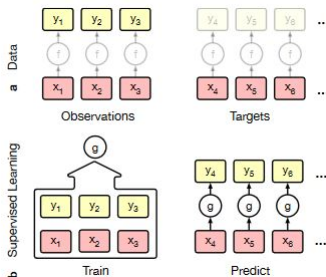
Tarandeep Singh

IIT Hyderabad

April 27, 2022

Introduction

- In deep supervised learning, given a dataset $\{(x_i, y_i)\}_{i=0}^{n-1}$ of n inputs where inputs $x_i \in X$ and outputs $y_i \in Y$, where there is an unknown function $f : X \rightarrow Y$, our algorithm returns an approximation function $g : X \rightarrow Y$, by minimizing loss b/w f and g .
- Generally, this g is randomly initialized with a limited information about distribution of f . Due to this, deep neural networks spends most of the time learning specific task and fails to share learning with different task.



(1)

Introduction

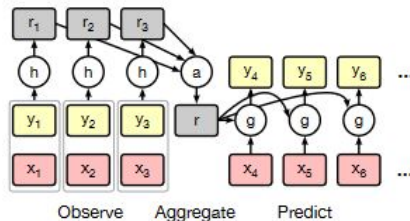
- In Bayesian learning methods, prior information about distribution of f is used based on the practitioner's knowledge about f .
- Bayesian inference is then based on conditioned functional space.
- But Bayesian learning methods, become computationally expensive as dimensionality of dataset increases. Moreover, it is sometimes hard to design models from prior information.
- If n is number of observations and m is number of targets, then complexity of these algorithms is $O((n + m)^3)$

Conditional Neural Processes

- In this paper, authors have proposed a new learning method that combines the function approximation power of supervised learning algorithms with knowledge and prior information about distribution.
- The new family of models called *Conditional Neural Processes*(CNP), proposed by authors define conditional distributions over functions given a set of observations.
- CNPs are permutation invariant of observations O and $T \subseteq X$

Conditional Neural Processes

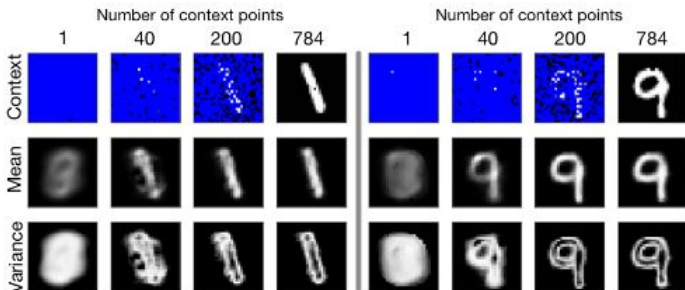
- The complexity of these algorithms is $O(n + m)$ making m predictions with n observations.



(2)

Experimental Results of CNPs

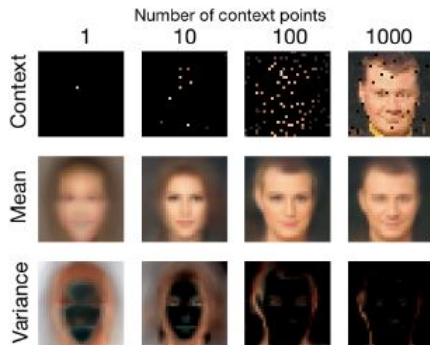
- **Pixel wise image regression** : Authors experimented with different number of context points (prior information or informed observations). They provided trained model with 1, 40, 200 and 728 context points and query the entire image.



(3)

Experimental Results of CNPs

- **Pixel wise image generation:** Authors experimented with different number of context points. They provide the model with 1, 10, 100 and 1000 context points and query the entire image.



(4)

References

- Marta Garnelo, Dan Rosenbaum, Christopher Maddison, Tiago Ramalho, David Saxton, Murray Shanahan, Yee Whye Teh, Danilo Jimenez Rezende, and S. M. Ali Eslami. Conditional neural processes. In Jennifer G. Dy and Andreas Krause, editors, Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018, volume 80 of Proceedings of Machine Learning Research, 1690–1699. PMLR, 2018. URL: <http://proceedings.mlr.press/v80/garnelo18a.html>.
- **Code:** <https://github.com/deepmind/neural-processes>