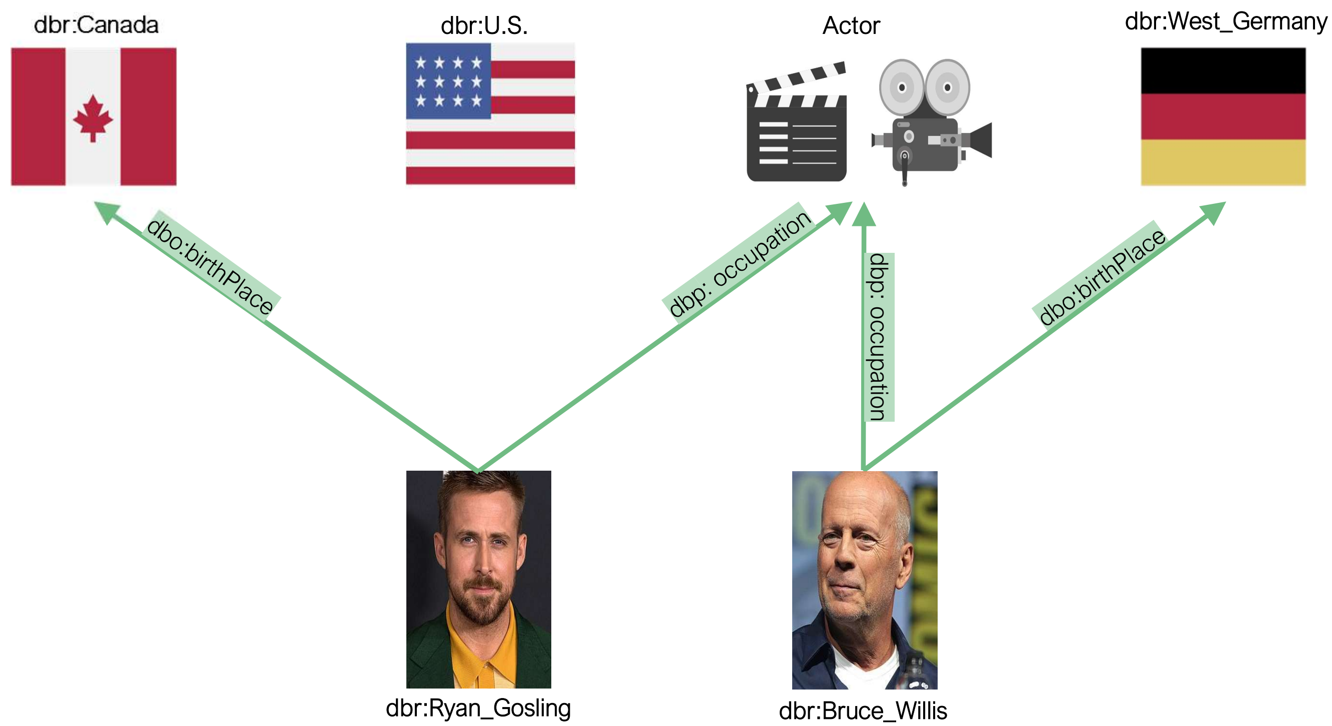
Improving Knowledge Graph Embeddings through Contrastive Learning with Negative Statements

Rita T. Sousa, Heiko Paulheim

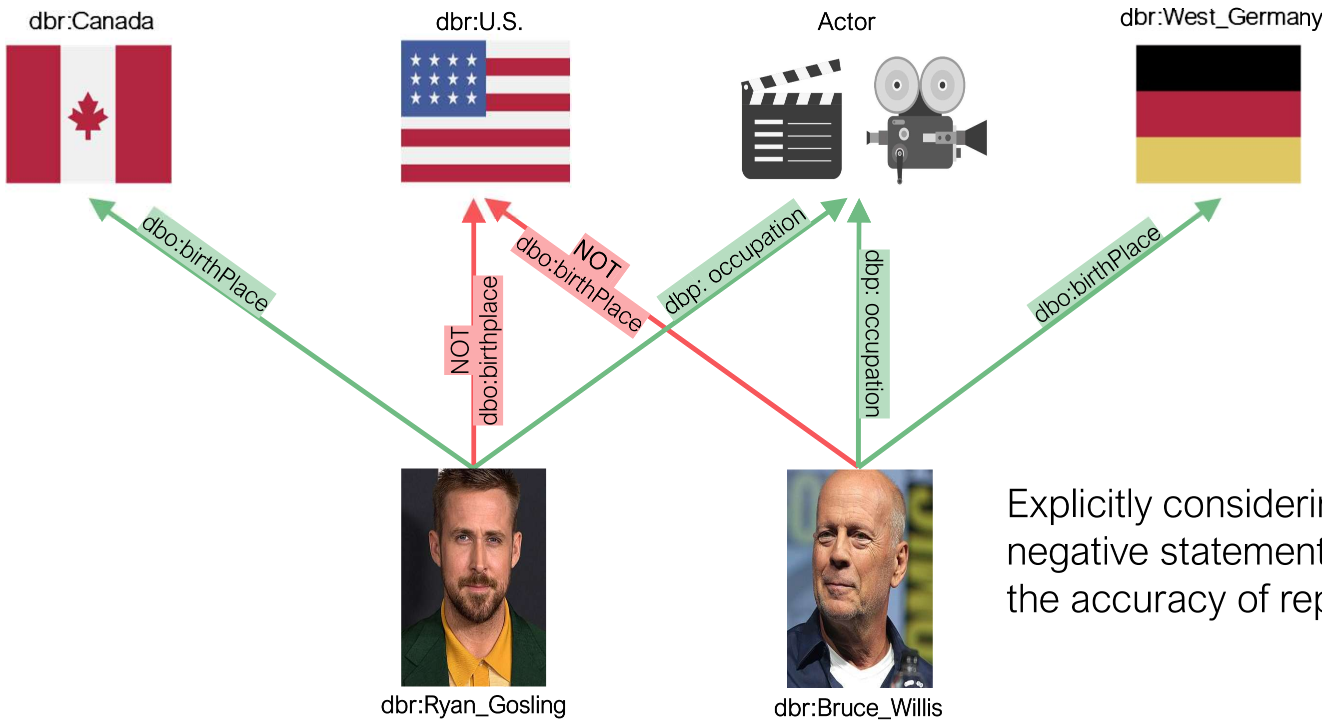
UNIVERSITY OF MANNHEIM

International Conference on Knowledge Capture (K-CAP)
December 10, 2025

The vast majority of knowledge graph (KG) relations are defined as positive statements.



However, negative statements can be defined.

