Multilingual Knowledge Graph Completion With Zero Seed Alignment (CS 249 Group 11 Course Project Report)

RUSTEM AYGUN*, University of California Los Angeles, USA
RAKESH BAL*, University of California Los Angeles, USA
RAHUL KAPUR*, University of California Los Angeles, USA
ASHWATH RADHACHANDRAN*, University of California Los Angeles, USA

Predicting missing facts in a knowledge graph (KG), also called Knowledge Graph Completion, is an important research problem. Due to the high cost of human labelling, current KGs are incomplete, and the problem amplifies when dealing with the knowledge represented in multiple languages. Prior works have been conducted to explore multilingual KG completion, some of which leverage limited seed alignment as a bridge to embrace the collective knowledge from multiple languages. However, there is very limited work on overcoming the problem of Multilingual KG Completion (MKGC) without using any prior seed alignments. In this project, we study Self Supervised Adaptive Graph Alignment (SS-AGA), a novel framework that combines the power of seed alignments and self-supervised learning to tackle the MKGC problem. SS-AGA uses many modules to achieve this, including Relation-aware Multilingual KG Network and Self-Supervised New Pair Generation. SS-AGA achieves State of the art results in 2 different datasets in MKGC problem. We study the details of SS-AGA, implementing the model as a multilingual baseline and then propose 3 novel approaches to modify the model in order to fully abandon the use of any prior seed alignments in the data while still maintaining the performance or even surpassing it. We show that one of our proposed approaches obtains better result than SS-AGA in the low resource Greek language of the DBP-5L dataset. We also show that 2 of our approaches outperform the monolingual baseline TransE in many languages which are not low-resource like Greek. Our experimental results show that the zero seed alignment setting has a lot of scope for improvement and can lead to real-world use as human annotation in seed alignments are hard to find. Studies in this direction can also help extend the power of graph neural networks in the KG completion task to more scarcely available languages. GitHub Page for the Course Project: https://github.com/rakeshbal99/CS249_Course_Project.

CCS Concepts: • **Computing methodologies** → Machine learning.

Additional Key Words and Phrases: Graph Neural Networks, Knowledge Graph Completion, Multilingual Data, Unsupervised Learning

ACM Reference Format:
Rustem Aygun, Rakesh Bal, Rahul Kapur, and Ashwath Radhachandran. 2022. Multilingual Knowledge Graph Completion With Zero Seed Alignment (CS 249 Group 11 Course Project Report). In *UCLA CS 249: Special Topics in Graph Neural Networks Spring 2022 Course Project Report*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Modern knowledge graphs (KGs) are far from complete, which makes knowledge graph completion (KGC) an important research area. KGs model real-world concepts and the relations that interconnect them. This structured representation of information is important because this knowledge is readily accessible and drives knowledge-driven use cases. Typically prior knowledge is stored in KGs as a triple (head, relation, tail) of information. One example of this, as seen in Figure 1 would be (LeBron James, Plays on, The Los Angeles Lakers). A KG contains multiple of these triples. As new things are discovered, knowledge graphs will inherently fall behind in how up-to-date their contents are. This is where KGC enters, as a successful implementation of this task and can be used to predict missing triples without the need for human annotation and thus, complete the KG.

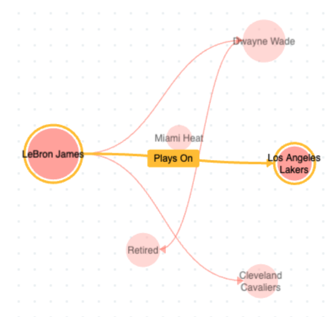

Fig. 1. Simple KG example

2 PROBLEM DEFINITION

The problem of knowledge graph incompleteness is exacerbated when working with data that has a multilingual representation, as human annotation is an expensive task with regard to time and manpower, especially for cases with low-resource languages.

Typically, KG embedding approaches stem from the information in a single KG. Past efforts have attempted to learn from each monolingual KG independently (Peng et al. [14], Xu et al. [19], Liang et al. [13]), and this leads to KGC underperforming in situations where a low-resource language is involved. However, multilingual knowledge graph completion is an important task since there is an argument that an embedding would be enhanced by the facts from multiple different KGs since they would each have their respective pros and cons in regards to data quality and scope. Multilingual KGC is mainly bottlenecked by the need for sufficient seed alignment information between the various KGs. Seed alignments are the mappings between similar concepts in two KGs. Alignments are helpful for two main reasons: 1) greater knowledge propagation

*All authors contributed equally to this research.

