

Title: Generate the concept representation using OMOP ontology graph

INTRODUCTION

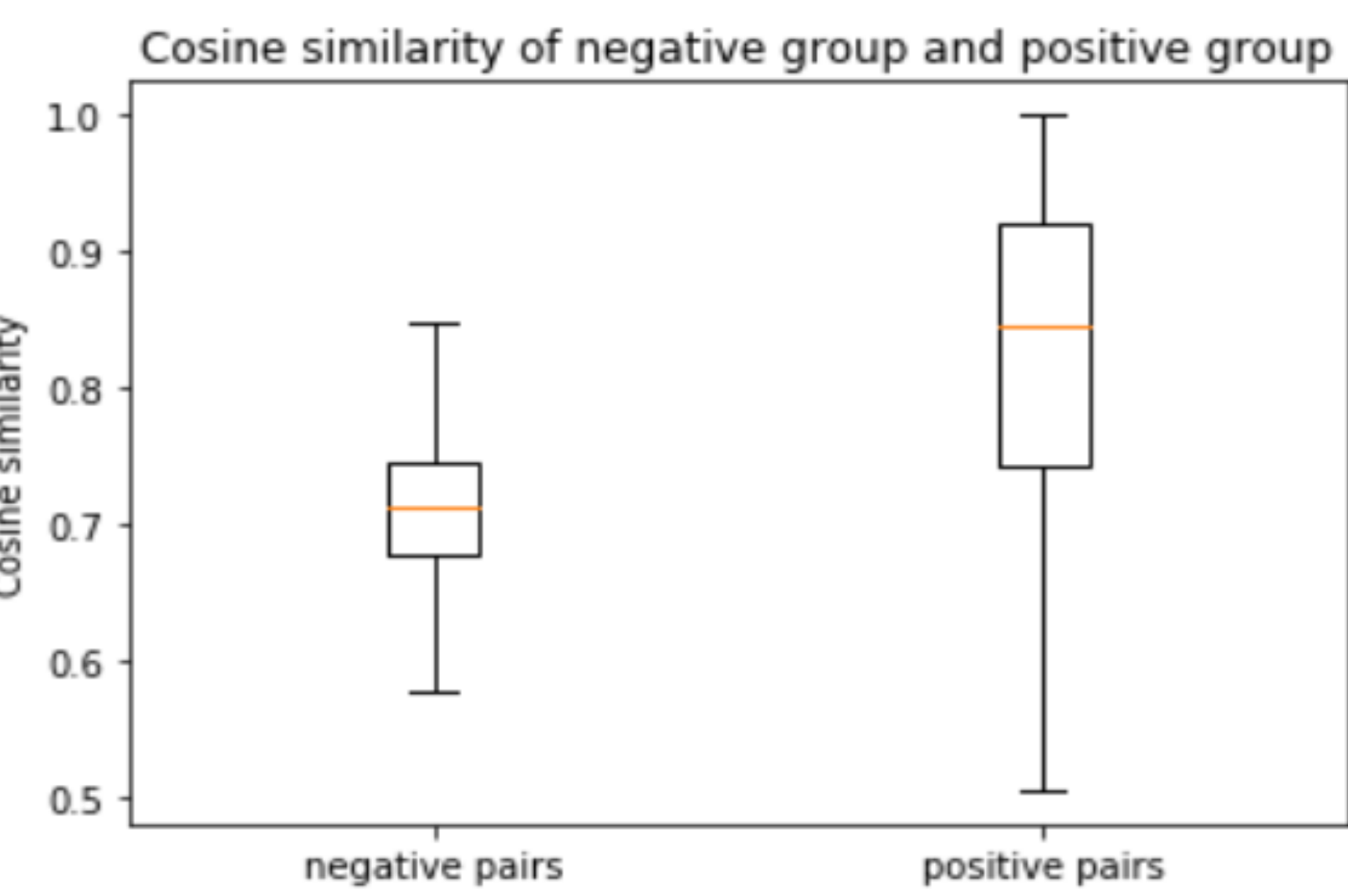
Learning distributed representation of OMOP concepts is powerful because

1. The representation can contain ontology information in OMOP CDM with low-dimensional vector
2. The representation can be applied to various machine learning tasks
3. The representation can be easily shared across distinct researchers and institutions

