



# Improving Convolutional Neural Networks for Knowledge Graph Completion

Kevin Cousof, Nilofar Moradi Farisar  
Waleed Ragheb, Mehdi Mirzapour

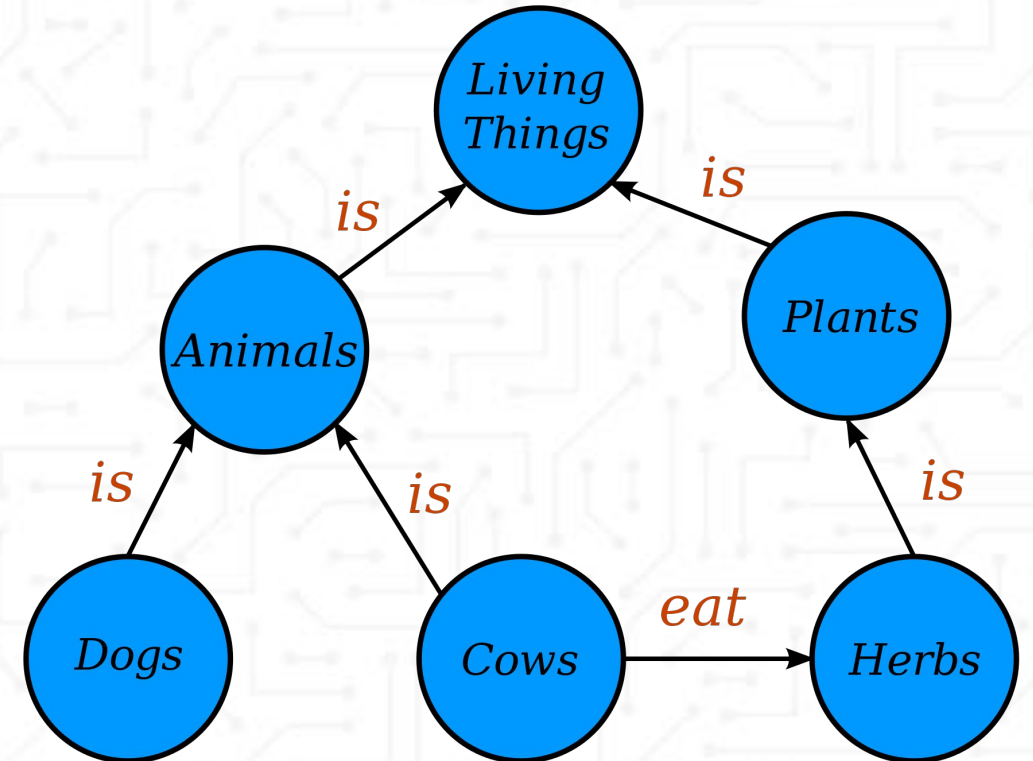
# Table of Contents

- 1. Introduction**
- 2. Related Work**
- 3. MConvKB**
- 4. Expérience**
- 5. Evaluation**
- 6. Conclusion et perspectives**

# Introduction

## Graphes de connaissances

- Structurent l'information sémantique sous forme de multi-graphes étiquetés
- Les sommets sont des entités, les arêtes des relations
- On notes les relations par des triplets
  - <Dogs, is, Animals>
  - <Cows, eat, Herbs>



# Introduction

## Graphes de connaissances

- Utile à de nombreuses fin telles que :
  - *Information Retrieval*
  - *Question Answering*
  - *Machine Translation*
- Quelques graphes de connaissances :
  - Freebase [10]
  - DBPedia [7]

# Introduction

## Prédiction de relations

- Les graphes de connaissances sont incomplets
- Besoin de méthodes pour ajouter l'information absente :

### *Knowledge Graph Completion*

- Classification de relations :  $\langle h, ?, t \rangle$
- Prédiction de relations :  $\langle h, r, ? \rangle$

# Related Work

## Modèles translationnels

- Idée centrale :
  - calculer des embeddings pour les entités et les relations
  - utiliser les embeddings de relations comme vecteur de translations
  - si  $\langle h, r, t \rangle$  existe alors  $h + r$  doit être proche de  $t$

## Réseaux de neurones

- cherchent à calculer des embeddings à partir de :
  - chemins et sous-graphes
  - informations structurelles

