

Теоретико-информационный анализ нейронных сетей & Оптимальные разбиения множеств на части меньшего диаметра

Александр Толмачев

Московский физико-технический институт, Сколковский институт науки и технологий

Научные руководители: Алексей Фролов / Андрей Райгородский

20 мая 2023 г.

Постановка задачи

Information Bottleneck

Теоретико-информационный анализ предлагает оценивать динамику двух значений взаимной информации в нейронных сетях: между входом и скрытым слоем, и между скрытым слоем и истинной меткой класса.

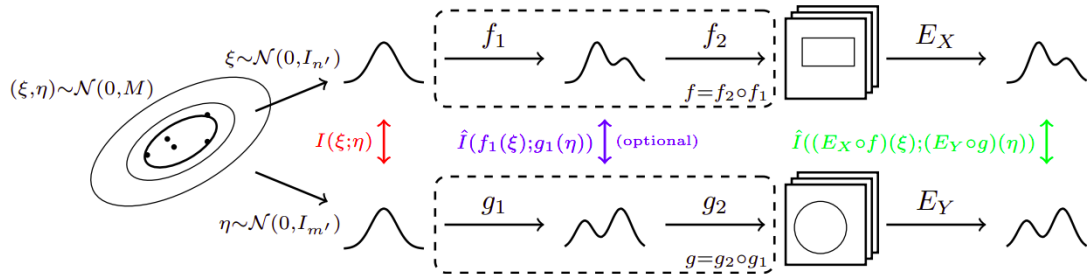
Проблема

Взаимная информация между многомерными случайными величинами трудно оценивать, поэтому предлагаемые ранее гипотезы об упомянутых выше величинах были проверены либо на маленьких “игрушечных” сетях, либо на нейронных сетях определенного вида.

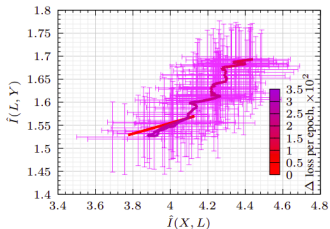
Предлагаемое решение

В нашем подходе предлагается оценивать взаимную информацию по **сжатым** представлениям, которые получены с помощью обученных автоэнкодеров.

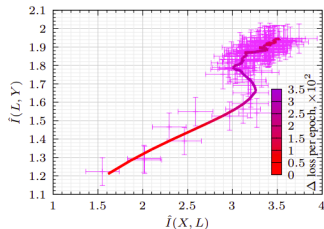
Краткое описание алгоритма для синтетических данных



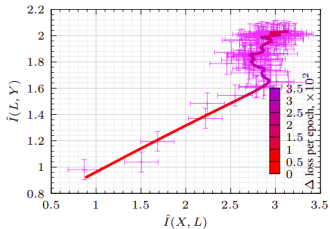
Результаты анализа обучени CNN-классификатора на датасете MNSIT



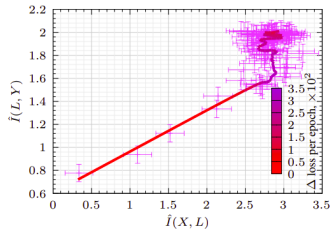
(a) L_2 (convolutional, LeakyReLU)



(b) L_3 (convolutional, LeakyReLU)



(c) L_4 (fully-connected, LeakyReLU)



(d) L_5 (fully-connected, LogSoftMax)

Спасибо за внимание! Однако, это еще не всё...

Разбиения множеств на плоскости

Определение 1

Пусть F — произвольное ограниченное множество на плоскости, $n \in \mathbb{N}$. Тогда

$$d_n(F) = \inf\{x \in \mathbb{R}^+ : \exists F_1, \dots, F_n : F \subseteq F_1 \cup \dots \cup F_n, \forall i \text{ diam}(F_i) \leq x\}.$$

Определение 2

Пусть

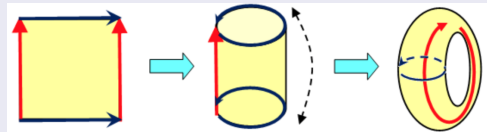
$$d_n = \sup_{\Omega \subset \mathbb{R}^2} d_n(\Omega) = \sup_{\Omega \subset \mathbb{R}^2} \inf \max_{1 \leq i \leq n} \text{diam } \Omega_i, \quad \text{diam } \Omega = 1.$$

Заметим, что $d_1 = d_2 = 1$ и $d_n \geq d_{n+1}$, т.к. одно из множеств разбиения может быть пустым, а также мы можем рассматривать покрытия замкнутыми выпуклыми множествами постоянной ширины 1.

Разбиения поверхности двумерного тора

Поверхность двумерного тора

Представим поверхность двумерного тора как факторпространство $\mathbb{R}^2/\mathbb{Z}^2$.
Неформально говоря, это квадрат со стороной 1, пары противоположных сторон которого “склеены”.