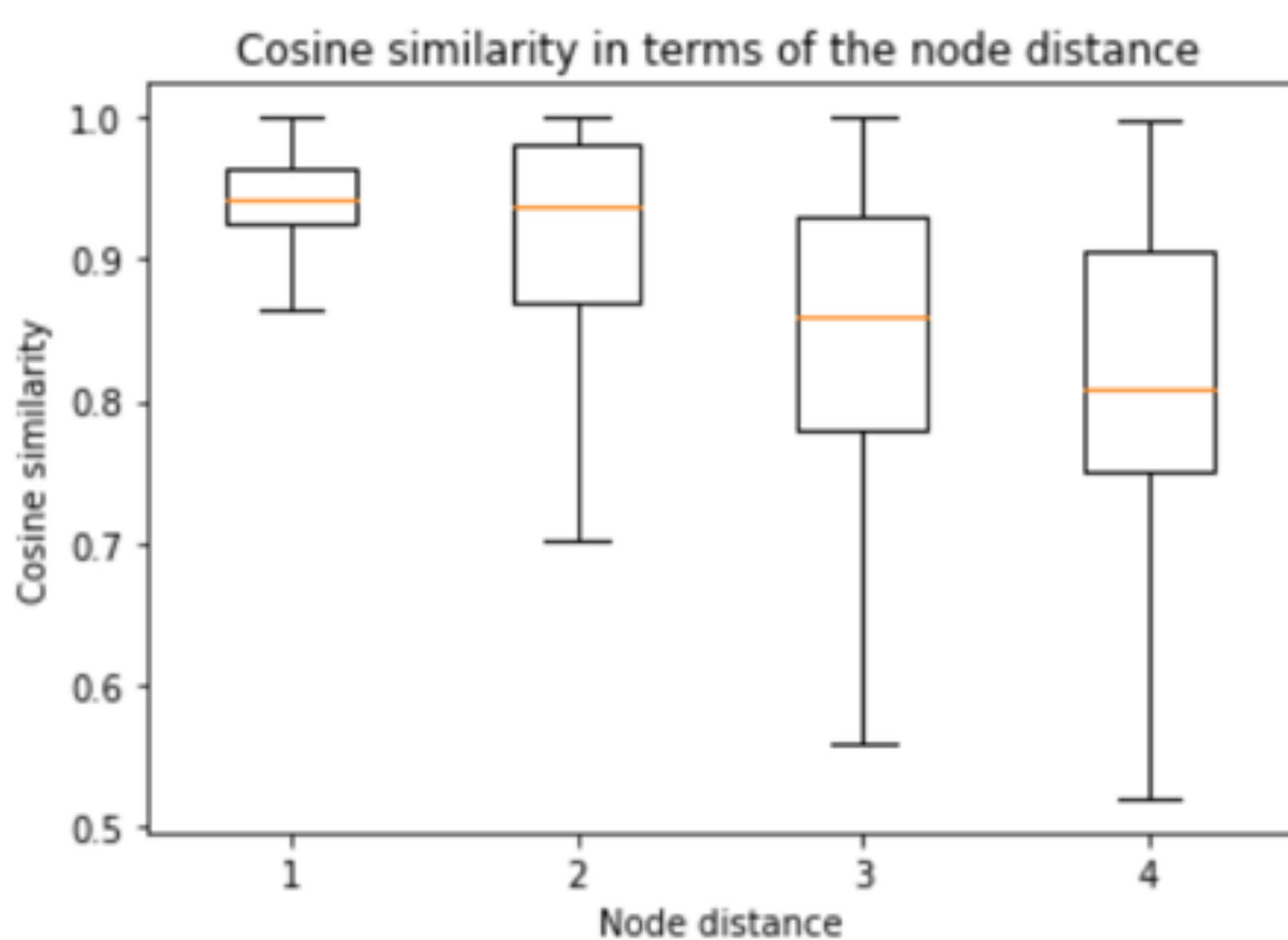
METHODS

1. Construct OMOP concept graph using concept_relationship table in the OMOP CDM
2. Learn distributed representation of the concept nodes using node2vec
3. Evaluate the representation using cosine similarity in terms of the three different metric: concept-set, node distance in the graph structure, and lexical similarity