# MConvKB

## Architecture

- Extension de ConvKB [9]
  - couche de classification à la VGG [5]
  - expérimentation avec différents embeddings pré-entraînés TransE [2] et TransD [4]

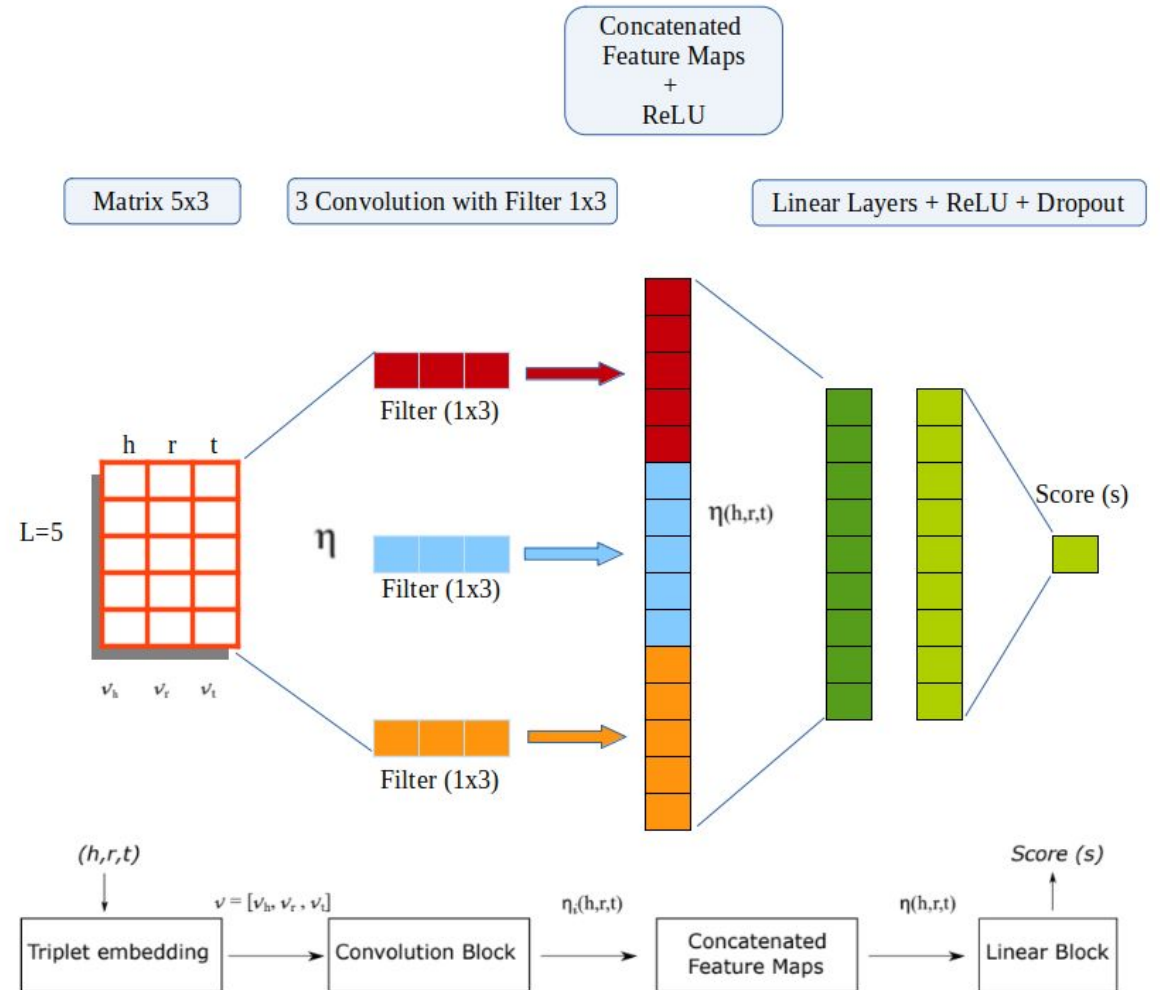
Model	Score function
TransE	$\ v_h + v_r - v_t\ _2^2$
TransD	$\ h_{\perp} + v_r - t_{\perp}\ _2^2$
ConvKB	$\text{concat}(g([v_h, v_r, v_t] * \Omega)) \cdot w$
MConvKB	$\text{concat}(g_1(\text{ReLU}(g_2(\text{ReLU}(g_3([v_h, v_r, v_t] * \Omega)))))) \cdot w$

TAB. 1 – Models score functions: with  $g_i$  as non-linear functions,  $\Omega$  as convolutional filters,  $w$  as weights;  $(*)$  and  $(\cdot)$  are convolutional and dot operators, respectively.