UCLA CS 249 S22 Report, June, 2022, Los Angeles, CA

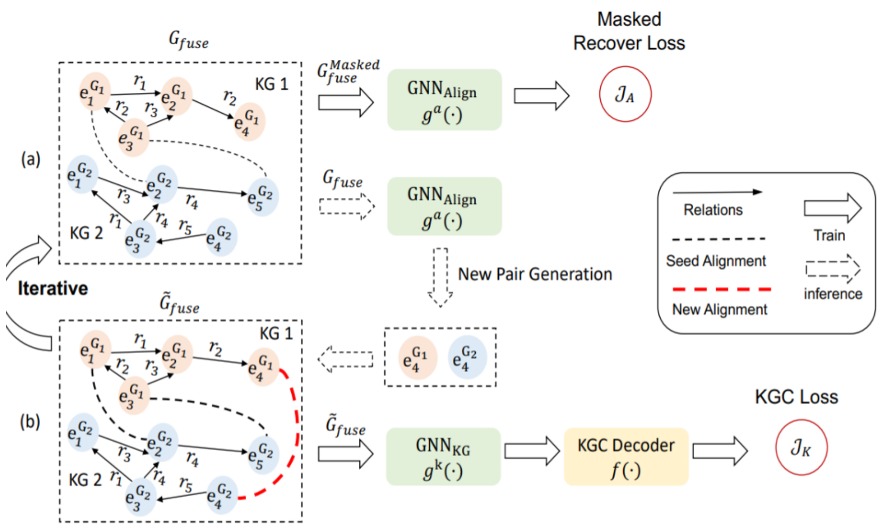

Fig. 2. SS-AGA

in sparse KGs 2) relieves the need for manually labelling across all languages.

Efforts to narrow language gaps have been started on multilingual KG embedding methods, but the lack of seed alignment (mappings between similar concepts in two KGs) prevents effective knowledge transfer. Sometimes the acquisition of alignment labels is a costly and noisy process, but simply using unsupervised methods (correlation or similarity between entities) can fall short of effective language alignment in KGs. We are aiming to address the setting where there is zero seed alignment present (as opposed to there being some pre-existing alignment pairs for knowledge transfer) since this is more difficult and more practical for real-world situations. And to reiterate, this problem is important because most current KGs are incomplete, and being able to fill in missing facts in these KGs can improve current technology (such as Google's SEOs).

In this work, we propose 3 approaches to circumvent the need for seed alignments in the SS-AGA model and show their effectiveness and efficiency over various language dataset KGs. To summarize, our work and contributions are as follows:

- We study the SS-AGA model in depth by using it as a baseline and obtaining its accuracies.
- We propose three approaches to avoid using seed alignments by the SS-AGA model for the Multilingual KG Completion task while still maintaining the performance.
- Results on DBP-5L dataset demonstrate that one of your approaches can outperform SS-AGA, a lower resource KG like Greek.

3 RELATED WORK

Monolingual knowledge graph embedding techniques aim to learn the relationship between the entities in a knowledge language graph and combine the embeddings with various scoring functions such as translation-based TransE (Bordes et al. [2]), TransH (Wang et al. [18]); rotation-based RotatE (Sun et al. [15]), and language-model-based KG-BERT (Yao et al. [21]) to infer the possible entity-relation-entity triples. Previously proposed monolingual knowledge graph embedding techniques do not consider the entity neighbor relationships (Bordes et al. [2]; Sun et al. [15]; Con [1]) have been proposed to address this limitation. Even though GNN-based KG embedding techniques improve the performance of the previously proposed basic graph embedding techniques in rich-resource languages, their KGC performance is limited when it comes to the low-resource languages due to the knowledge graph sparseness.

Multilingual knowledge graph completion approaches to transfer the knowledge learned from one language to another to solve the KG sparseness problem. However, generating a model that can successfully perform this kind of knowledge transfer is far from trivial for two reasons. First, achieving high-performance multilingual knowledge graph completion requires learning the accurate distance of the aligned entities in different languages without losing their contextual semantics in their respective KGs. If the aligned entities are treated equally by ignoring this fact, MKGC models might bring irrelevant information for given KG queries (Huang et al. [10]). Second, MKGCs should perform well in limited seed alignment settings because finding aligned entities in different languages requires tremendous human effort. As a result, many real-world KGs have a limited number of aligned entities with the other languages.

Some of the existing MKGC approaches learn the knowledge graph embeddings for each language independently and then minimize the distance between the aligned entities in different languages (Chen et al. [6]; Zhang et al. [22]; Sun et al. [17]). These approaches lose the contextual importance of the aligned entities by treating them equally. Also, they perform poorly in limited seed alignment settings. Some recent studies address the limited seed alignment problem by generating new alignment pairs without supervision during the training phase(Chen et al. [4];Chen et al. [6]). These approaches generate noise during the training and could achieve limited MKGC success. Moreover, they are not immune to the problem of equal treatment of parallel entities.