RESULTS

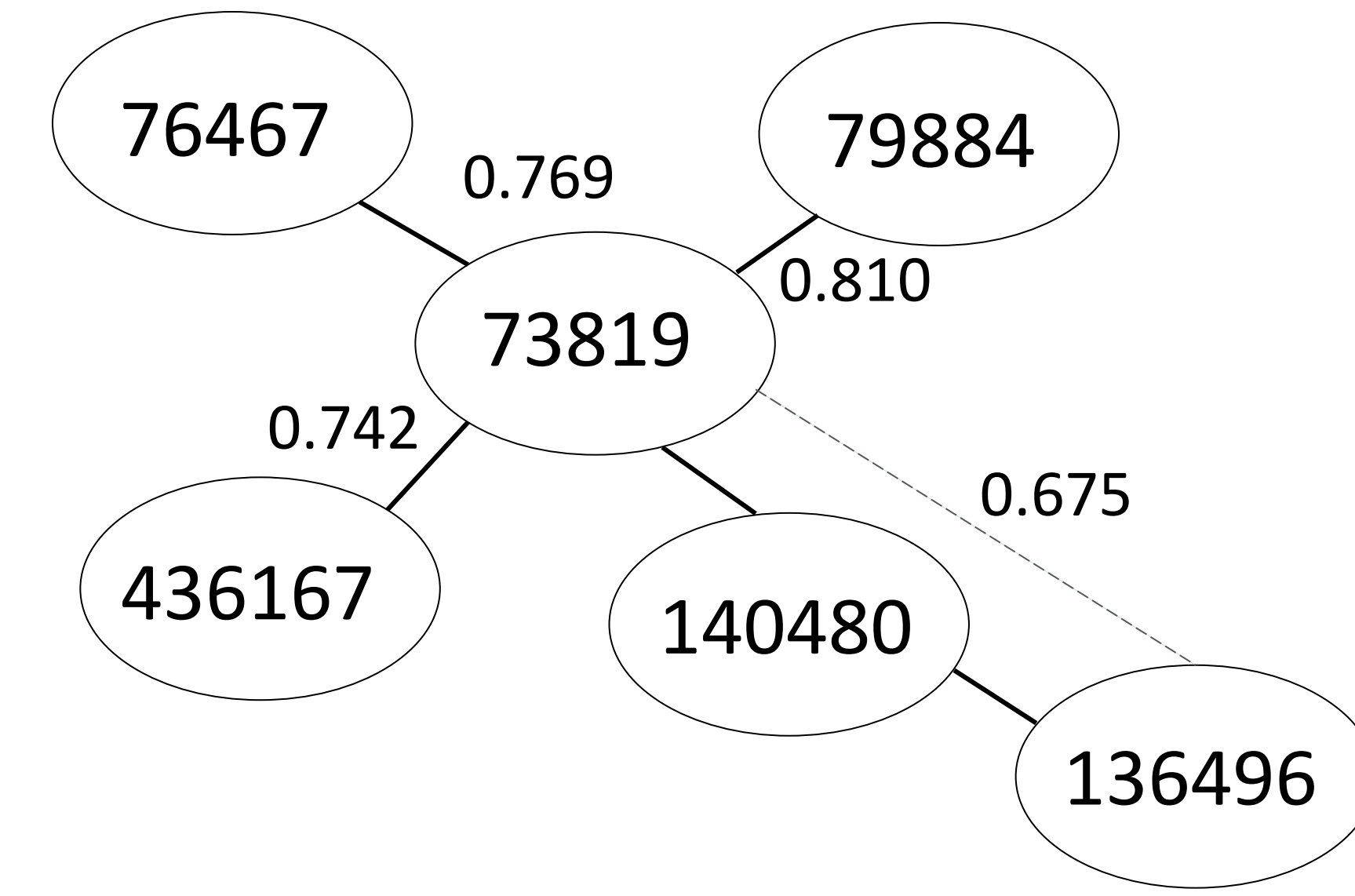


*Positive(negative) pair: two concepts (not)belonging to the same phenotype group