# MConvKB

## Architecture

1. Plongement des triplets
2. Extraction de features
3. Classification



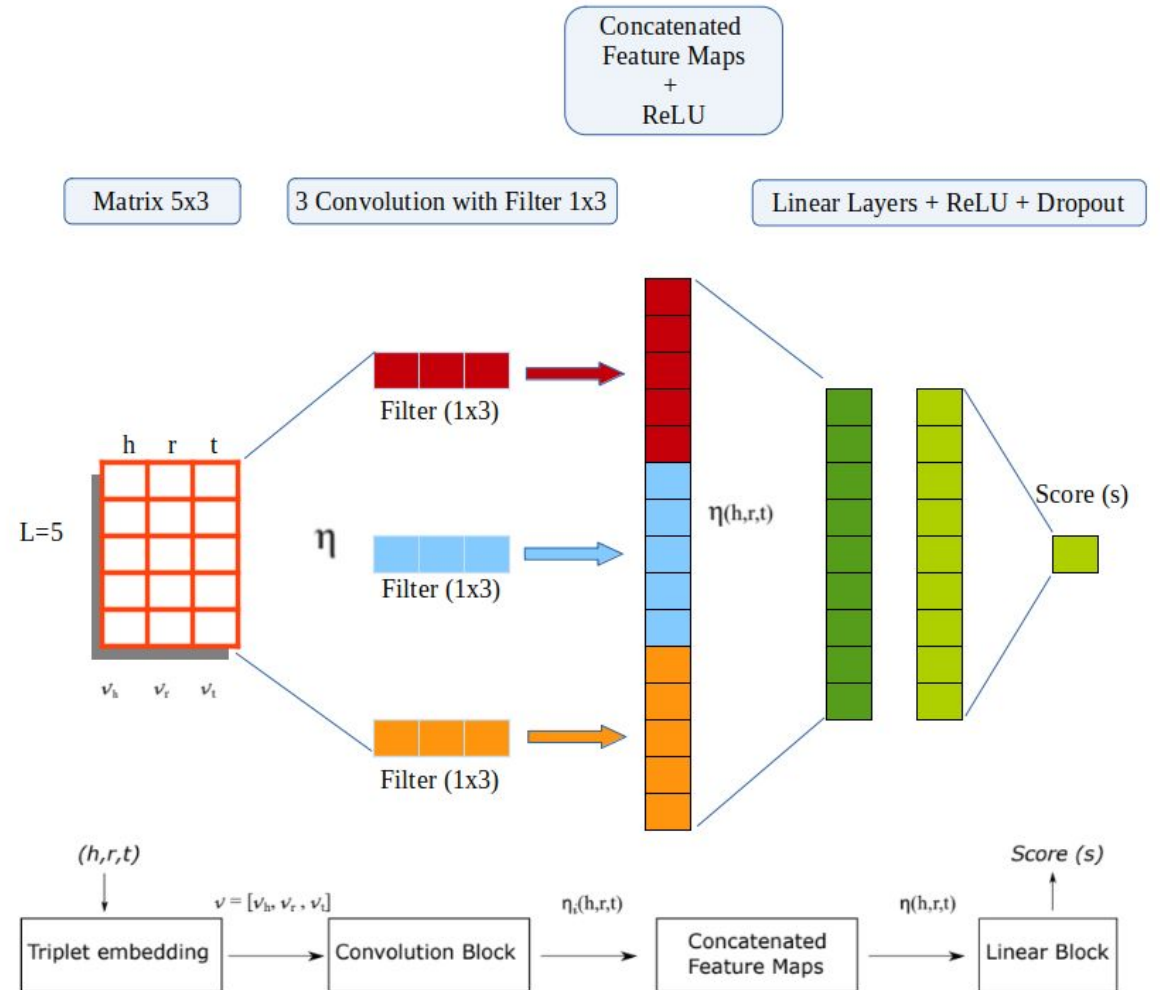


# MConvKB

## Architecture

### 1. Plongement des triplets

- Concaténation des embeddings issue d'un modèle pré-entraîné
- On obtient une matrice qui représente l'embedding du triplet