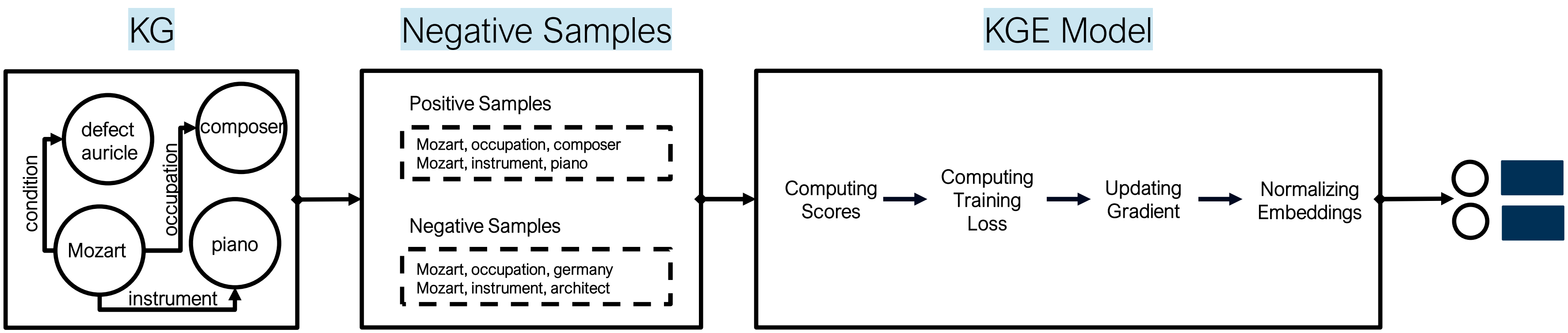


Explicitly considering interesting negative statements improves the accuracy of representations.

UNIVERSITY OF MANNHEIM

However, little attention has been given to the exploration of negative statements by KG embedding (KGE) approaches.

Training KGE models



Typically involves corrupting true triples by randomly replacing their head or tail entity with another entity from the KG.

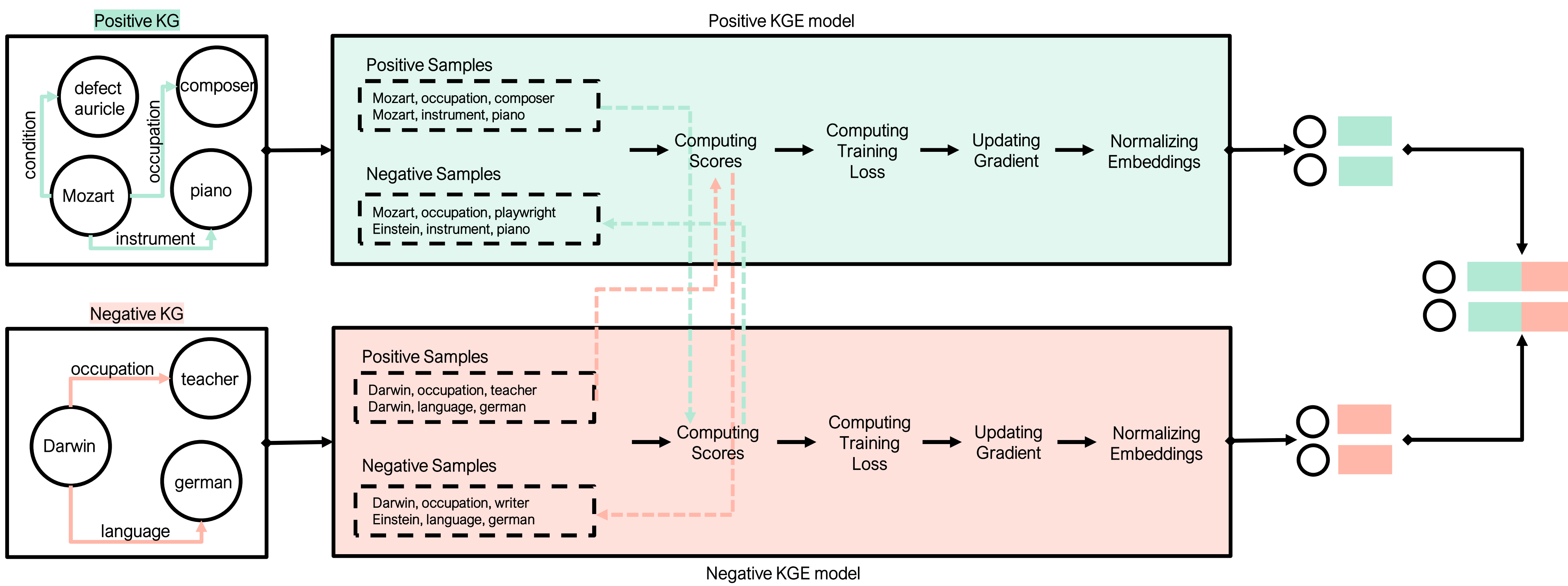
The goal is to assign higher scores to positive samples and lower scores to negative samples.

KGs are built under the Open World Assumption, while KGE models adopt the Closed World Assumption or Local Closed World assumption.

Methodology

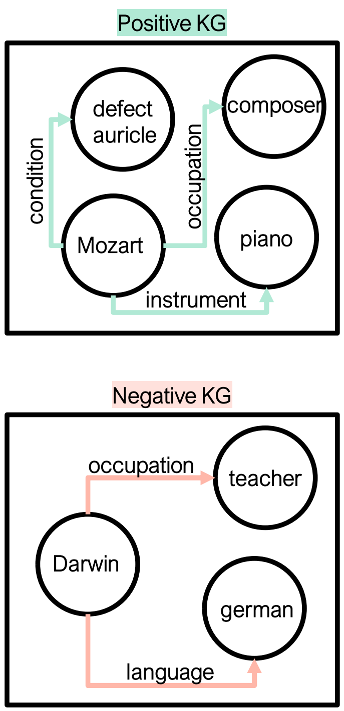
UNIVERSITY OF MANNHEIM

It is KGE-agnostic and can be integrated into any KGE model that defines a scoring function and employs negative sampling during training.



