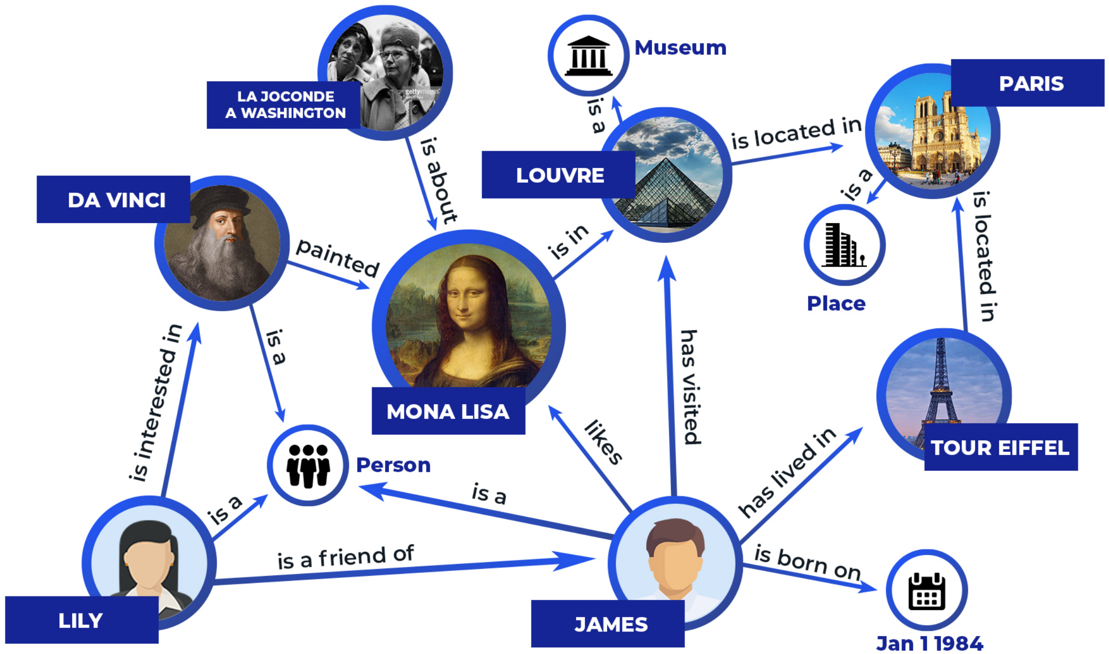
Knowledge Graph Embedding Compression

Mrinmaya Sachan

Toyota Technological Institute at Chicago
ETH Zurich

Knowledge Graph Representations



<u>**Entities**</u>

Da Vinci

Mona Lisa

Louvre

...

<u>**Relations**</u>

painted

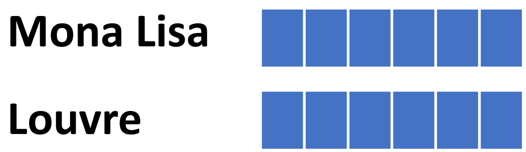
is in

...

Large Knowledge Graph Representation Layer

KGs can be very large – millions or billions of entities in a KG, resulting in a large KG representation layer

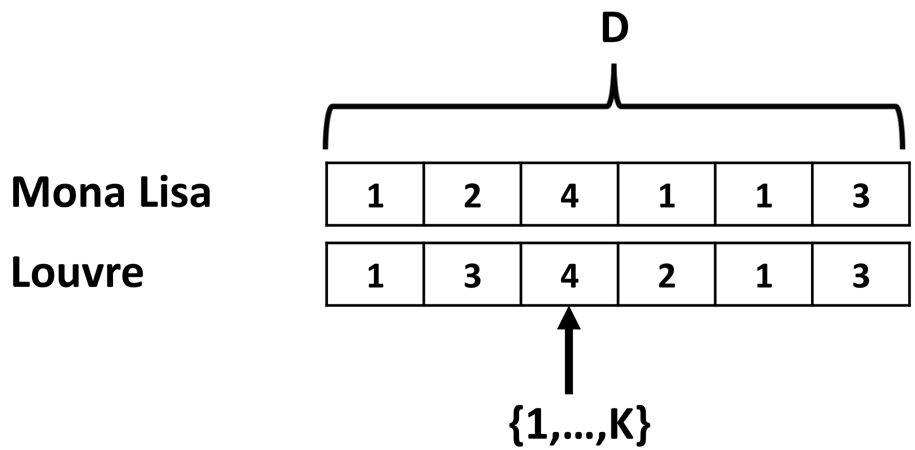
Redundancy in KG representations – many entities are actually similar!



