Stability in Non-Contrastive Learning

Topic

Mitigating Representational Collapse

Background

Contrastive learning heavily relies on negative samples to push representations apart, preventing the model from mapping all inputs to a constant output. However, non-contrastive architectures, such as BYOL or SimSiam, achieve state-of-the-art results without relying on negative pairs.

Assignment Task

Explain the underlying mechanisms that prevent non-contrastive self-supervised learning frameworks from falling into trivial solutions, also known as representational collapse. Analyze the mathematical or structural roles of components such as asymmetric architectures, stop-gradient operations, and momentum encoders in bypassing shortcut learning.

Submission Expectation

Prepare a rigorous, self-contained written response that defines all key assumptions, uses precise technical terminology, and supports the argument with mathematical, architectural, or conceptual reasoning where appropriate.