

Self-supervised learning

Автор: Виданов Андрей
Руководитель: Барабанщикова Полина

30 марта 2023 г.

Аннотация

В последние годы все активнее развиваются техники обучения без учителя, которые не требуют предварительной разметки на данных. Для изображений является полезной задача создания представлений, поскольку они могут использоваться в различных задачах машинного обучения для повышения качества финальных моделей. Переход к подобным методам является перспективным, но сопряжен с рядом новых задач. Предлагается рассмотреть методы на основе многокомпонентной функции потерь в частности VICReg и TiCo, а также их модификации. Эксперименты проводятся на наборе данных "CIFAR10" в качестве финального классификатора используется простая линейная модель

Введение

Развитие техник самообучения[6]. В данной области ставится множество задач и применяются различные, в том числе генеративные, методы [4] Как показывают эксперименты, применение классических для традиционного машинного обучения техник, таких как batch norm, может улучшать результаты до передовых в данной области знания [5].

Однако мотивы применения тех или иных подходов зачастую являются сугубо практическими(улучшение качества модели), но не теоретически обоснованными, что ставит под вопрос универсальность их применения для задач своего типа. Теоретические изыскания в данном вопросе активно проводятся [2], тем не менее множество работ сосредоточено на получении эффективных эвристик. Специфика задачи генерации эмбедингов для картинок требует особых методов. В частности, для задач связанными с изображениями нет методов для создания отрицательных пар(пара изображений разных классов, подробнее на paperwithcode[7]), обучаясь на которых модель могла бы улавливать чем отличаются кардинально разные картинки. Избежать данной проблемы можно спроектировав модели, не требующие негативных пар. Модель минимизирует расстояние между эмбедингами похожих картинок. Но данные методы склонны к вырождению, иначе говоря сталкиваются с коллапсом(явление, когда модель сопоставляет всем

картинкам одно и то же представление). В результате чего стоит крайне аккуратно подбирать лосс-функции[2] или проектировать модель. [1], [3].

Приведенные выше техники эффективны на практике, но теоретическая мотивация их использования до сих пор остается непрозрачной. Задачей исследования является нахождения зависимостей между применениями различных техник и качеством получаемых эмбедингов, а также риска столкнуться с коллапсом. Предполагается, модификации лосс-функции приведут к улучшению финальных показателей в прикладных задачах.

Предлагаемые изменения в функции потерь являются точечными из-за чего получится лучше определить их влияние на результат модели.

Измеримыми целями исследования являются численные показатели полученные в ходе применения модификаций к функции потерь. В частности, метод обучается на тренировочной выборке, далее на представлении на тренировочной выборке обучается простой линейный классификатор, после чего на тестовой части данных классификатор предсказывает класс изображения, затем считается точность предсказания.

Список литературы

- [1] Xinlei Chen Kaiming He - "Exploring Simple Siamese Representation Learning"
- [2] Adrien Bardes, Jean Ponce, Yann LeCun - "VICREG: VARIANCE-INVARIANCE-COVARIANCE REGULARIZATION FOR SELF-SUPERVISED LEARNING"
- [3] Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, Stephane Deny ´
- [4] Xiao Liu, Fanjin Zhang, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, Jie Tang - "Self-supervised Learning: Generative or Contrastive"
- [5] Abe Fetterman Josh Albrecht - Understanding Self-Supervised and Contrastive Learning with "Bootstrap Your Own Latent"(BYOL)
- [6] Самообучение. Проблематика и постановка задачи
- [7] Задача классификации по представлениям