

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ

А.Т. Агишев

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ ОБРАБОТКИ ДАННЫХ

**Сборник лекций для студентов магистратуры,
обучающихся по образовательной программе
«7М07125 - Электроника и системы управления»**

Алматы, 2025

Лекция 5. Самоорганизующиеся карты Кохонена: теория и применение

Цель лекции

Рассмотреть принципы построения и функционирования самоорганизующихся карт Кохонена (Self-Organizing Maps, SOM), их математическую основу, этапы обучения и практические области применения в интеллектуальных системах обработки данных.

Основные вопросы:

1. Идея самоорганизации и обучения без учителя.
2. Архитектура и элементы самоорганизующихся карт.
3. Математическая модель и алгоритм обучения.
4. Этапы формирования карты и топологическая организация.
5. Влияние параметров обучения на результат.
6. Применение SOM для анализа, кластеризации и визуализации данных.

Краткие тезисы:

1. Идея самоорганизации. Самоорганизующиеся карты - это нейронные сети, которые обучаются без учителя, то есть без заранее известных правильных ответов. Их задача - выявить скрытую структуру данных и упорядочить многомерные входные векторы на двумерной карте, сохраняя топологические отношения между ними. Такая сеть способна группировать схожие объекты и представлять сложные данные в наглядной форме.

2. Архитектура карты Кохонена. Сеть состоит из двух слоёв:

- **входного слоя**, принимающего многомерный вектор признаков;
- **выходного (картографического) слоя**, организованного в виде двумерной решётки нейронов.

Каждый нейрон выходного слоя имеет собственный *весовой вектор*, характеризующий его положение в пространстве признаков.

3. Математическая модель и обучение. Пусть входной вектор имеет вид

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

а весовой вектор нейрона i -

$$\mathbf{w}_i = (w_{i1}, w_{i2}, \dots, w_{in})$$

Обучение осуществляется в несколько шагов:

1. **Выбор входного вектора** из обучающего множества.
2. **Поиск нейрона-победителя (Best Matching Unit, BMU)** - нейрона, чей весовой вектор ближе всего к входному по метрике Евклида:

$$i^* = \arg \min_i || \mathbf{x} - \mathbf{w}_i ||$$

3. Обновление весов нейрона-победителя и его соседей:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) \cdot h_{i,i^*}(t) \cdot (\mathbf{x} - \mathbf{w}_i(t))$$

Где $\alpha(t)$ - скорость обучения, $h_{i,i^*}(t)$ - функция соседства, уменьшающая влияние с расстоянием от победителя.

4. Постепенное уменьшение $\alpha(t)$ и радиуса соседства с течением обучения.

Таким образом, карта «самоорганизуется», формируя топологически согласованное представление данных.

4. Топологическая организация и интерпретация карты.

Нейроны, отвечающие на схожие входные векторы, располагаются рядом друг с другом на карте. Результатом обучения является двумерная карта, на которой области соответствуют различным кластерам данных. SOM визуализирует многомерное пространство признаков, сохраняя их взаимную близость - это удобно для анализа и интерпретации сложных зависимостей.

5. Параметры обучения. Качество работы SOM зависит от выбора параметров:

- начальных значений весов;
- скорости обучения $\alpha(t)$;
- радиуса функции соседства;
- количества итераций.

На ранних этапах обучения карта перестраивается активно, затем процесс стабилизируется. Оптимальные значения подбираются экспериментально.

6. Применение самоорганизующихся карт. Самоорганизующиеся карты используются для:

- кластеризации и группировки данных без разметки;
- анализа и сжатия многомерных признаков;
- визуализации больших массивов данных;
- обнаружения аномалий и закономерностей;
- распознавания образов и сигналов;
- анализа технических и экономических показателей.

В интеллектуальных системах SOM применяются как инструмент предварительной обработки данных, а также как часть гибридных моделей совместно с другими нейросетевыми и статистическими методами.

Вопросы для контроля, изучаемого материал:

- 1) В чём заключается принцип самоорганизации в картах Кохонена?
- 2) Как определяется нейрон-победитель в процессе обучения?

- 3) Опишите роль функции соседства и скорости обучения в SOM.
- 4) Почему самоорганизующиеся карты удобны для визуализации многомерных данных?
- 5) В каких областях инженерии и аналитики применяются SOM?
- 6) Какие факторы влияют на качество и стабильность обучения карты?

Рекомендуемый список литературных источников:

1. Hopgood A. A. Intelligent Systems for Engineers and Scientists. — 3rd ed. — Boca Raton : CRC Press / Taylor & Francis, 2012. — 682 p.
2. Sterling T., Anderson M., Brodowicz M. High Performance Computing: Modern Systems and Practices. — Amsterdam : Elsevier / Morgan Kaufmann, 2017. — 728 p.
3. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. 4th Edition. Pearson, 2021.