

Анализ смещения распределений при использовании сравнительного подхода в обучении представления данных

Мария Александровна Никитина

Московский физико-технический институт

Кафедра: Интеллектуальный анализ данных

Научный руководитель: кандидат ф.-м. наук Р. В. Исаченко

2023

Цель исследования

Цель

Устранение смещения распределения, вызванного появлением ложноположительных элементов в задаче поиска представления без учителя.

Задача

Нахождение оптимальной лосс-функции, устраняющей смещение, и её применение в алгоритме обучения модели.

Идея

Выразить распределение положительных элементов через декомпозицию полного распределения:

$$p_x^+(\mathbf{x}') = \frac{p(\mathbf{x}') - \tau^- p_x^-(\mathbf{x}')}{\tau^+},$$

где τ^+ , τ^- – вероятности нахождения положительных и отрицательных элементов соответственно.

Contrastive learning

Contrastive learning – подход при котором обучение происходит не только по принципу близости, но и по принципу различия.

Положим вектор x некоторого объекта в качестве основного. Тогда вектор схожего объекта назовём x^+ – позитивный элемент. Он должен быть как можно ближе к x . А вектор отличного объекта x^- – как можно дальше, как негативный элемент.



Рис.: Принцип contrastive learning. Взято из [Schroff et al. 2015].

Список литературы



Ching-Yao Chuang, Joshua Robinson, Lin Yen-Chen, Antonio Torralba, Stefanie Jegelka (2020)

Debiased Contrastive Learning



Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, Dilip Krishnan (2021)

Supervised Contrastive Learning



CTing Chen, Simon Kornblith, Mohammad Norouzi, Geoffrey Hinton (2020)

A Simple Framework for Contrastive Learning of Visual Representations



Sohn, Kihyuk (2016)

Improved Deep Metric Learning with Multi-class N-pair Loss Objective



Florian Schroff, Dmitry Kalenichenko, James Philbin (2015)

FaceNet: A Unified Embedding for Face Recognition and Clustering

Ложноположительные и ложноотрицательные элементы

Классический лосс, не учитывающий смещения:

$$\mathcal{L}_{N-pair}(f) = -\log \frac{\exp(f(\mathbf{x})^T f(\mathbf{x}_i^+))}{\exp(f(\mathbf{x})^T f(\mathbf{x}_i^+)) + \sum_{i=1}^N \exp(f(\mathbf{x})^T f(\mathbf{x}_i^-))}$$

Решение для устранения смещения ложноотрицательных элементов [Chuang et al., 2020]:

$$L_{\text{DebiasedNeg}}^{N,M}(f) = \mathbb{E}_{\substack{\mathbf{x} \sim p; \mathbf{x}^+ \sim p_x^+, \\ \{\mathbf{u}_i\}_{i=1}^N \sim p_x^N, \\ \{\mathbf{v}_j\}_{j=1}^M \sim p_x^{+M}}} \left[-\log \frac{e^{f(\mathbf{x})^T f(\mathbf{x}^+)}}{e^{f(\mathbf{x})^T f(\mathbf{x}^+)} + Ng(\mathbf{x}, \{\mathbf{u}_i\}_{i=1}^N, \{\mathbf{v}_j\}_{j=1}^M)} \right]$$

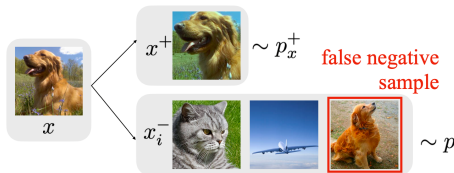


Рис.: Взято из [Chuang et al., 2020].

DebiasedPos

В отличие от $L_{\text{DebiasedNeg}}^{N,M}(f)$ здесь устраняется смещение ложноположительных элементов.

$$\tilde{L}_{\text{DebiasedPos}}^N(f) = \mathbb{E}_{\substack{\mathbf{x} \sim p \\ \mathbf{x}^- \sim p_x^-}} \left[-\log \frac{\mathbb{E}_{\mathbf{x}' \sim p} h(\mathbf{x}, \mathbf{x}') - \tau^- \mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}_i^-)}{\mathbb{E}_{\mathbf{x}' \sim p} h(\mathbf{x}, \mathbf{x}') + (N\tau^+ - \tau^-) \mathbb{E}_{\mathbf{x}^- \sim p_x^-} h(\mathbf{x}, \mathbf{x}_i^-)} \right]$$

Финальная оценка:

$$L_{\text{DebiasedPos}}^N(f) = \mathbb{E}_{\substack{\mathbf{x} \sim p \\ \{\mathbf{u}_i\}_{i=1}^N \sim p_x^- \\ \mathbf{v} \sim p_x^+}} \left[-\log \frac{P_{\text{emp}} - \tau^- P_{\text{emp}}^-}{P_{\text{emp}} + (N\tau^+ - \tau^-) P_{\text{emp}}^-} \right].$$

$P_{\text{emp}}, P_{\text{emp}}^-$ — эмпирические оценки матожиданий.

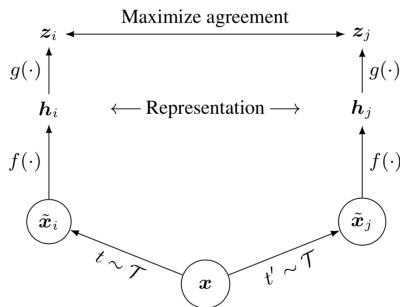
Theorem

Для произвольного представления f и произвольного $\delta > 0$ существует достаточно большое N , что

$$|\tilde{L}_{\text{DebiasedPos}}^N(f) - L_{\text{DebiasedPos}}^N(f)| \leq \left[\left(1 + \frac{\tau^-}{\tau^+} + \delta\right) \sqrt{\frac{\pi}{2N}} + \left(1 + \frac{1}{\tau^+}\right) \sqrt{\frac{\pi}{2N+2}} \right] e^{3/2}$$

SimCLR

CTing Chenet al., 2020



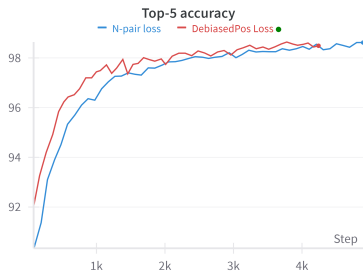
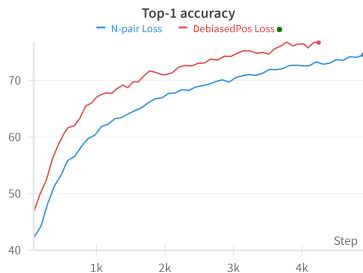
- ▶ \mathcal{T} — семейство аугментаций
- ▶ Семплируются 2 аугментации $t, t' \sim \mathcal{T}$, применяются к каждому объекту.
- ▶ Обучаем сеть-энкодер $f(\cdot)$ и MLP сеть-проекцию $g(\cdot)$, максимизируя соответствие представлений.

Вычислительный эксперимент

Цель

Сравнить работу $L_{\text{DebiasedNeg}}^N(f)$, $L_{\text{DebiasedPos}}^N(f)$ и $L_{N\text{-pair}}(f)$ на модели SimCLR.

Первичный результат



1. Оценка сходимости $L_{\text{DebiasedPos}}$ при выборе нескольких положительных элементов
2. Формула для $L_{\text{DebiasedPos}}$ в случае отказа от предположения о равномерном распределении классов
3. Эксперименты в случае выбора нескольких положительных элементов
4. Эксперименты с комбинированием $L_{\text{DebiasedPos}}$ и $L_{\text{DebiasedNeg}}$