The recently proposed Self-Supervised Adaptive Graph Alignment (SS- AGA) framework ((Huang et al. [10])) addresses the aforementioned limitations of the previous models via a GNN-based self-supervised model and out-performs them in limited seed alignment settings. The proposed models in our current study are inspired by SS-AGA.

In addition to these studies, some of the previous work tackled entity alignment problems only without MKGC. These approaches are based on a learnable transformation matrix(MTransE : Chen et al. [5]), collective neighborhood aggregation(CG-MuA : Zhu et al. [24]), utilizing attribute information (Zhang et al. [22]; Chen et al. [4]), bootstrapping (BootEA - Sun et al. [16]), rule-based structural difference finding (Cao et al. [3]), parameter sharing based alignment improvement (Zhu et al. [23])

4 SS-AGA MODEL DESCRIPTION

In this section, we will briefly introduce the Self-Supervised Adaptive Graph Alignment model since all of our zero seed alignment models are based on the modified versions of it. Please refer to Figure 2 for an illustration of the method.

SS-AGA is a GNN method that addresses the limited seed alignment problem via self-supervised pair generation. The model first merges different language KGs into a single KG called G_{fuse} by considering the relation between the aligned pairs $\left(e_i, r_{\text{align}}, e_j\right)$ and $\left(e_j, r_{\text{align}}, e_i\right)$ as a new edge type $r_{\text{align.}}$. In the training step, some of the initial alignment edges are selected randomly and then masked. Later, the GNN encoder $g^a(\cdot)$ is trained to recover the masked edges. The trained $g^a(\cdot)$ is used to create new alignment edges. The newly generated edges are incorporated into the $\widetilde{G}_{\text{fuse}}$ graph and another GNN encoder called $g^k(\cdot)$ is trained to learn the contextual information of the entities. In the final step, the KGC decoder function is applied to the learned embeddings in the previous step to calculate the triple scores. The detailed explanations of each different module are given below.

4.1 GNN models

SS-AGA uses two separate attention-based relation-aware GNN encoders, $g^a(\cdot)$ and $g^k(\cdot)$ to learn masked alignment recovery and contextual embeddings respectively. $g^a(\cdot)$ and $g^k(\cdot)$ have the same GNN structure, and they are trained for different tasks to capture the relevant embeddings of their target tasks in a more precise manner. Moreover, both of them consist of L stacked layer to capture multi-hop neighbour information of the entities.

4.2 Masked Alignment Recovery GNN

$g^a(\cdot)$ is trained based on the lost function $\mathcal{J}_A$ to learn to recover the masked alignments. In addition to graph-related embeddings, the textual embeddings of the KG entities are also learned via a multilingual pre-trained language model (Devlin et al. [9]) called mBERT. Since mBERT is a multilingual word embedding model, it can capture the rich semantic relationships of the entities in different languages.

$$\mathcal{J}_A^{G_i \leftrightarrow G_j} = \sum_{\substack{(e_h,e_t)\in\Gamma_{ij}^p \\ (e_{h'},e_{t'})\in\Gamma_{ij}^n}} \left[\left\|\widetilde{\boldsymbol{e}}_h^a - \widetilde{\boldsymbol{e}}_t^a\right\|_2 - \left\|\widetilde{\boldsymbol{e}}_{h'}^a - \widetilde{\boldsymbol{e}}_{t'}^a\right\|_2 + \gamma_a\right]$$

$$\mathcal{J}_A = \sum_{1\le i<j\le M} \mathcal{J}_A^{G_i \leftrightarrow G_j},$$

After learning both the graph and textual embeddings, the maximum of the cosine similarity of these two separate embeddings is found to define the similarity score between two entities. The model generates a new alignment between two unaligned entities if they are the nearest neighbours based on the cross-domain similarity local scaling (CSLS) measure (Conneau et al. [8]). The newly generated aligned pairs are added to G_{fuse} to learn the contextual embeddings in the next step. In this way, SS-AGA addresses the limited seed alignment problem via self-supervised alignment generation.

$$\operatorname{sim}\left(e_i, e_j\right) = \max\left(\cos\left(\boldsymbol{e}_i^a, \boldsymbol{e}_j^a\right), \cos\left(\boldsymbol{e}_i^{\text{text}}, \boldsymbol{e}_j^{\text{text}}\right)\right)$$

$$\operatorname{CSLS}\left(e_i, e_j\right) = 2\operatorname{sim}\left(e_i, e_j\right) - s\left(e_i\right) - s\left(e_j\right)$$

$$\text{subject to } s\left(e_i\right) = \frac{1}{K}\sum_{e_{i'}\in\mathcal{N}(e_i)} \operatorname{sim}\left(e_i, e_{i'}\right),$$