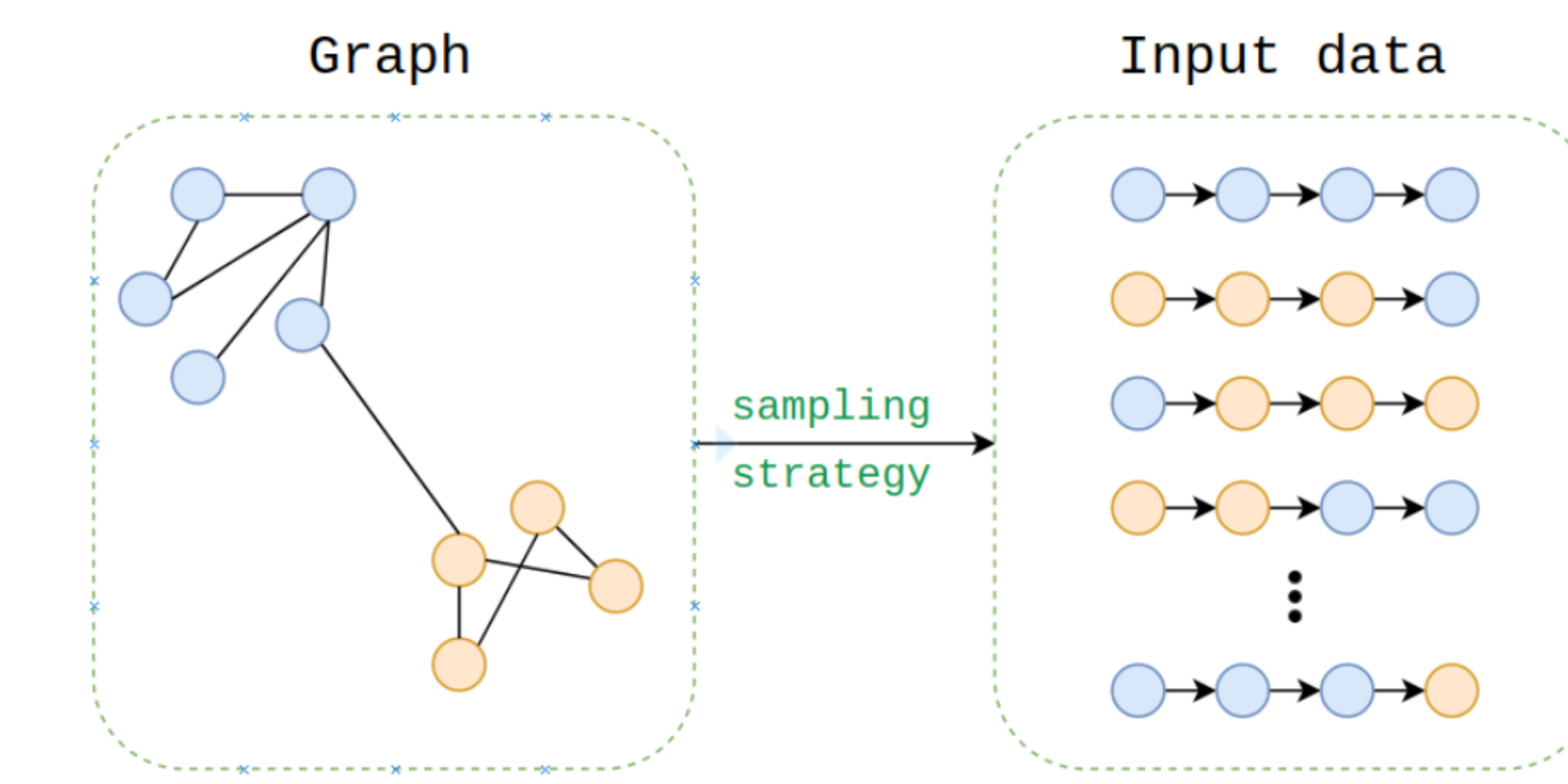