Methodology

Building the Positive and Negative KG



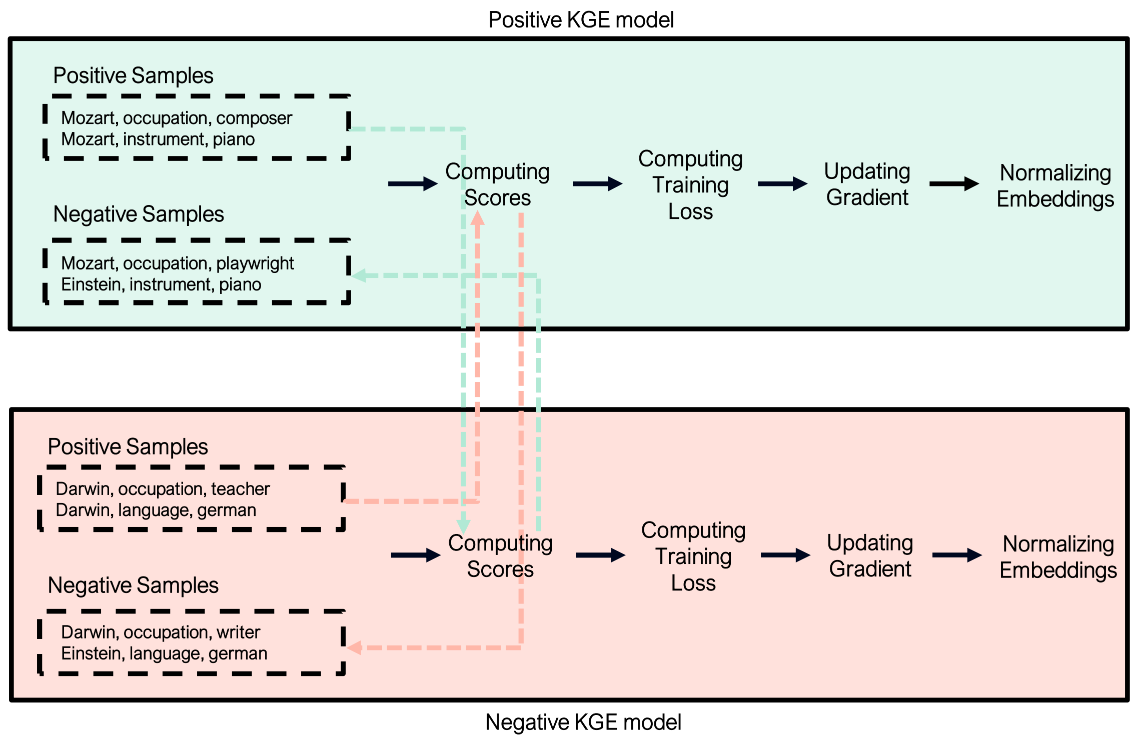
- Two separate RDF graphs are built: one from positive statements and another from explicitly defined negative statements.
- When the KG is backed by an ontology, the transformation follows the OWL to RDF Graph Mapping guidelines defined by W3C.

Methodology

Initializing KGE models & Contrastive Learning for Generating Negative Samples

Two-stage approach to negative sampling:

1. In the initial phase, the standard random corruption is used.
2. After cl_phase epochs a contrastive learning-based strategy is used.



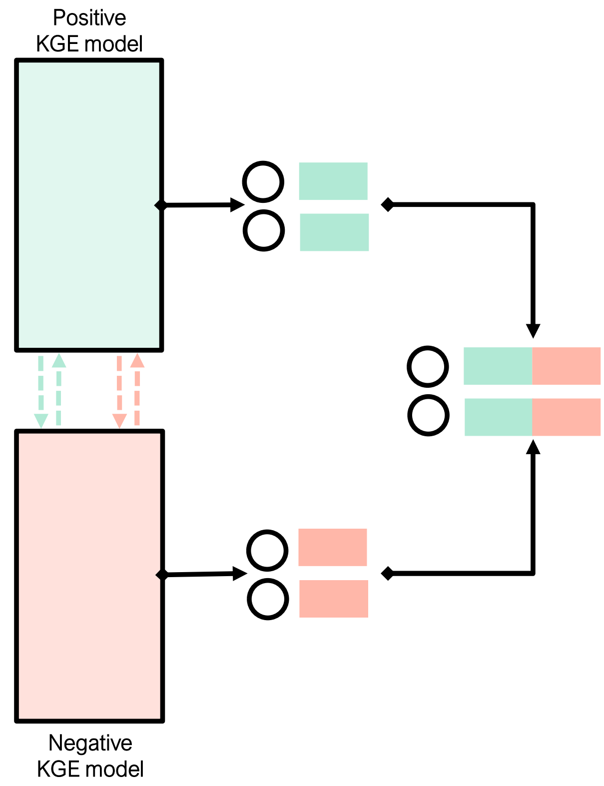
The contrastive learning strategy enables the dynamic generation of negative samples that are more meaningful and challenging as training progresses.

Methodology

Generating the Final Representation

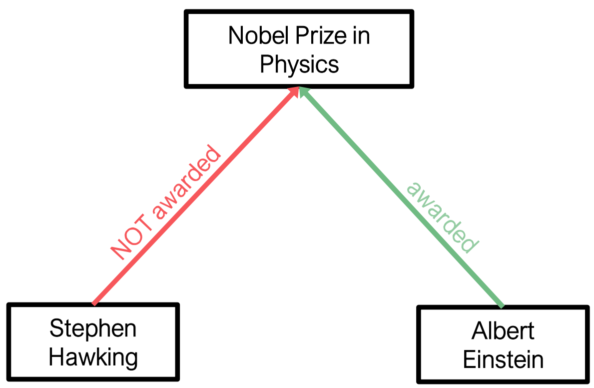
- Each node and relation in the KG is associated with two distinct representations: one learned from positive statements, and another learned from negative statements.

- To construct the final entity representation, the two embeddings using vector concatenation.