4.2.1 Contextual Embedding GNN. $g^k(\cdot)$ is trained on $\widetilde{G}_{\text{fuse}}$ to learn contextual relationships between the entities. Based on the learned embeddings in this step, the KGC decoder function is used to generate triple scores. The loss function for this task is defined as follows:

$$\mathcal{J}_K = \sum_{\substack{(e_h,r,e_t)\in\mathcal{T}_m \\ (e_{h'},r,e_{t'})\notin\mathcal{T}_m \\ m=1,\ldots,M}} \left[f((e_h', r, e_t') - f((e_h, r, e_t)) + \gamma\right]$$

Thanks to the separate contextual embedding model, SS-AGA avoids the equal treatment of the aligned entities problem that exists in all the previous MKGC approaches. For more detailed information about the SS-AGA framework and the given formulas above, please refer to the original SS-AGA paper (Huang et al. [10]).

5 METHODS

5.1 Approach 1: SS-AGA + mBERT gen

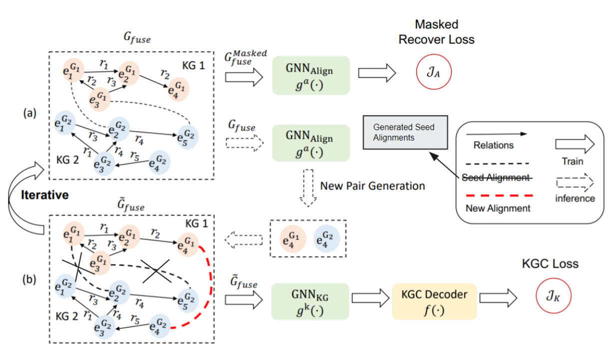

Fig. 3. Approach 1: SS-AGA + mBERT gen.

Dataset	#Entity	#Relation	#Triple	#Aligned Links
EN	13,996	831	80,167	16,916
FR	13,176	178	49,015	16,877
ES	12,382	144	54,066	16,347
JA	11,805	128	28,774	16,263
EL	5,231	111	13,389	9,042

Table 1. Statistics of DBP-5L dataset. #Aligned Links denotes the number of alignment pairs where one of the aligned entities belongs to that KG.

The first approach is to obtain textual embeddings of the entities to generate seed alignments and then feed them to the SS-AGA model. A schematic of this given in Figure 3 which is a modified version of the original Figure 2.

We extract the textual representations of entities (e.g. "Los Angeles") and use a pretrained multilingual BERT model to obtain its embedding in a latent vector space (768 in this case). We use the distilled mBERT model from sentence transformers [1] library for this purpose. After this we calculate the cosine similarity of embeddings of all entities in a particular language with embeddings of all entities in another language for each pair of languages. Hence, we get a 2D matrix of cosine similarities for all pairs of languages. For example, if we have el-en pair ->we get a 2D matrix of shape (5231, 13996) of cosine similarities. After this, we take the top k values from each of the matrices, and the entities in the corresponding languages are treated as seed alignments. We choose k in the following manner:

- If a pair has 'el' (Greek) as one of the languages we take k as *2000*.
- For rest of the pairs of languages we take k as *5000*.

The motivation behind such k selection is to roughly align with the number of seed alignments provided in the original dataset. Next, we feed these generated seed alignments to the SS-AGA model, keeping the rest of all the modules exactly the same as before. A statistical distribution of matching seed alignments pairs that are generated to the pairs already present in the dataset (groundtruth) is given in Figure 7. The figure shows low matching of seed pairs generated to seed pairs present and hence shows that even though the textual representations of the entities might be very similar, the structural representations of the entities can be very different and hence, we need to use better techniques to solve the problem. We will discuss the results in Section 7.

5.2 Approach 2: SS-AGA + no seeds

In this approach, we remove all the seed alignments present either in the dataset or in the generated ones in Section 5.1. A schematic of this given in Figure 4 which is a slightly modified version of the original Figure 2. We freeze the alignment loss module for some epochs (hyperparameter m), and we generate new seed pairs for those m epochs using the New Pair Generation module. After that, we unfreeze the alignment loss module, and it starts to train. The KGC Loss module is untouched in this approach, and the rest of the procedure is exactly the same. We will discuss the evaluation results and performance variance on hyperparameter m in Section 7.