Discrete KG representations

Inspired from previous work in discrete representation learning (van den Oord et al., 2017; Chen et al., 2018; Chen and Sun, 2019) for language and vision, this paper presents discrete KG representation learning

KD codes:

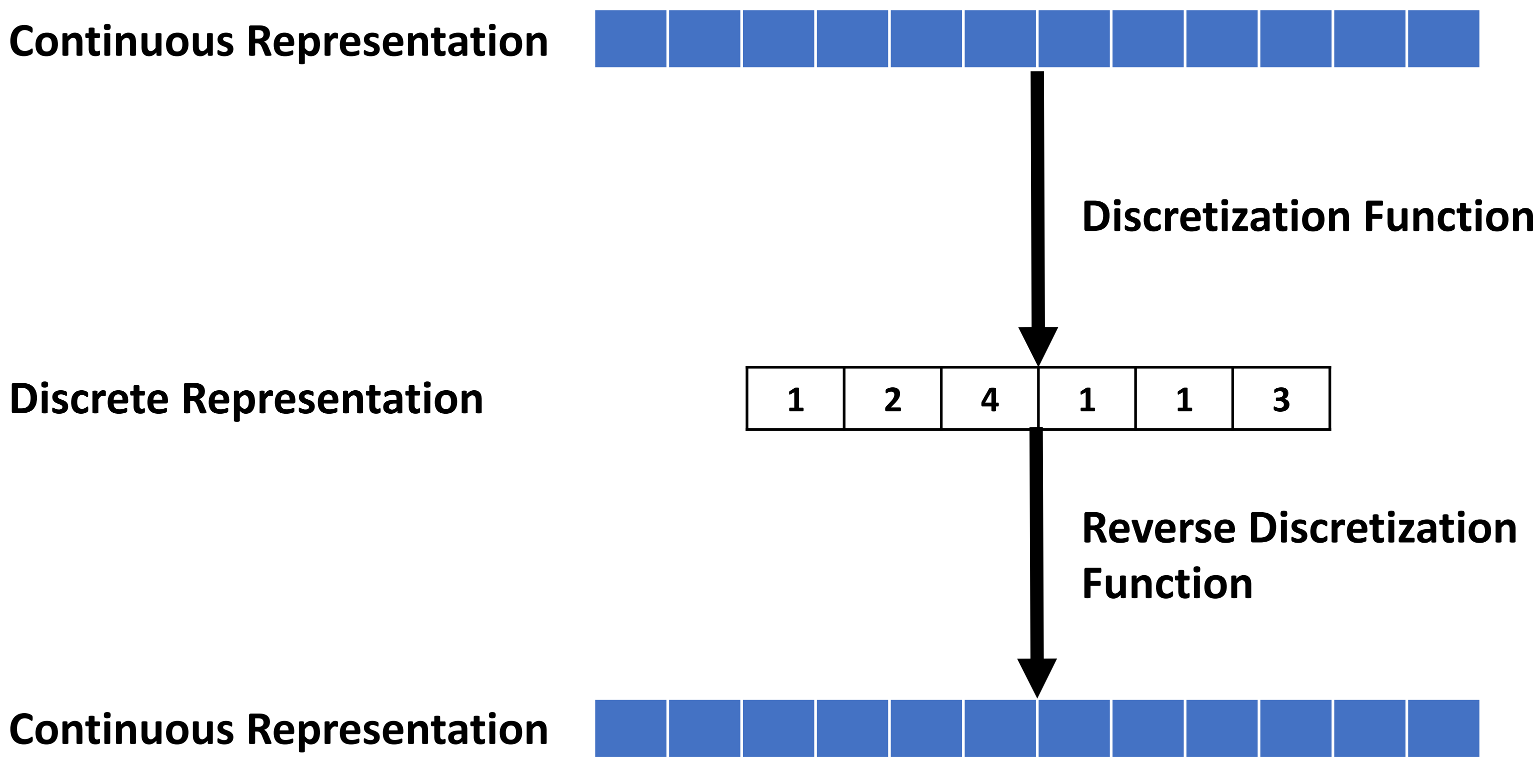