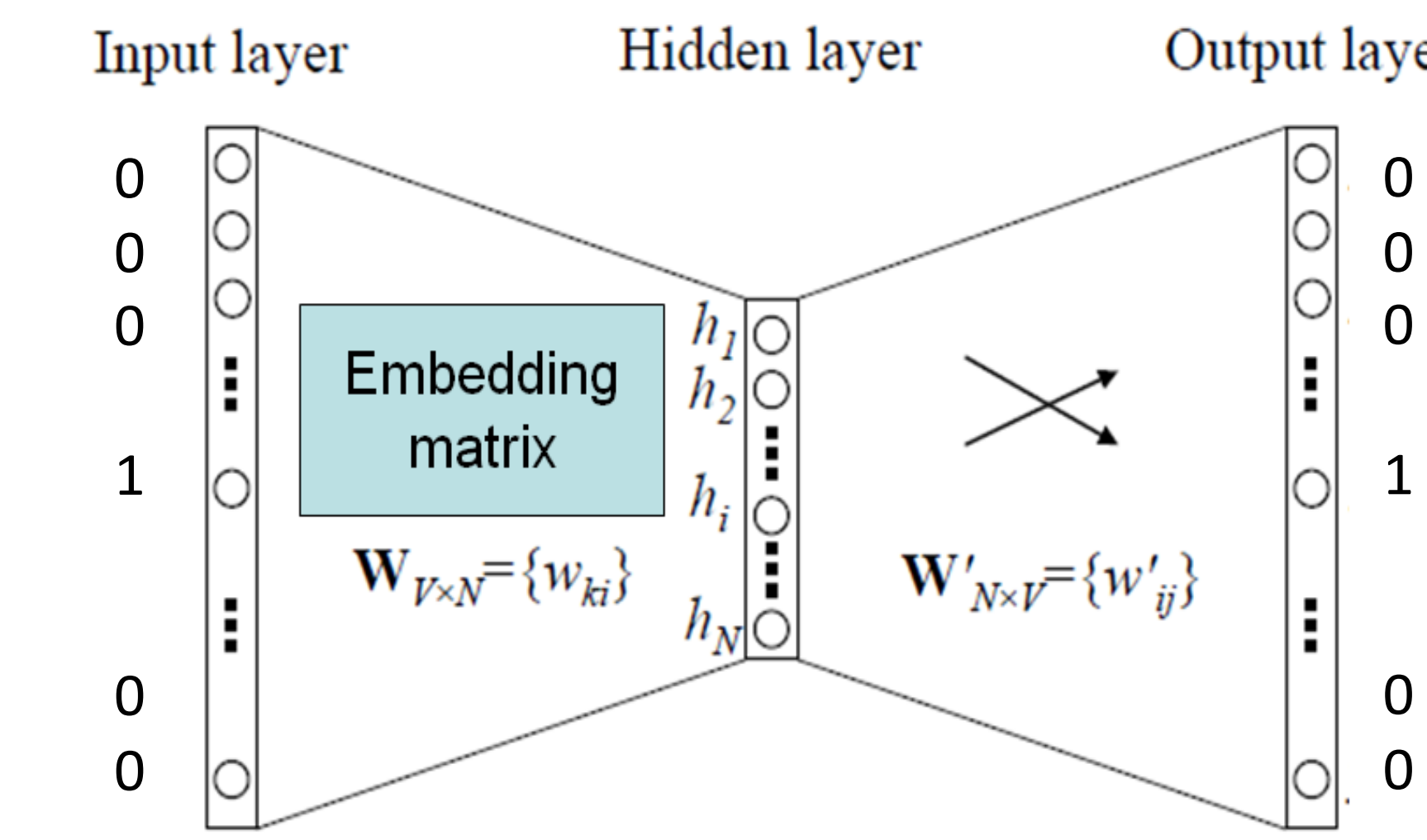