[1] https://www.sbert.net/index.html

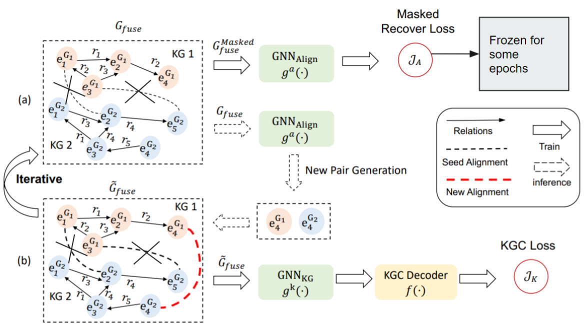

Fig. 4. Approach 2: SS-AGA + no seeds

5.3 Approach 3: SS-AGA + mBERT fused

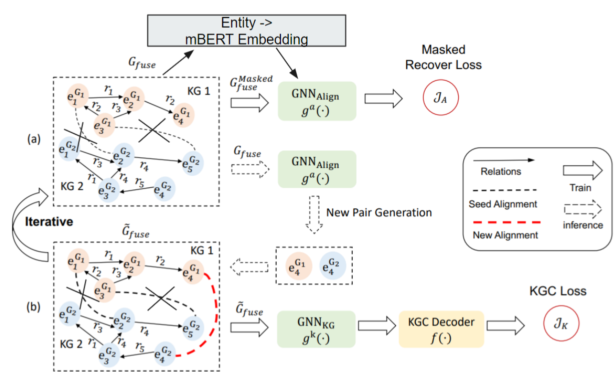

Fig. 5. Approach 3: SS-AGA + mBERT fused

In this approach, we add mBERT model to the SS-AGA model and train the entire module end to end. A schematic of this is given in Figure 4, which is again a slightly modified version of the Figure 5. This will finetune the mBERT model to accommodate the structural representations of the entities along with their textual embeddings and, in theory, can lead to good results. We discuss challenges in training the model in this approach in section 7.

Method	Metric	JA	EL	ES	FR
Monolingual Baselines					
TransE	H@1	21.1	13.1	13.5	17.5
	H@10	48.5	43.7	45	48.8
	MRR	25.3	24.3	24.4	27.6
Multilingual Baselines					
SS-AGA	H@1	34.6	30.8	25.5	27.1
	H@10	66.9	58.6	61.9	65.5
	MRR	42.9	35.3	36.6	38.4
SS-AGA + Bert Gen Seeds	H@1	17.9	16.8	14.1	17.2
(Approach 1)	H@10	45.5	57.9	50.4	53
	MRR	27.7	30.9	26.4	29.1
SS-AGA + No Seeds	H@1	20.4	17.3	9.7	13.3
(Approach 2)	H@10	51.9	59.8	52.5	55.2
	MRR	31.1	31.1	23.4	27.1

Table 2. Results on Approach 1 and 2 on different metrics

6 EXPERIMENTS

6.1 Dataset

We ran all of our experiments in the widely used benchmark dataset called **DBP-5L** [7]. It contains five language-specific KGs from DBpedia [11] i.e. English (EN), French (FR), Spanish (ES), Japanese (JA), Greek (EL). As the original dataset only contains structural information, the authors from [10] additionally crawled the text information for these entities and relations based on the given URLs. Table 1 lists the statistics of the dataset. The English (EN) KG is the most populated one and the Greek (EL) KG is the least populated one. We use the same split as per [7] to split the facts in each KG into three parts: 60% for training, 30% for validation and weight learning, and 10% for testing.

6.2 Baselines

We used TransE [2], and SS-AGA [10] models as the monolingual and multilingual baselines for our work, respectively. The reason for choosing TransE is twofold:

- It is a very popular technique in KG completion literature and Graph Networks in general.
- SS-AGA uses TransE as its KG Decoder, and hence using it as a baseline would be beneficial to study the effectiveness of our approaches.

6.3 Evaluation Protocol

We follow the same evaluation protocol as suggested by [10] where in the testing phase, given each query $f(e_h, r, ?e_t)$, we compute the plausibility scores $f(e_h, r, \tilde{e_t})$ for triples formed by each possible tail entity $\tilde{e_t}$ in the test candidate set and rank them. We report the mean reciprocal ranks (MRR), accuracy (Hits@1), and the proportion of correct answers ranked within the top 10 (Hits@10) for testing. We also adopt the filtered setting from previous works based on the premise that the candidate space has excluded the triples that have been seen in the training set ([18], [20]).

7 RESULTS

The main results of all of our experiments are given in Table 4 and Table 3. In this section, we will discuss the results on various approaches.