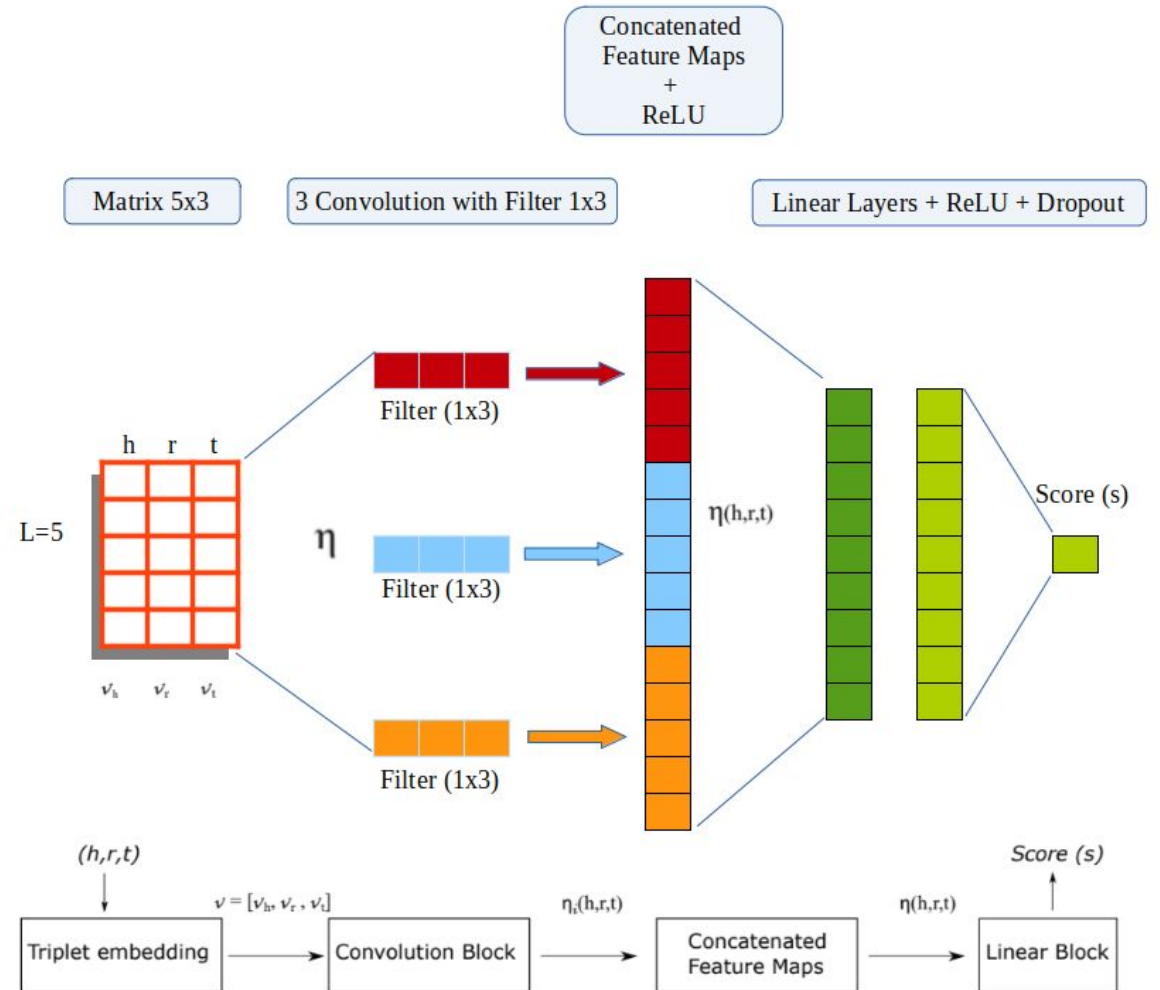


# MConvKB

## Architecture

### 2. Extraction de features

- Couche convolutionnelle avec des filtres 1x3
- Concaténation des features map



# MConvKB

## Architecture

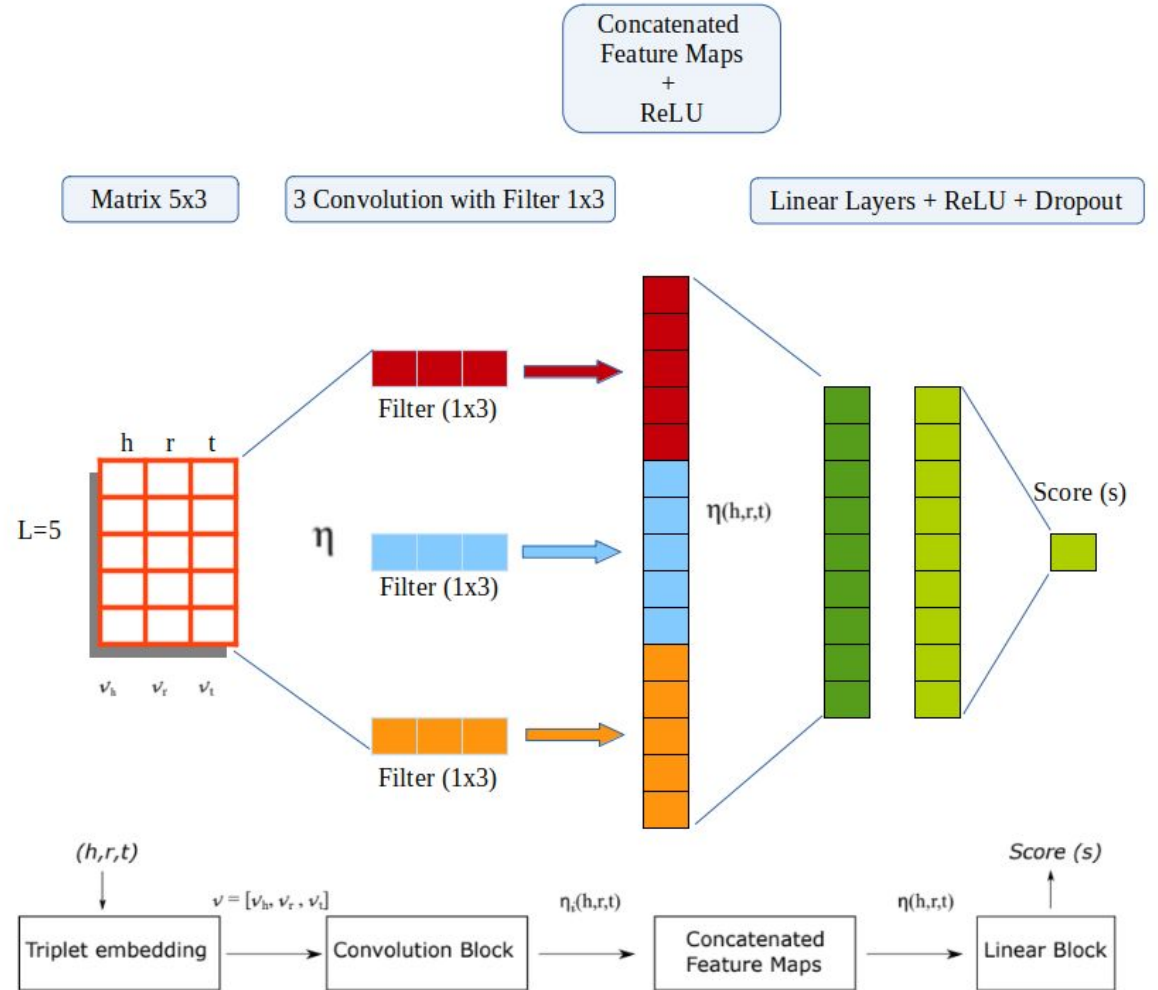
### 3. Classification

- Trois couches fully-connected
- Fonction de score :  $s = \psi_{\theta}(h, r, t)$
- Target score :

$$\hat{s}(h, r, t) = \begin{cases} 1 & (h, r, t) \in \mathcal{V} \\ -1 & (h, r, t) \in \mathcal{V}' \end{cases}$$