Learnt in a way to capture the the semantics and relational structure in KGs

Key advantage: Embedding compression

Discrete KG representation learning

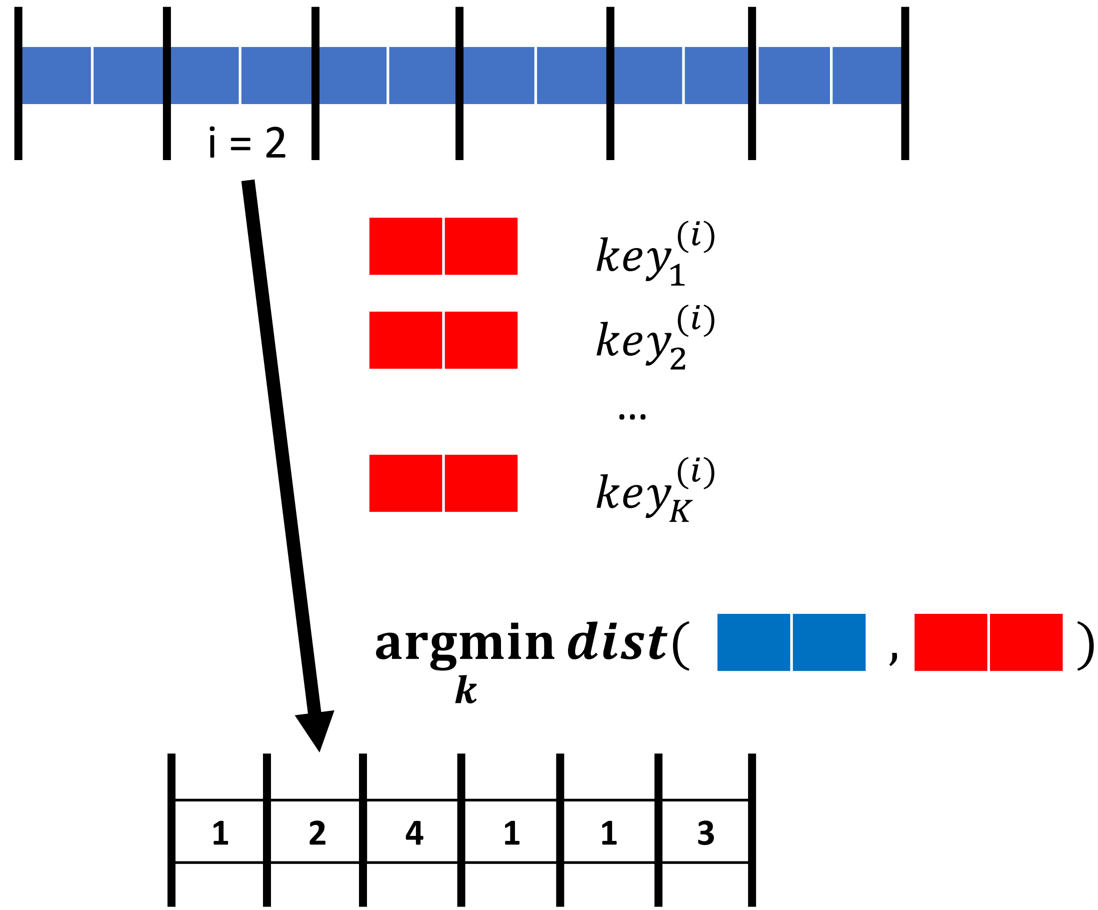
(van den Oord et al., 2017; Chen et al., 2018; Chen and Sun, 2019)



Discretization Function

(van den Oord et al., 2017; Chen et al., 2018; Chen and Sun, 2019)