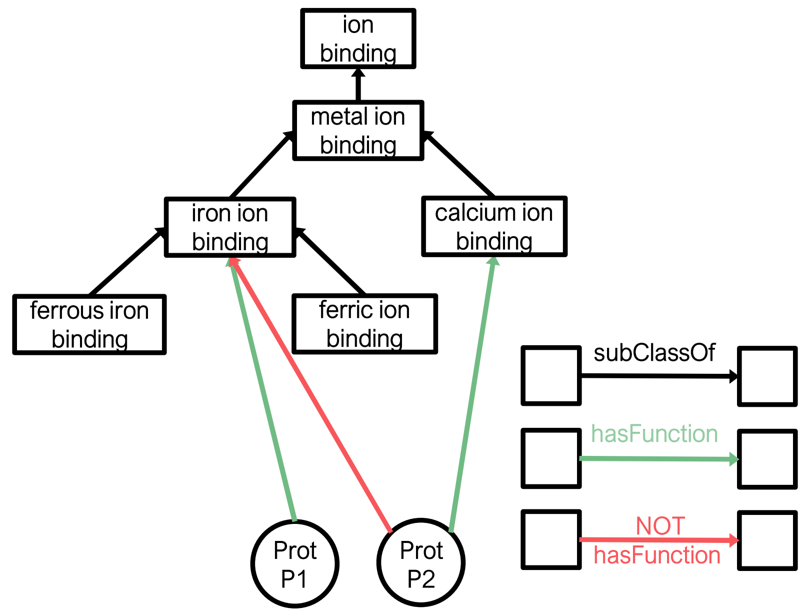


Evaluation
Data

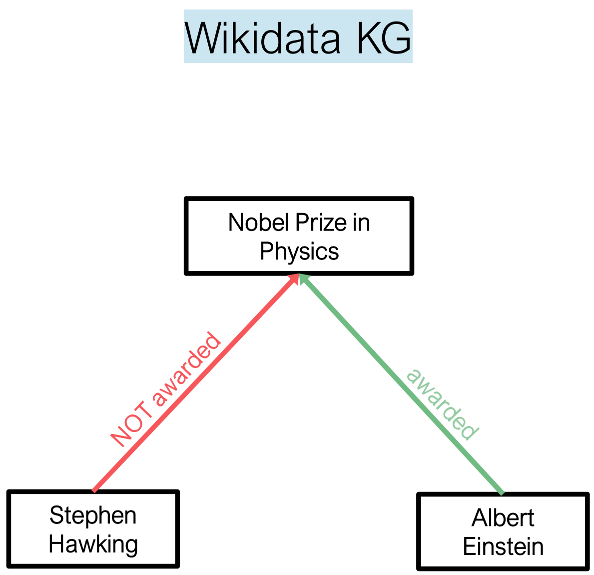
Wikidata KG



Gene Ontology (GO) KG



Evaluation
Data



Positive Statements: Wikidata is a vast collection of statements describing millions of entities.

Negative Statements: A statistical inference method called peer-based inference[1] is used.

Task: Link prediction.

[1] Hiba Arnaout, Simon Razniewski, Gerhard Weikum, and Jeff Z. Pan. 2021. Negative knowledge for open-world wikidata. In The Web Conference. 544–551.

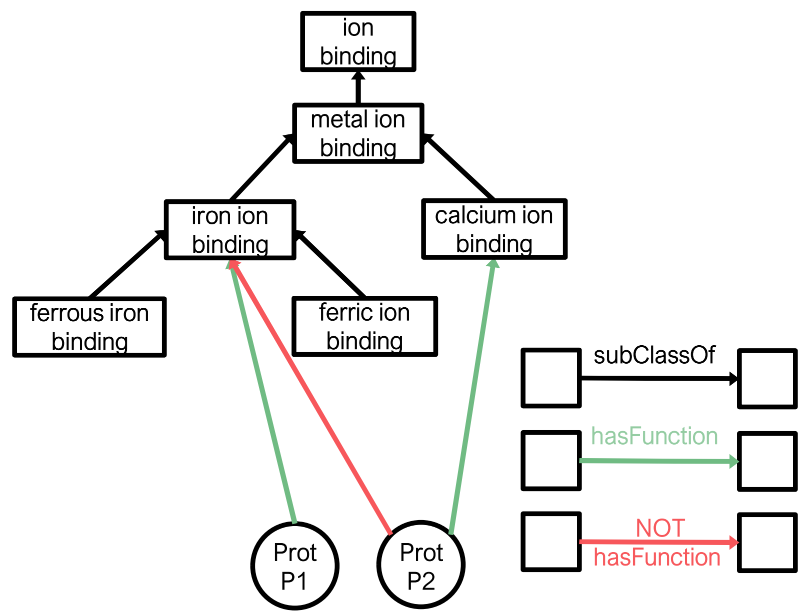
Evaluation
Data

Gene Ontology (GO) KG

Positive Statements: GO KG integrates both the GO itself and the GO annotation.

Negative Statements: The phylogenetic trees reveal where functions are lost over time[2] and can be recorded negative annotations.

Task: Triple classification.



[1] Hiba Arnaout, Simon Razniewski, Gerhard Weikum, and Jeff Z. Pan. 2021. Negative knowledge for open-world wikidata. In The Web Conference. 544–551.
[2] Alex Warwick Vesztrocy and Christophe Dessimoz. 2020. Benchmarking Gene Ontology function predictions using negative annotations. Bioinformatics 36, Supplement_1 (07 2020), i210–i218.

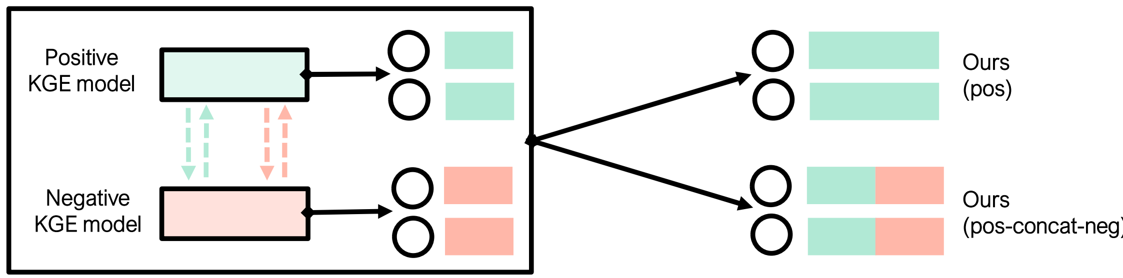
Evaluation
Set-up & Baselines

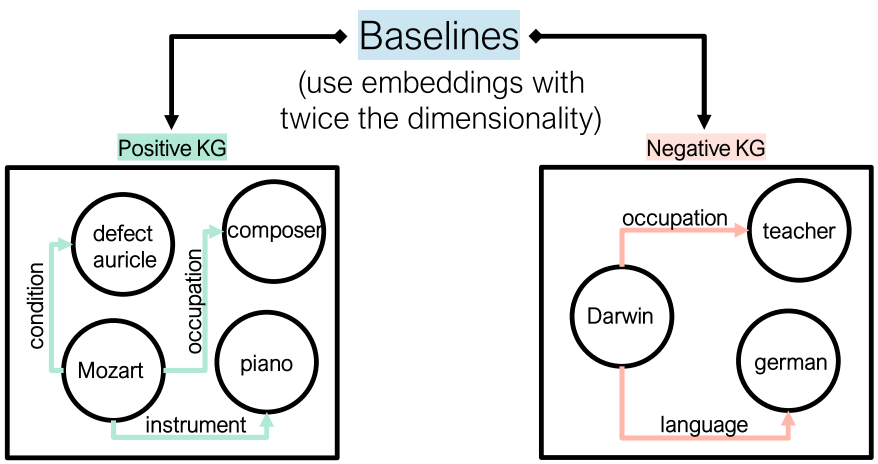
UNIVERSITY OF MANNHEIM

3 KGE models

KGE	Scoring Function
TransE	$f(h,r,t) = -\|h + r - t\|_p$
DistMult	$f(h,r,t) = \sum_i h_i \cdot r_i \cdot t_i$
ComplEx	$f(h,r,t) = \mathrm{Re}\left(\sum_i h_i \cdot r_i \cdot t_i\right)$

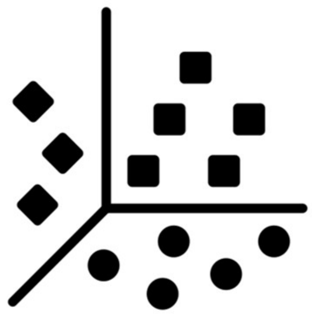
2 variations of the proposed approach



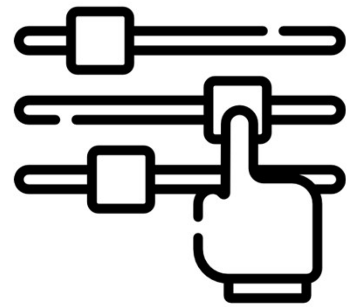


Results

Report performance results for link prediction on Wikidata and triple classification on GO KG.



