



Contrastive Learning & Predictability Minimization

Representation learning has been greatly improved with the advance of contrastive learning methods. By pre-training models on large amounts of unlabelled data, one can achieve state-of-the-art results with a fraction of the labelled training data. This makes data gathering orders of magnitude cheaper.

One elegant and intriguing contrastive learning objective was introduced by Zbontar et. al. [Zbo+21].

Rather than teasing apart negative samples, they aim to reduce information redundancy in the embedding space. To do so, they have a loss term minimizing the correlation between the embedding features.

A more thorough approach to reduce information redundancy would be to use predictability minimization, as introduced by Schmidhuber in 1992 [Sch92]. So in this thesis we aim to combine contrastive learning with predictability minimization to achieve state-of-the-art results and to better understand the role of information redundancy in contrastive learning.

Requirements: Strong motivation, proficiency in Python, ability to read papers and work independently. Prior knowledge in deep learning is preferred. We will have weekly meetings to address questions, discuss progress and think about future ideas. I would strongly recommend reading the cited papers before contacting me about this topic :)

Interested? Please contact us for more details!

Contact

- Béni Egressy: begressy@ethz.ch, ETZ G94

References

- [Sch92] Jürgen Schmidhuber. “Learning factorial codes by predictability minimization”. In: *Neural computation* 4.6 (1992), pp. 863–879.
- [Zbo+21] Jure Zbontar et al. “Barlow twins: Self-supervised learning via redundancy reduction”. In: *arXiv preprint arXiv:2103.03230* (2021).