Continuous Representation



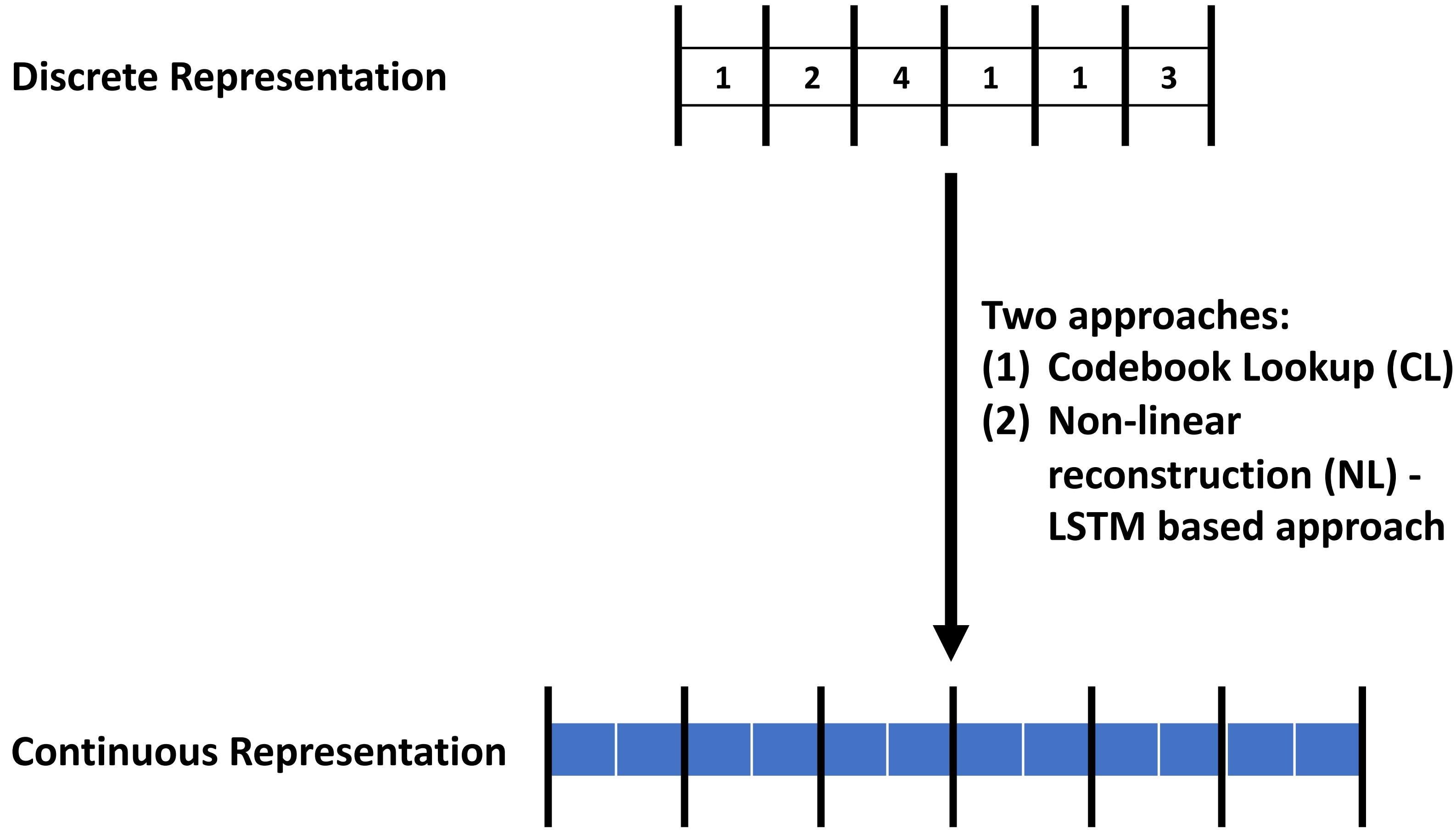
Discrete Representation

Two approaches:

1. **Vector Quantization (VQ)**
2. **Tempering Softmax (TS)**

Reverse-Discretization Function

(van den Oord et al., 2017; Chen et al., 2018; Chen and Sun, 2019)



Experiments

Datasets:

- Freebase 15k (Bordes et al., 2013)
- Freebase 15k-237 (Toutanova et al., 2015)
- Wordnet 18 (Bordes et al., 2013)
- Wordnet 18RR (Dettmers et al., 2018)