7.1 Approach 1

In EL, ES and FR languages, *Approach 1* outperformed the monolingual baseline *TransE* in all the evaluation metrics. In JA language, *Approach 1* outperformed *TransE* only in the MRR metric. Also, *Approach 1* was not able to compete with SS-AGA model in almost all of the languages except for EL where it was almost at par with SS-AGA in Hits@10 metric. This shows that EL language being a low resource KG, SS-AGA is not able to fully utilise the power of provided seed alignments, and hence *Approach 1* with its generated seed alignments is able to compete with SS-AGA. As a summary, *Approach 1* is clearly better than *TransE* in almost all languages and not at par with SS-AGA in general except for EL.

7.2 Approach 2

In contrast to Approach 1, *Approach 2* outperforms monolingual baseline *TransE* in all languages in almost all metrics (except in ES and FR for Hits@1 metric). This shows that *Approach 2* is superior to *TransE* in general. Also, *Approach 2* is better than *Approach 1* in JA and EL languages in all metrics and ES and FR languages in the Hits@10 metric. This shows even though *Approach 2* uses no prior seeds, it is able to outperform *Approach 1* in most cases. Although, *Approach 2* does not perform at par with SS-AGA in most languages in most metrics, it is much closer to SS-AGA given the fact it does not use any ssed and even outperforms SS-AGA in EL language in the Hits@10 metric. This again shows that EL language

Method	Metric	m = 0	m = 3	m = 6	m = 9	m = 12
SS-AGA + No Seeds	H@1	14.9	17.3	17.5	14.7	15.7
(EL Dataset)	H@10	55.7	59.8	56.9	56.7	58.6
	MRR	29.6	31.1	30.8	29.8	30.5

Table 3. Results on Approach 2 on hyperparamter m

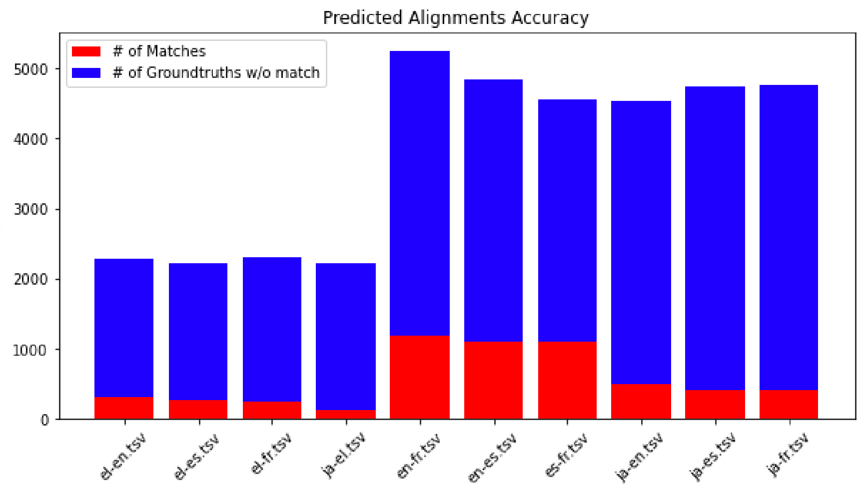

Fig. 6. Approach 1: Unsupervised Generation of Alignments Statistics

being a low resource KG, SS-AGA is not able to fully utilise the power of provided seed alignments and hence *Approach 2* is able to outperform SS-AGA.

We also provide performance of *Approach 2* on changing the hyperparameter m in table 3. The results show that there is a sweet spot of $m = 3$ where the performance is maximized, as shown in Figure ??. For lower m, the alignment model doesn't have any significant seeds to start training with and hence the performance of the overall model is affected. For higher m, since the alignment module is frozen longer, the new seed pairs generated are not of high quality, and hence when the alignment module is unfreezed, it is not able to generalise, affecting performance.

7.3 Approach 3 and Challenges

Even though we were able to implement *Approach 3* from a programming perspective, the GPU resources available at our disposal prevented us from getting any results with this approach. The input subgraph in the original implementation [2] at the stage where the bert model is introduced was very large. Hence we constantly received CUDA out of memory error for even very small batch sizes (2) and smaller values of hyperparameter k, which controls the number of neighbours in the computational subgraph. We hope to obtain some results in the future with this approach by mitigating the resource constraint problem and we believe the results would be very intriguing.

We also faced challenges in experimenting with the English (EN) KG of the dataset as it is a very large KG, and hence we were consistently obtaining *CUDA* out of memory error. Hence, we don't have any results for EN KG.