The quality of the embeddings on both KGs by applying clustering metrics.



Ablation studies to analyze the impact of cl_phase and embedding size.

Results

Link Prediction on Wikidata

Inferring a missing entity in a triple, such as predicting the head entity or the tail entity.
Any corrupted triples that appear in the training set are excluded.

	KGE Model	MRR	Hits@10	Hits@1
TransE	pos	10.06%	18.20%	5.52%
	neg	9.34%	17.28%	5.08%
	ours (pos)	10.15%	18.56%	5.52%
	ours (pos-concat-neg)	10.40%	19.32%	5.36%
DistMult	pos	6.86%	12.64%	3.76%
	neg	4.05%	7.48%	2.24%
	ours (pos)	7.90%	17.52%	3.40%
	ours (pos-concat-neg)	9.63%	20.56%	4.64%
ComplEx	pos	4.66%	10.32%	1.84%
	neg	4.77%	8.60%	2.96%
	ours (pos)	8.86%	19.12%	4.16%
	ours (pos-concat-neg)	10.07%	20.80%	4.84%

- Our approach consistently outperforms the baselines for MRR and Hits@10.
- Using the negative KG yields competitive results.

Results

Triple Classification on Gene Ontology

The relation between two proteins is predicted as a binary classification task by combining their embeddings with the Hadamard product and training a Random Forest using 5-fold CV.

	KGE Model	Pr	Re	F1	AUC
TransE	pos	58.85%	58.68%	58.51%	62.80%
	neg	64.07%	64.06%	64.05%	68.49%
	ours (pos)	60.98%	60.98%	60.97%	64.74%
	ours (pos-concat-neg)	**67.34%**	**67.32%**	**67.31%**	**74.13%**
DistMult	pos	81.67%	81.66%	81.66%	88.91%
	neg	**83.70%**	**82.89%**	**82.78%**	90.58%
	ours (pos)	77.27%	77.07%	77.03%	86.03%
	ours (pos-concat-neg)	83.03%	82.68%	82.64%	**90.93%**
ComplEx	pos	79.34%	79.02%	78.94%	86.70%
	neg	81.45%	80.20%	79.99%	87.93%
	ours (pos)	78.97%	78.97%	78.97%	87.54%
	ours (pos-concat-neg)	**81.91%**	**81.91%**	**81.91%**	**89.55%**

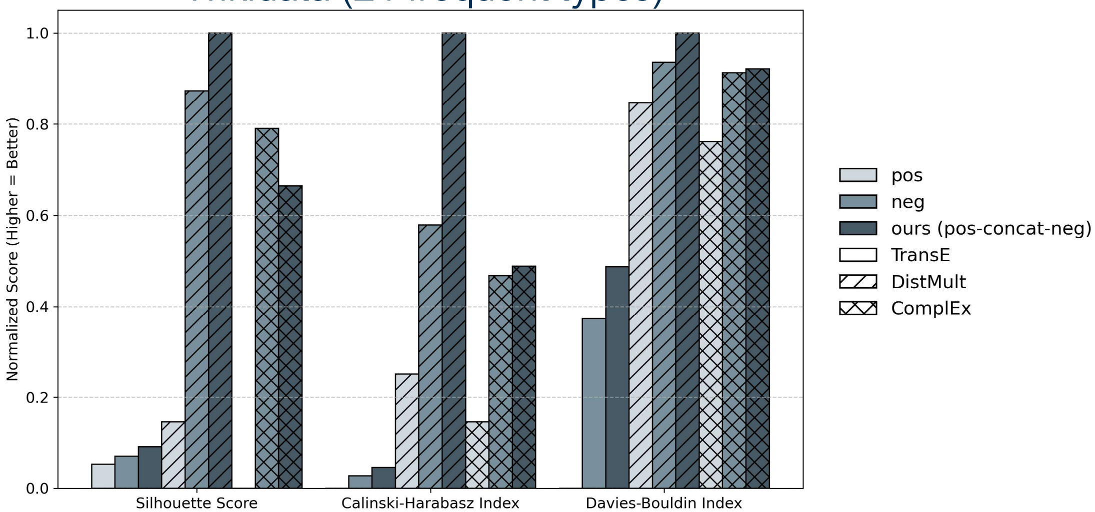
- Except for DistMult, it leads to improvements across all performance metrics.
- Weaker KGE models, such as TransE, benefit the most from our approach.

Results

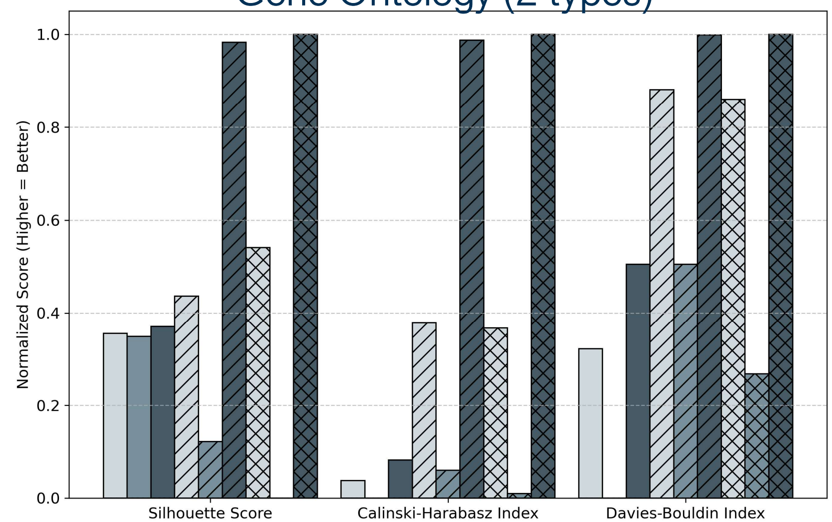
Embeddings Evaluation using Clustering Metrics

Provides a proxy for how well the representations reflect underlying semantic distinctions.