- Loss :

$$\mathcal{L} = \sum_{(h,r,t) \in \{\mathcal{V} \cup \mathcal{V}'\}} \log(1 + \exp(s * \hat{s}))$$



# Expérience

## Dataset

- RezoJDM [6], un réseau lexico-sémantique pour le français
- RezoJDM16k [8]

Resource	Entities	triples	Types
RezoJDM16k (Train)	16k	666k	53
RezoJDM16k (Validation)	16k	83k	53
RezoJDM16k (Test)	16k	83k	53

TAB. 2 – *Dataset statistics of the RezoJDM16k split*

# Expérience

## Setup expérimental

- Initialisation à l'aide de plongement pré-entraînés TransE et TransD (dim=200)
- Optimiseur : AdaGrad
- $lr = 0.01$
- Régularisation norme 2 avec  $\lambda = 0.2$  et  $\lambda_2 = 0.01$
- 50 epochs, 500 batches
- Dropout à 0.5
- 64 filtres

# Évaluation

Model	MMR	MR	H@10	H@3	H@1
TransD	0.208	189.186	0.474	0.278	0.064
ConvKB	0.218	186.653	0.493	0.275	0.078
MConvKB	<b>0.337</b>	<b>158.275</b>	<b>0.590</b>	<b>0.384</b>	<b>0.218</b>

TAB. 3 – *Evaluation results on RezoJDM16k with TransD embeddings*

Model	MMR	MR	H@10	H@3	H@1
TransE	0.179	203.310	0.432	0.242	0.041
ConvKB	0.202	224.951	0.444	0.247	0.081
MConvKB	<b>0.295</b>	<b>174.763</b>	<b>0.536</b>	<b>0.330</b>	<b>0.183</b>

TAB. 4 – *Evaluation results on RezoJDM16k with TransE embeddings*

# Conclusion

- Travaux en cours sur une MConvKB, une architecture pour la tâche de prédiction de relations
  - ajout d'une couche de classification à la VGG
  - initialisation à partir de différents plongement pré-entraînés
- Résultats encourageants avec une amélioration de 10% par rapport à la baseline ConvKB
- Perspectives :
  - évaluation sur WN18RR [3] et FB15K-237 [10]
  - expérimenter avec la profondeur et la largeur de la couche convolutionnelle



Merci pour votre attention !



# Bibliographie

- [1] Bollacker, K., C. Evans, P. Paritosh, T. Sturge, and J. Taylor (2008). Freebase: a collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data, pp. 1247–1250.
- [2] Bordes, A., N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko (2013a). Translating embeddings for modeling multi-relational data. Advances in neural information processing systems 26.
- [3] Dettmers, T., P. Minervini, P. Stenetorp, and S. Riedel (2018). Convolutional 2d knowledge graph embeddings. In Thirty-second AAAI conference on artificial intelligence.
- [4] Ji, G., S. He, L. Xu, K. Liu, and J. Zhao (2015). Knowledge graph embedding via dynamic mapping matrix. In Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers), pp. 687–696.
- [5] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
- [6] Lafourcade, M. (2007). Making people play for Lexical Acquisition with the JeuxDeMots prototype. In SNLP'07: 7th International Symposium on Natural Language Processing, Pattaya, Chonburi, Thailand, pp. 7.

# Bibliographie

- [7] Lehmann, J., R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morse, P. Van Kleef, S. Auer, et al. (2015). Dbpedia—a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web* 6(2), 167–195.
- [8] Mirzapour, M., W. Ragheb, M. J. Saeedizade, K. Cousot, H. Jacquenet, L. Carbon, and M. Lafourcade (2022). Introducing RezoJDM16k: a French Knowledge Graph DataSet for Link Prediction.
- [9] Nguyen, D. Q., T. D. Nguyen, D. Q. Nguyen, and D. Phung (2017). A novel embedding model for knowledge base completion based on convolutional neural network. *arXiv preprint arXiv:1712.02121*.
- [10] Toutanova, K. and D. Chen (2015). Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd workshop on continuous vector space models and their compositionality*, pp. 57–66.