8 DISCUSSION AND CONCLUSION

In this work, we experimented with the Multilingual Knowledge Graph Completion problem in a zero seed alignment setting and developed various approaches to tackle the problem. We specifically used the recently proposed SoTA model called SS-AGA, which uses prior seed alignments for the objective. We designed 3 approaches for solving the problem, and 2 of them outperformed the monolingual baseline *TransE* in almost all languages in all metrics. Also, one of our approaches was able to outperform SS-AGA model in EL KG in the Hits@10 metric. Even though our results are not suitable for real-world application as of now, we believe our results provide a very viable ground for future work in this direction, leading to much better results in the zero-shot setting. In this regard, we believe our approach 3 might provide much better results than the other two approaches, and we hope to obtain some results from that approach as soon as possible by mitigating the resource constraint problem.

As future work, we propose to design a new model where we discard any sort of seed alignment requirement for the model. We take intuition from [12] and plan to devise a similar model where

[2] https://github.com/amzn/ss-aga-kgc

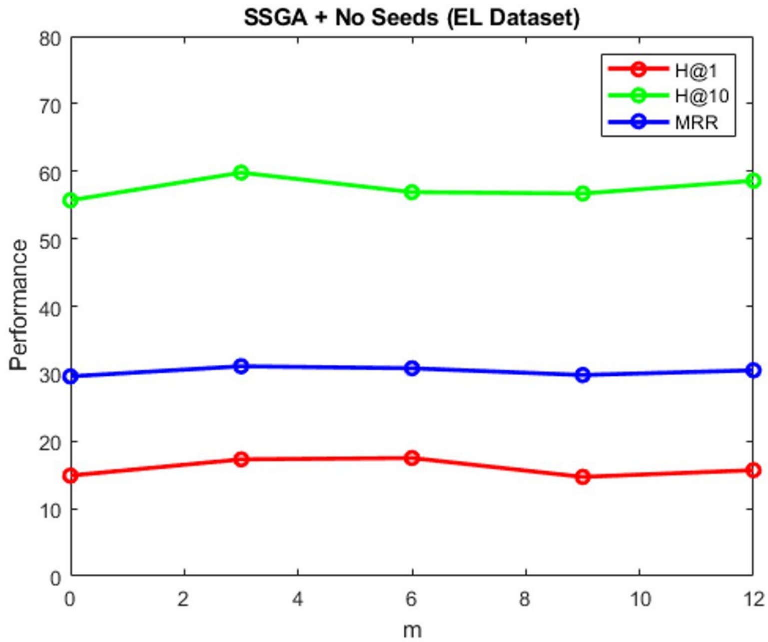

Fig. 7. Approach 2: Plot of evaluation metric

we use only a single encoder and train the single encoder for all KGs without the need to fuse the KGs to a single KG. This would be an even more unsupervised approach without needing any seed alignments and hence reducing human annotations and providing the ability to extend to lower resource KGs.

9 WORKLOAD DISTRIBUTION

Give in table **??**

ACKNOWLEDGMENTS

We would like to thank our mentors **Zijie Huang** and **Roshni Iyer** for their continued support throughout the project.

REFERENCES

[1] 2018. Convolutional 2D Knowledge Graph Embeddings. *Proceedings of the AAAI Conference on Artificial Intelligence* 32, 1 (Apr. 2018). https://doi.org/10.1609/aaai.v32i1.11573

[2] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *Advances in Neural Information Processing Systems*, C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger (Eds.), Vol. 26. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf

[3] Yixin Cao, Zhiyuan Liu, Chengjiang Li, Zhiyuan Liu, Juanzi Li, and Tat-Seng Chua. 2019. Multi-Channel Graph Neural Network for Entity Alignment. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 1452–1461. https://doi.org/10.18653/v1/P19-1140

[4] Muhao Chen, Yingtao Tian, Kai-Wei Chang, Steven Skiena, and Carlo Zaniolo. 2018. Co-training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-lingual Entity Alignment. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*. International Joint Conferences on Artificial Intelligence Organization, 3998–4004. https://doi.org/10.24963/ijcai.2018/556

[5] Muhao Chen, Yingtao Tian, Mohan Yang, and Carlo Zaniolo. 2017. Multilingual Knowledge Graph Embeddings for Cross-Lingual Knowledge Alignment. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence* (Melbourne, Australia) *(IJCAI'17)*. AAAI Press, 1511–1517.

[6] X. Chen, Muhao Chen, Changjun Fan, Ankith Uppunda, Yizhou Sun, and Carlo Zaniolo. 2020. Multilingual Knowledge Graph Completion via Ensemble Knowledge Transfer. In *FINDINGS*.