Wikidata (21 frequent types)



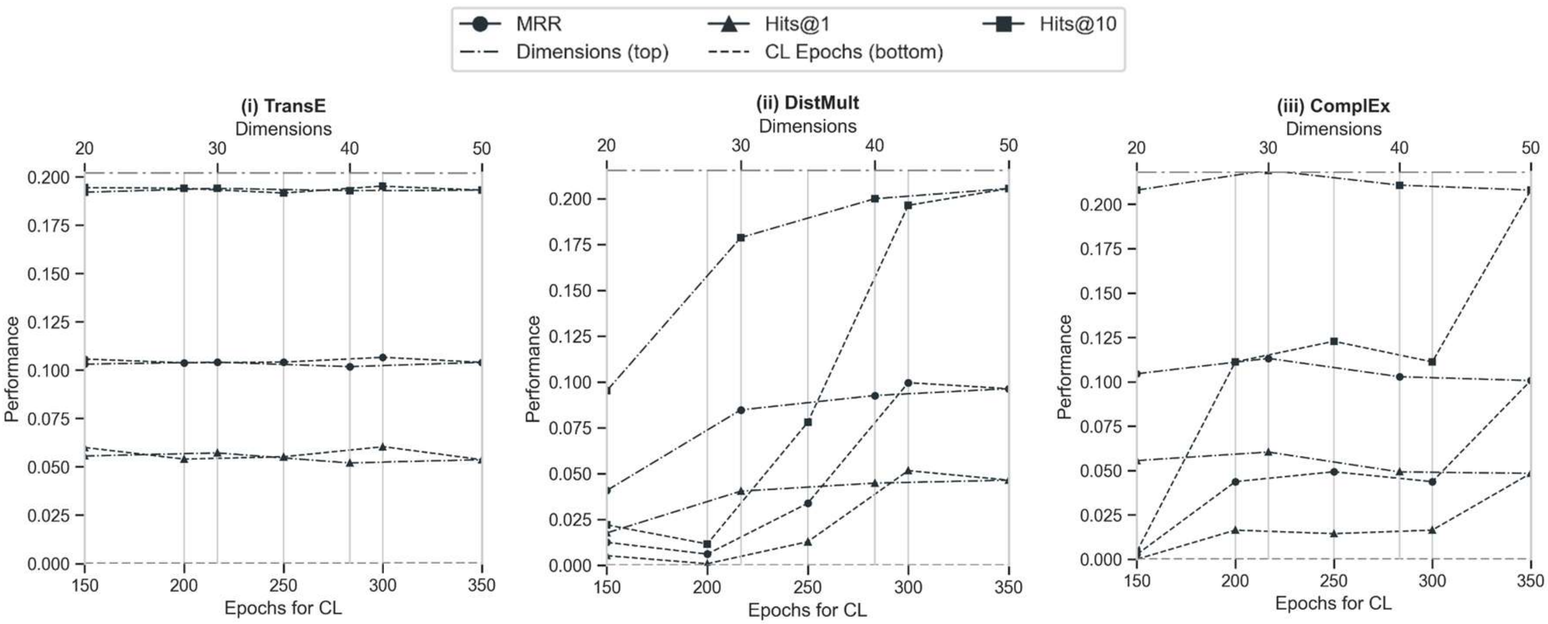
Gene Ontology (2 types)



- Our approach demonstrates superior clustering performance.
- For Wikidata KG, the baseline for negative KG outperforms the baseline for positive KG, likely due to differing relation distributions across entity types.

Results

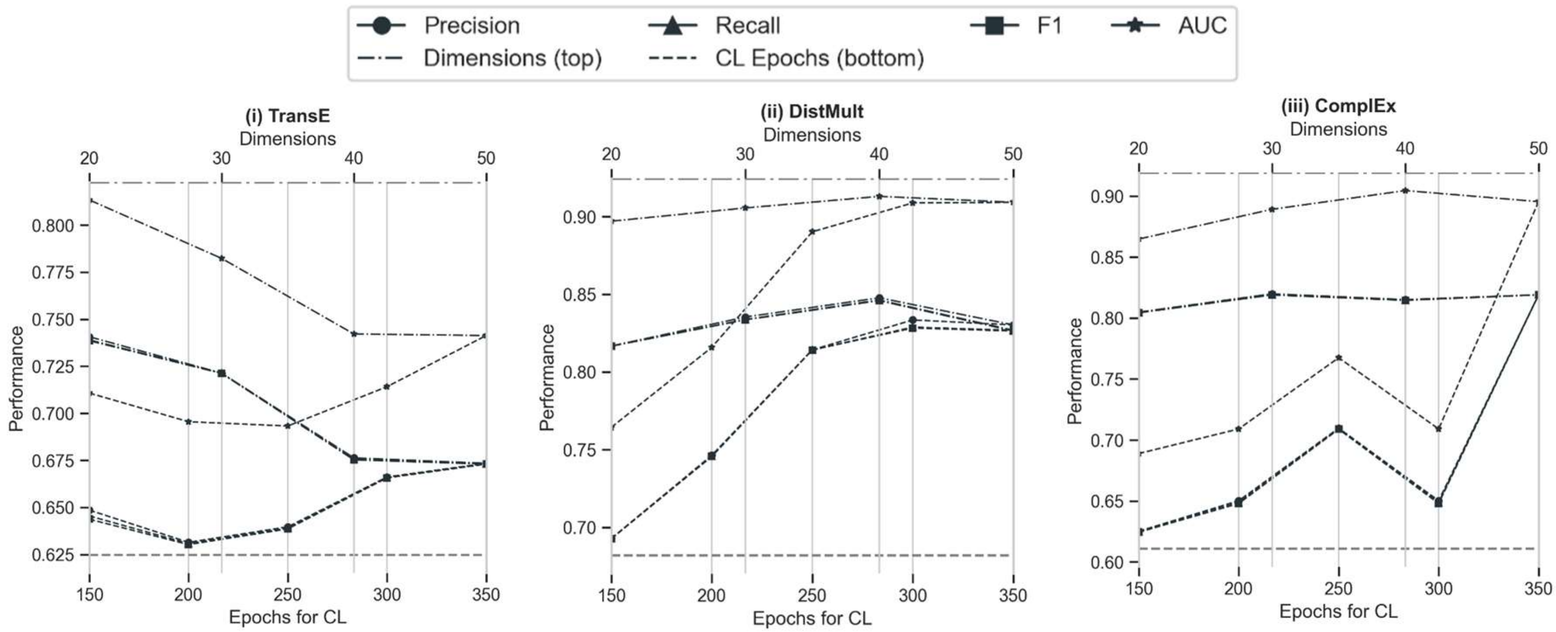
Ablation Studies for Wikidata



- Delaying the transition to the contrastive learning strategy appears be beneficial.
- Our approach demonstrates robustness across different embedding sizes.

Results

Ablation Studies for Gene Ontology



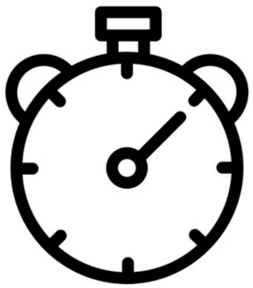
- Delaying the transition to the contrastive learning strategy appears be beneficial.
- Our approach demonstrates robustness across different embedding sizes.