Link Prediction tasks

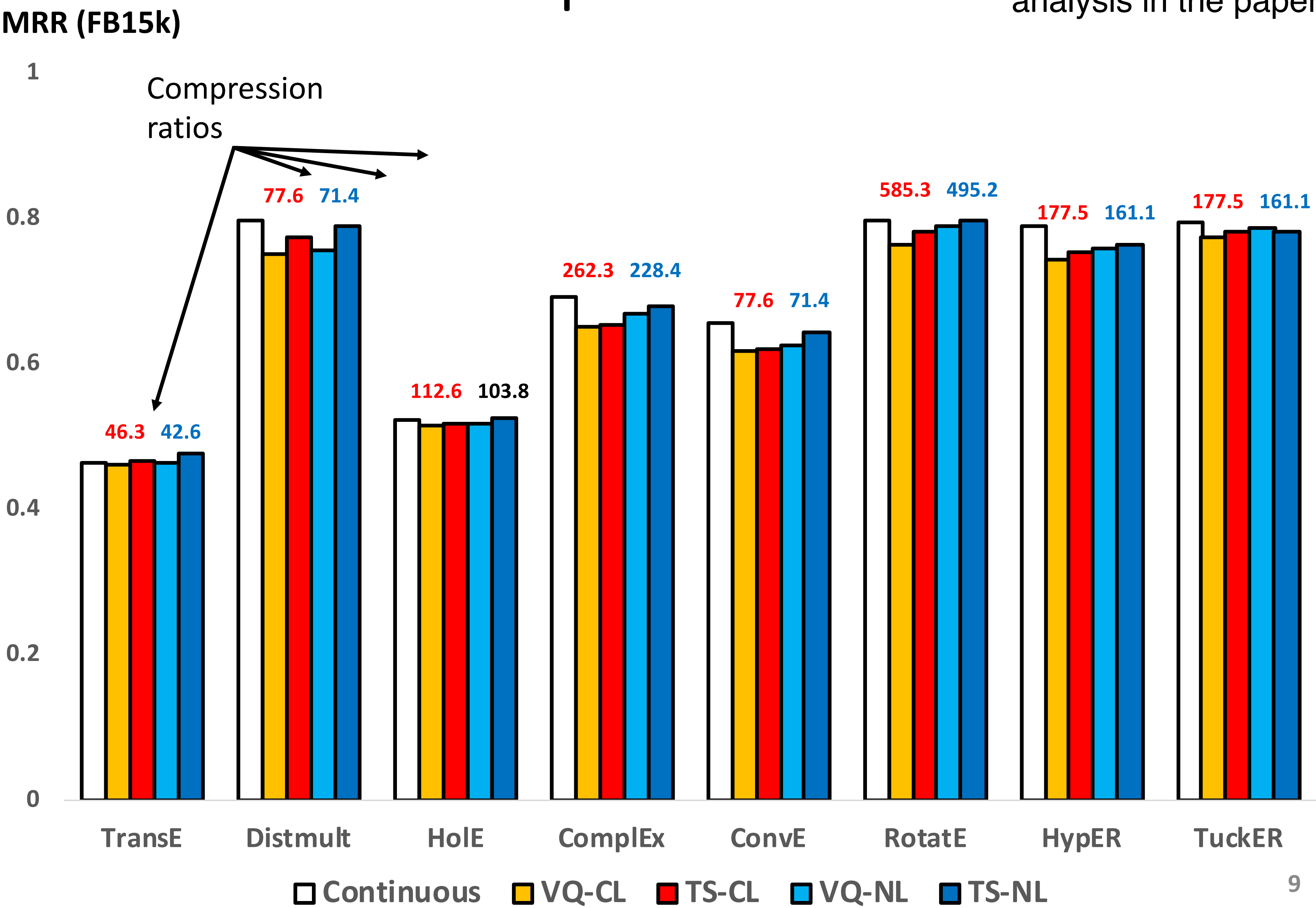
Metrics:

- Mean reciprocal rank (MRR)
- Hits@10 (H@10).

Compression Ratio (CR) = $\frac{storage(continuous)}{storage(discrete)}$

Experiments

*** More results and analysis in the paper



Thank you ☺