*Node distance: the minimum distance in the concept graph

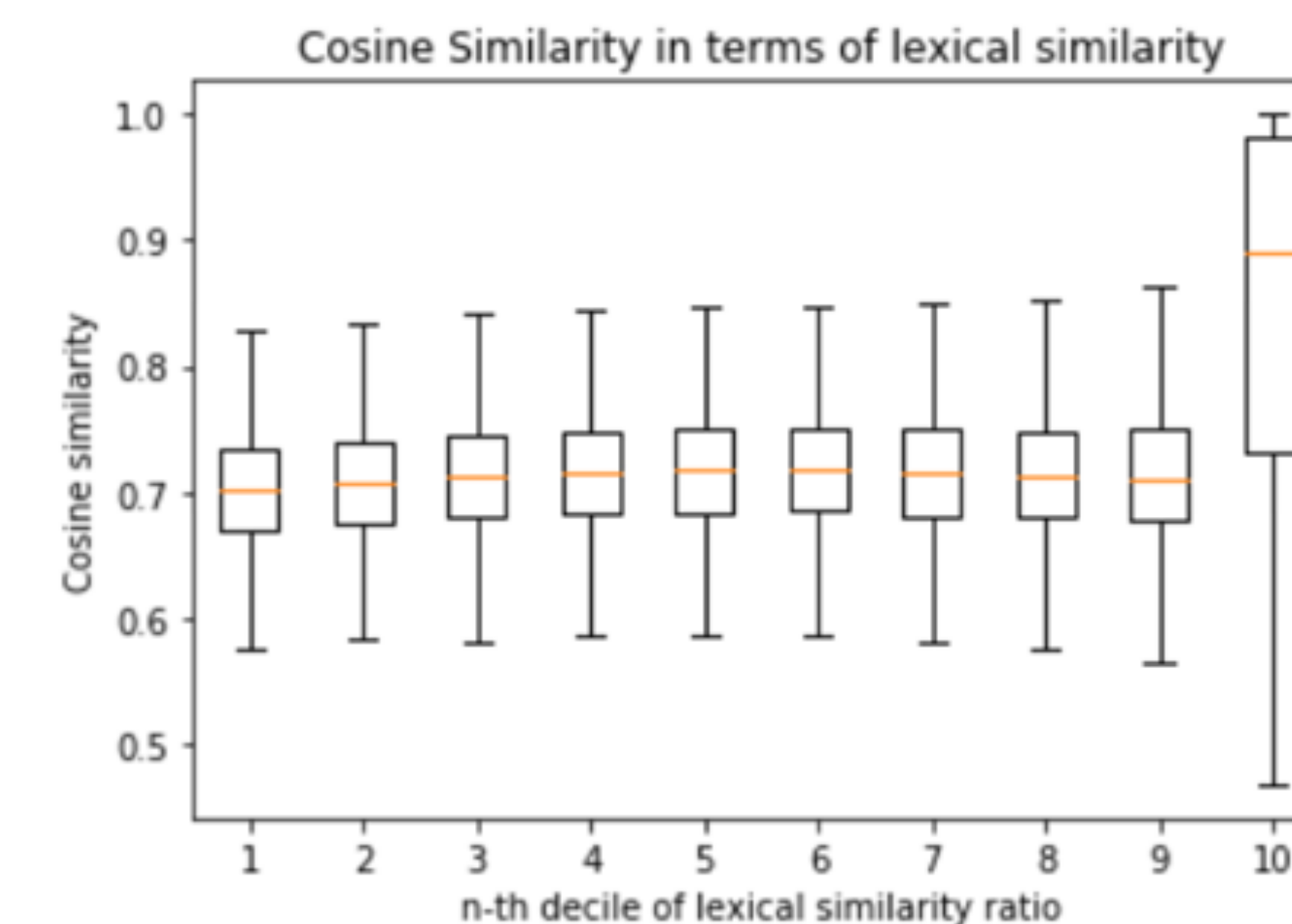
Vector representation of concepts in OMOP CDM containing information of their relationships.



Concept ID	Name
73819	Pain of breast
76467	Congenital hypoplasia of breast
79884	Galactorrhea not associated with childbirth
436167	Lactocele
140480	Impetigo
136496	Cellulitis and abscess of face



Concept ID	Name
Concept Nodes	2,349,171
Edges	15,705,202



Take a picture to download the full paper

PRESENTER: Jay(Junghwan) Lee

Junghwan Lee; Cong Liu, PhD
 Ning Shang, PhD;
 Xinzhuo Jiang, MS;
 Krishna Sai Dheeraj Kalluri, MS;
 Kai Chen; Chao Pang, PhD;
 Karthik Natarajan, PhD*;
 Patrick Ryan, PhD*;
 Chunhua Weng, PhD*