Conclusions

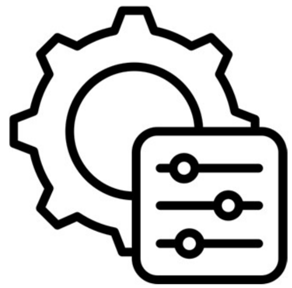
UNIVERSITY OF MANNHEIM

- Current KGE models cannot effectively handle explicitly defined negative statements.
- Our approach combines dual-model training with an adapted negative sampling mechanism grounded in contrastive learning.
- Our approach outperforms state-of-the-art KGE models on two KGs and tasks.
- Our approach can be easily incorporated into any scoring-based KGE model for any KG and task.

Future Work



Efficient strategy to generate candidate negative samples in contrastive learning.



Systematic hyperparameter optimization.

Implementation of an asynchronous contrastive training strategy.

Thank you for your attention!

rita.sousa@uni-mannheim.de

https://ritatsousa.github.io/

Additional Slides

Additional Results

Link Prediction on Wikidata

Semantic awareness, sem@k, measures the proportion of triples that the predicted entity (head or tail) belongs to the same type as the corresponding entity in the ground-truth triple.

	KGE Model	Head	Tail	Average
TransE	pos	53.12%	**62.48%**	57.80%
	neg	20.96%	58.72%	39.84%
	ours (pos)	**63.36%**	57.84%	**60.60%**
	ours (pos-concat-neg)	53.04%	62.40%	57.72%
DistMult	pos	62.16%	40.00%	51.08%
	neg	25.28%	46.64%	35.96%
	ours (pos)	71.84%	46.48%	59.16%
	ours (pos-concat-neg)	**74.40%**	**58.96%**	**66.68%**
ComplEx	pos	76.72%	34.08%	55.40%
	neg	37.44%	**63.20%**	50.32%
	ours (pos)	**80.24%**	39.20%	59.72%
	ours (pos-concat-neg)	74.88%	57.60%	**66.24%**

- Our approach not only improves ranking metrics but also leads to more semantically plausible predictions.

Additional Results

Random Negative statements

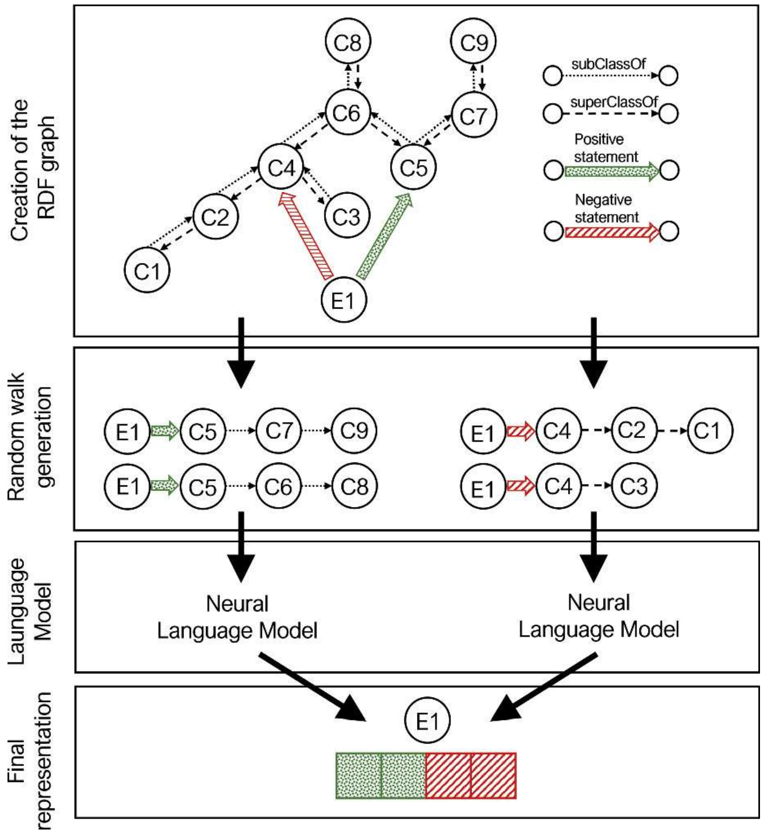
Wikinegata

	KGE Model	MRR	Hits@10	Hits@1
TransE	pos	10.06%	18.20%	**5.52%**
	neg	9.34%	17.28%	5.08%
	neg random	0.02%	0.00%	0.00%
	ours (pos)	10.15%	18.56%	**5.52%**
	ours (pos) with random	11.03%	20.62%	6.08%
	ours (pos-concat-neg)	**10.40%**	**19.32%**	5.36%
	ours (pos-concat-neg) with random	9.56%	17.46%	5.27%
DistMult	pos	6.86%	12.64%	3.76%
	neg	4.05%	7.48%	2.24%
	neg random	0.01%	0.00%	0.00%
	ours (pos)	7.90%	17.52%	3.40%
	ours (pos) with random	8.63%	16.05%	4.78%
	ours (pos-concat-neg)	**9.63%**	**20.56%**	**4.64%**
	ours (pos-concat-neg) with random	1.38%	2.96%	0.61%
ComplEx	pos	4.66%	10.32%	1.84%
	neg	4.77%	8.60%	2.96%
	neg random	0.02%	0.00%	0.00%
	ours (pos)	8.86%	19.12%	4.16%
	ours (pos) with random	**9.27%**	**19.17%**	**4.78%**
	ours (pos-concat-neg)	**10.07%**	**20.80%**	**4.84%**
	ours (pos-concat-neg) with random	2.12%	4.05%	1.05%

Gene Ontology

	KGE Model	Pr	Re	F1	AUC
TransE	pos	58.85%	58.68%	58.51%	62.80%
	neg	64.07%	64.06%	64.05%	68.49%
	neg with random	53.67%	53.66%	53.63%	55.48%
	ours (pos)	60.98%	60.98%	60.97%	64.74%
	ours (pos) with random	57.82%	57.81%	57.79%	61.83%
	ours (pos-concat-neg)	**67.34%**	**67.32%**	**67.31%**	**74.13%**
	ours (pos-concat-neg) with random	58.44%	58.44%	58.44%	61.78%
DistMult	pos	81.67%	81.66%	81.66%	88.91%
	neg	**83.70%**	**82.89%**	**82.78%**	90.58%
	neg random	54.39%	54.39%	54.39%	55.51%
	ours (pos)	77.27%	77.07%	77.03%	86.03%
	ours (pos) with random	77.63%	77.56%	77.55%	85.50%
	ours (pos-concat-neg)	83.03%	82.68%	82.64%	**90.93%**
	ours (pos-concat-neg) with random	77.11%	77.02%	77.00%	83.48%
ComplEx	pos	79.34%	79.02%	78.94%	86.70%
	neg	81.45%	80.20%	79.99%	87.93%
	neg random	53.06%	53.06%	53.06%	54.90%
	ours (pos)	78.97%	78.97%	78.97%	87.54%
	ours (pos) with random	77.96%	77.81%	77.78%	85.00%
	ours (pos-concat-neg)	**81.91%**	**81.91%**	**81.91%**	**89.55%**
	ours (pos-concat-neg) with random	76.28%	76.28%	76.28%	83.83%