Актуальность задачи

- Разбиение поверхности тора является естественным обобщением известной задачи по разбиениям плоских множеств
- Изучение структуры близких к оптимальным разбиений поверхности тора позволит заметить новые закономерности и улучшить оценки для разбиения множеств пространства большей размерности
- Отдельно отметим, что такая задача рассматривается **впервые** в данной работе

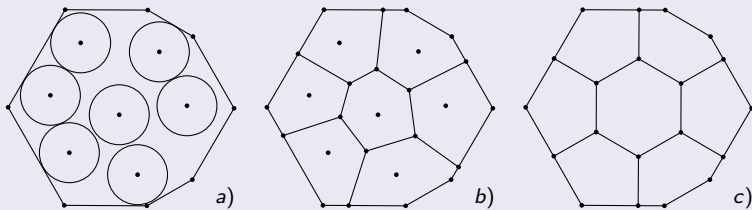
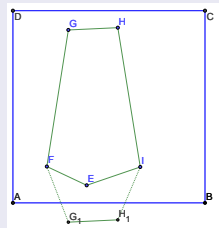
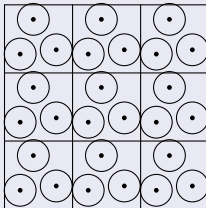
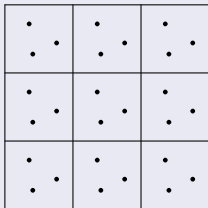


Рис.: Грубое приближение в задаче плотной упаковки кругов (a), диаграмма Вороного (b) и окончательное разбиение (c)

Адаптация алгоритма для тора

- 1 Множество $\Omega \subset \mathbb{R}^2$ является квадратом 3×3
- 2 Каждая точка присутствует в 9 экземплярах в каждом квадратице
- 3 При обновлении текущего оптимума $X^* := X_1$ дополнительно проверяем что данное разбиение "корректно"



Разбиение множеств на торе. Верхние оценки

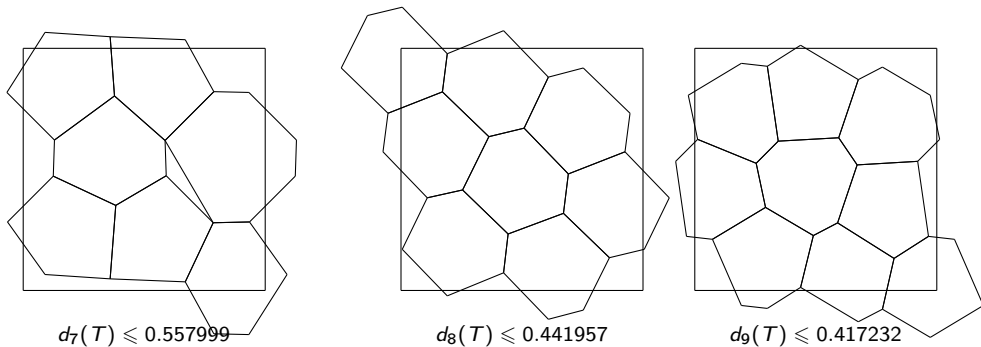


Рис.: Разбиение тора на 7, 8, 9 частей

Разбиение множеств на торе. Верхние оценки

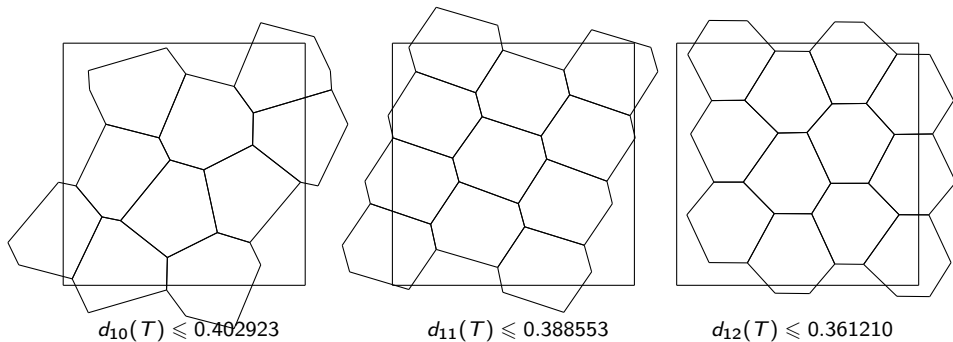
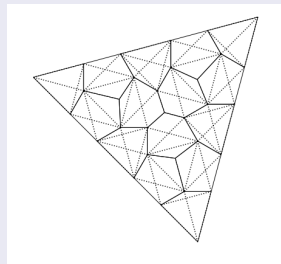
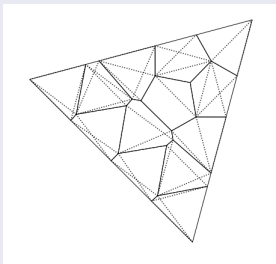
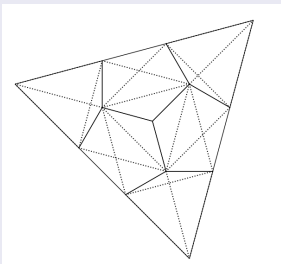


Рис.: Разбиение тора на 10, 11, 12 частей

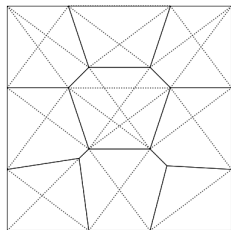
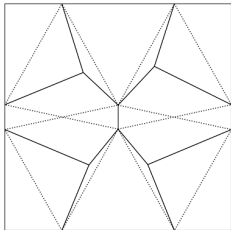
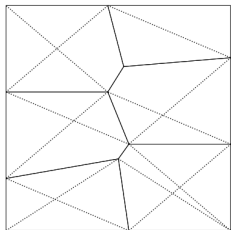
Разбиения правильного треугольника: $n = 6(7), 13, 15$

- $d_{\Delta}(6) = d_{\Delta}(7) = \frac{1}{1+\sqrt{3}}$
- $d_{\Delta}(15) \leq \frac{1}{1+2\sqrt{3}}$



Разбиения квадрата: $n = 6, 8, 9$

- $d_{\square}(6) \leq \frac{1}{3\sqrt{3}} \sqrt{7 - \frac{200}{a}} + a; \quad a = \sqrt[3]{2(383 + 129^{3/2})}$
- $d_{\square}(8) \leq \frac{1}{12} \sqrt{80 - 16\sqrt{7}}$



Спасибо за внимание!