Нейросетевой подход для решения NP-трудной задачи минимизации энергии в несубмодулярной модели Изинга

Суглобов Кирилл Алексеевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: ksuglobov@outlook.com

Научный руководитель — *Сердюк Юлиан Анатольевич*

Модель Изинга — это модель статистической физики, предложенная для объяснения магнетизма [1]. Она представляет собой кристаллическую решётку, заданную графом $G(\mathcal{V}, \mathcal{E})$, вершины $\mathcal{V}$ которого соответствуют спинам $x_i \in \{-1, +1\}$, $i \in \mathcal{V}$, а рёбра $\mathcal{E}$ обозначают взаимодействие между соседними спинами $i, j \in \mathcal{E}$.

Спины взаимодействуют с соседями через парные потенциалы $\theta_{ij}^P(x_i, x_j) \in \mathbb{R}$ и с внешним магнитным полем через унарные потенциалы $\theta_i^U(x_i) \in \mathbb{R}$. Это формирует энергию модели Изинга $E(x) \in \mathbb{R}$.

Переопределим спины как $x_i \in \{0, 1\}$, что делает потенциалы и энергию неотрицательными: $\theta_i^U, \theta_{ij}^P, E(x) \in \mathbb{R}_+$. Требуется минимизировать энергию системы, эквивалентной исходной модели Изинга:

$$E(x) = \sum_{i \in \mathcal{V}} \theta_i^U(x_i) + \sum_{(i,j) \in \mathcal{E}} \theta_{ij}^P(x_i, x_j) + E_0 \to \min_x, \tag{1}$$

где $x \in \{0, 1\}^{|\mathcal{V}|}$, $E_0 = \mathrm{const}$. Задача (1) возникает в компьютерном зрении [2], где энергия моделирует согласованность меток пикселей.

Субмодулярность Энергия, и сама модель, субмодулярна, если:

$$\theta_{ij}^P(0,0) + \theta_{ij}^P(1,1) \le \theta_{ij}^P(0,1) + \theta_{ij}^P(1,0), \quad \forall\,(i,j) \in \mathcal{E}. \tag{2}$$

Для субмодулярных моделей задача минимизации энергии лежит в классе P и решается алгоритмом Min-Cut [2]. Если условие (2) нарушается, задача становится NP-трудной. Для несубмодулярных моделей используется алгоритм QPBO [2], который даёт частичное решение: часть спинов совпадает с оптимальным решением, а оставшиеся не определены. При небольшом их количестве, решение можно доопределить до оптимального перебором.

В данной работе для нахождения оптимального решения, необходимого для оценки качества, применяется именно этот подход.

Нейросетевой подход Рассматривается двумерная квадратная модель Изинга размера $n \times n$, где каждый спин взаимодействует с четырьмя соседями (по вертикали и горизонтали). Конфигурация спинов представлена матрицей $x \in \{0,1\}^{n \times n}$.

В данной работе предложена нейросеть f_φ, которая принимает тензор потенциалов θ (унарных θ^U и парных θ^P) и преобразует его в матрицу вероятностей q^φ:

$$f_\varphi : \theta \mapsto q^\varphi, \quad \text{где } q^\varphi_{i,j} = \mathbb{P}(x_{i,j} = 1). \tag{3}$$

Нейросеть (3) обучается минимизировать релаксированную энергию на выборке моделей, полученных из распределения $p(\theta)$. Функционал потерь имеет вид:

$$\mathbb{E}_{\theta \sim p(\theta)} E(q^\varphi, \theta) \to \min_\varphi,$$

где $E(q^\varphi, \theta)$ — энергия модели с потенциалами θ, релаксированная за счёт замены булевых спинов $x_{i,j}$ на вероятности $q^\varphi_{i,j}$.

Оценка качества Качество решения измеряется нормированной энергией (0 — оптимум, 100 — случайное решение) и точностью: $\text{Accuracy} = \frac{1}{n^2} \sum_{i,j} \mathbb{I}(x_{i,j} = x^*_{i,j})$, где x^* — оптимальное решение.

Решение нейросети сравнивается с оптимальным, а также с эвристиками: унарной ($x_i = \arg\min \theta^U_i(x_i)$) и парной (однородной) ($x_i \equiv 0/1$). Они минимизируют соответствующие вклады в энергию.

Генерация данных Рассматриваются несубмодулярные модели Изинга размера 8×8: $\theta^U_i, \theta^P_{ij} \overset{\text{iid}}{\sim} \mathcal{U}[0,1]$. Обучающая выборка формируется через семплирование с отклонением: модели решаются QPBO и случаи с > 7 неопределёнными спинами отбрасываются, иначе решение доопределяется до оптимального с помощью перебора.

В данной работе показано, что такой метод сохраняет исходное распределение несубмодулярных взаимодействий и низкое качество тривиальных эвристик, то есть сохраняет сложности задачи.

Архитектура нейросети Потенциалы модели Изинга $n \times n$ записываются в виде тензора размерности $n \times n \times 10$: 2 канала на унарные потенциалы, и 8 каналов — на парные.

Для решения задачи выбрана и модифицирована архитектура нейросети U-Net с механизмом 3D-внимания [3]. Это позволяет учитывать локальные и глобальные зависимости между спинами, как в задачах сегментации [2].

Результаты Нейросеть обучалась на 32,000 моделях и валидировалась на 4,000. Лучшей оказалась архитектура U-Net глубины 2 с 3D-вниманием (обозначение модели: `d2-s16-w3-attn_3d`).

Результаты тестирования (метрики энергии и точности) на выборке из 10,000 моделей представлены в Таблице 1.

Нейросеть демонстрирует скорость, сопоставимую с комбинацией QPBO и перебора на малых моделях 8×8. Однако, в отличие от классического подхода, её время работы не зависит от числа неопределенных спинов. Это даёт решающее преимущество в задачах, где перебор вычислительно неприемлем.

Иллюстрации

Метрика	**Оптимум**	**Нейросеть**	**Унарная эвр.**	**Парная эвр.**
Энергия (↓)	0.00	**0.66** [0.16; 1.51]	50.14 [42.28; 58.19]	90.56 [83.26; 97.09]
Точность (↑), %	100.00	**95.31** [93.75; 98.44]	67.19 [64.06; 71.88]	53.12 [50.00; 56.25]

Таблица 1. Качество решений для несубмодулярных моделей Изинга размера 8×8; 10,000 семплов; формат: `Median [P25; P75]`.

Предложен метод решения задачи минимизации энергии в несубмодулярной модели Изинга, основанный на нейросети U-Net с 3D-вниманием. На задаче размера 8×8 метод достигает точности 95.31% при константном времени работы, что потенциально даёт преимущество перед классическими подходами (QPBO с экспоненциальным перебором) для задач большего размера.

Результаты демонстрируют перспективность сочетания релаксации дискретных функционалов и глубокого обучения для решения NP-трудных задач комбинаторной оптимизации, однако масштабируемость метода требует дополнительного исследования.

Литература

1. Ising E. Beitrag zur Theorie des Ferromagnetismus // Zeitschrift für Physik. 1925. Vol. 31. pp. 253–258.
2. Kolmogorov V., Rother C. Minimizing Nonsubmodular Functions with Graph Cuts: A Review // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007. Vol. 29, №. 7. pp. 1274–1279.
3. Oktay O., Schlemper J., Le Folgoc L. et al. Attention U-Net: Learning Where to Look for the Pancreas // arXiv preprint `arXiv:1804.03999`. 2018.