Related Work
TrueWalks



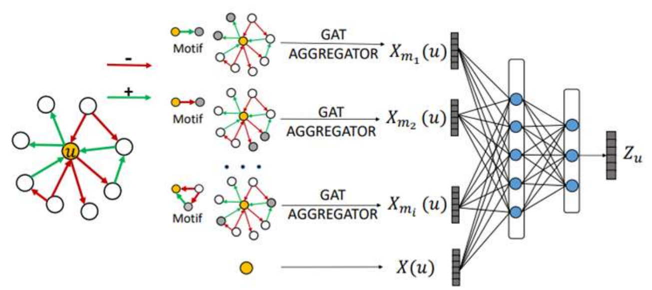
Rita T Sousa, Sara Silva, Heiko Paulheim, and Catia Pesquita. 2023. Biomedical knowledge graph embeddings with negative statements. In International Semantic Web Conference. Springer, 428–446.

TrueWalks generates two distinct embeddings for each entity: one capturing the positive semantics and another capturing the negative semantics.

- For the positive embedding, it generates walks based on positive and *subClassOf* relationships
- For the negative embedding, it generates walks using negative and *superClassOf* relationships

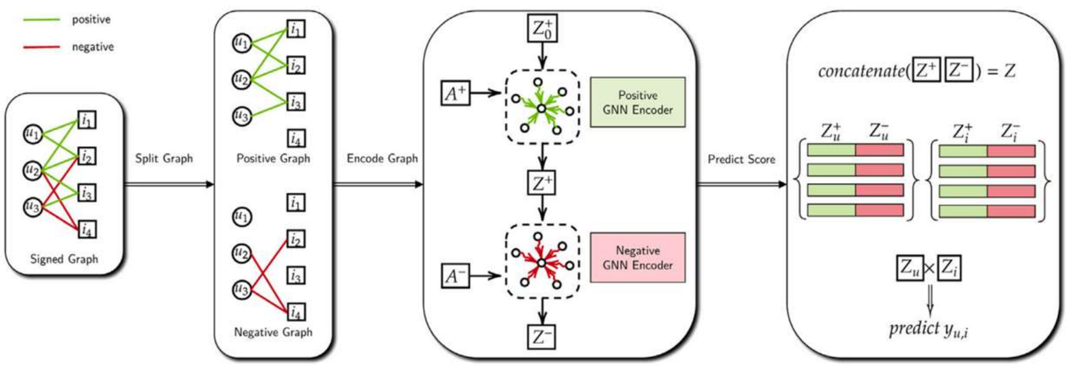
Related Work
GNN-based approaches

SiGAT uses social theories (balance and status theory) to categorize neighbor nodes into motifs. Then, it applies a GAT-based aggregation to combine information from those categorized neighbors.



Junjie Huang, Huawei Shen, Liang Hou, and Xueqi Cheng. 2019. Signed graph attention networks. In International Conference on Artificial Neural Networks. Springer, 566–577.

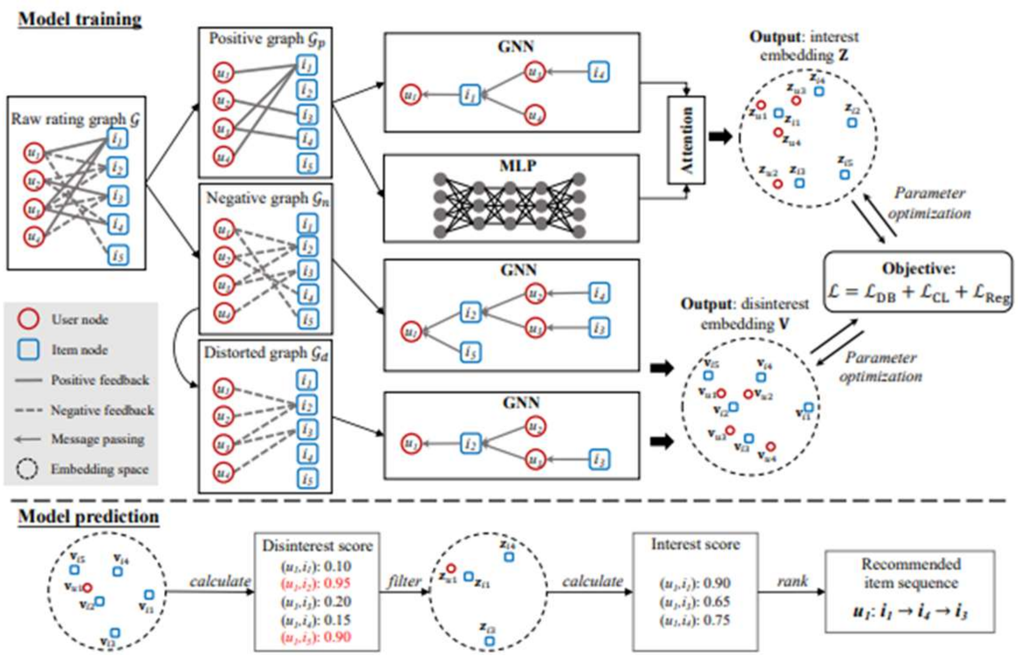
SiGRec creates two separate embeddings per node (positive and negative) and combines them by concatenation. It also introduces the sign cosine loss, a loss function designed to handle various types of negative feedback.



Junjie Huang, Ruobing Xie, Qi Cao, Huawei Shen, Shaoliang Zhang, Feng Xia, and Xueqi Cheng. 2023. Negative can be positive: Signed graph neural networks for recommendation. Information processing and management 60, 4 (2023), 103403.

Related Work

GNN-based approaches



Ziyang Liu, Chaokun Wang, Jingcao Xu, Cheng Wu, Kai Zheng, Yang Song, Na Mou, and Kun Gai. 2023. PANE-GNN: Unifying positive and negative edges in graph neural networks for recommendation. arXiv:2306.04095 (2023).

PANE-GNN partitions the graph into two distinct bipartite graphs based on positive and negative feedback and then generates an interest embedding and a disinterest embedding with positive and negative edges. For the negative graph, a distortion is introduced to denoise the negative feedback.