

# Графы в машинном обучении

Романов Владимир БПМИ192

Национальный исследовательский университет  
«Высшая школа экономики» (Москва)

23 ноября 2021 г.

# Какие задачи на графы нам интересны?

## Задачи на вершины

- link-prediction
- node-classification

## Задачи на графы

- graph-classification

# Vertex embeddings (random walks)

Рассмотрим embedding, которые описывают локальные свойства

Мы хотим, чтобы соседи имели близкий вектор, а далекие разные

# Vertex embeddings (random walks)

Рассмотрим embedding, которые описывают локальные свойства

Мы хотим, чтобы соседи имели близкий вектор, а далекие разный

Формулировка напоминает аналогичную задачу для текста

Идея: попытаемся применить word2vec для этой задачи

# DeepWalk (Алгоритм)

Embedding:  $\Phi : V \rightarrow \mathbb{R}^d$

Loss: 
$$\sum_{u \text{ и } v \text{ «близко»}} -\ln P(u \mid \Phi(v))$$

## Алгоритм

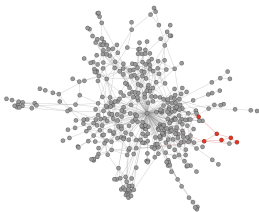
- 1 Переберем вершину  $s \in V$
- 2 Рассмотрим случайный путь  $\mathcal{W}_s$  размера  $t$
- 3 Прорелаксируем SkipGram пройдясь по  $\mathcal{W}_s$  окном  $w$
- 4 Повторим процесс  $\gamma$  раз

Проблема: SkipGram — делать Softmax долго

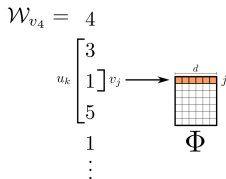
Решение: воспользуемся деревом отрезков

$$P(u | \Phi(v)) = \prod_{b_i \in B} P(b_i | \Phi(v))$$

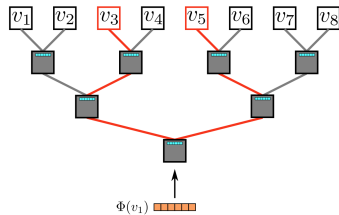
## Итоговый алгоритм:



(a) Random walk generation.



(b) Representation mapping.



(c) Hierarchical Softmax.