[7] Xuelu Chen, Muhao Chen, Changjun Fan, Ankith Uppunda, Yizhou Sun, and Carlo Zaniolo. 2020. Multilingual knowledge graph completion via ensemble knowledge transfer. *arXiv preprint arXiv:2010.03158* (2020).

[8] Alexis Conneau, Guillaume Lample, Marc'Aurelio Ranzato, Ludovic Denoyer, and Hervé Jégou. 2018. Word translation without parallel data. In *International Conference on Learning Representations*. https://openreview.net/forum?id=H196sainb

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[10] Zijie Huang, Zheng Li, Haoming Jiang, Tianyu Cao, Hanqing Lu, Bing Yin, Karthik Subbian, Yizhou Sun, and Wei Wang. 2022. Multilingual Knowledge Graph Completion with Self-Supervised Adaptive Graph Alignment. *arXiv preprint arXiv:2203.14987* (2022).

[11] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web* 6, 2 (2015), 167–195.

[12] Liunian Harold Li, Haoxuan You, Zhecan Wang, Alireza Zareian, Shih-Fu Chang, and Kai-Wei Chang. 2020. Unsupervised vision-and-language pre-training without parallel images and captions. *arXiv preprint arXiv:2010.12831* (2020).

[13] Zongwei Liang, Junan Yang, Hui Liu, and Keju Huang. 2021. A Semantic Filter Based on Relations for Knowledge Graph Completion. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 7920–7929. https://doi.org/10.18653/v1/2021.emnlp-main.625

[14] Xutan Peng, Guanyi Chen, Chenghua Lin, and Mark Stevenson. 2021. Highly Efficient Knowledge Graph Embedding Learning with Orthogonal Procrustes Analysis. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.

Task	People
Literature Review	Rakesh, Rustem, Ashwath and Rahul
Implementing Approach 1	Rakesh and Ashwath
Implementing Approach 2	Rakesh and Rustem
Preparing Slides	Rakesh, Rustem, Ashwath
Writing Report	Rakesh, Rustem, Ashwath and Rahul

Table 4. Work Load Distribution

naacl-main.187

[15] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. 2019. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net. https://openreview.net/forum?id=HkgEQnRqYQ

[16] Zequn Sun, Wei Hu, Qingheng Zhang, and Yuzhong Qu. 2018. Bootstrapping Entity Alignment with Knowledge Graph Embedding. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*. International Joint Conferences on Artificial Intelligence Organization, 4396–4402. https://doi.org/10.24963/ijcai.2018/611

[17] Zequn Sun, Qingheng Zhang, Wei Hu, Chengming Wang, Muhao Chen, Farahnaz Akrami, and Chengkai Li. 2020. A Benchmarking Study of Embedding-Based Entity Alignment for Knowledge Graphs. *Proc. VLDB Endow.* 13, 12 (jul 2020), 2326–2340. https://doi.org/10.14778/3407790.3407828

[18] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge Graph Embedding by Translating on Hyperplanes. In *AAAI*.

[19] Jingwen Xu, Jing Zhang, Xirui Ke, Yuxiao Dong, Hong Chen, Cuiping Li, and Yongbin Liu. 2021. P-INT: A Path-based Interaction Model for Few-shot Knowledge Graph Completion. In *Findings of the Association for Computational Linguistics: EMNLP 2021*. Association for Computational Linguistics, Punta Cana, Dominican Republic, 385–394. https://doi.org/10.18653/v1/2021.findings-emnlp.35

[20] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2014. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575* (2014).

[21] Liang Yao, Chengsheng Mao, and Yuan Luo. 2019. KG-BERT: BERT for Knowledge Graph Completion. https://doi.org/10.48550/ARXIV.1909.03193

[22] Qingheng Zhang, Zequn Sun, Wei Hu, Muhao Chen, Lingbing Guo, and Yuzhong Qu. 2019. Multi-view Knowledge Graph Embedding for Entity Alignment. In *IJCAI*. 5429–5435. https://doi.org/10.24963/ijcai.2019/754

[23] Hao Zhu, Ruobing Xie, Zhiyuan Liu, and Maosong Sun. 2017. Iterative Entity Alignment via Joint Knowledge Embeddings. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 4258–4264. https://doi.org/10.24963/ijcai.2017/595

[24] Qi Zhu, Hao Wei, Bunyamin Sisman, Da Zheng, Christos Faloutsos, Xin Luna Dong, and Jiawei Han. 2020. *Collective Multi-Type Entity Alignment Between Knowledge Graphs*. Association for Computing Machinery, New York, NY, USA, 2241–2252. https://doi.org/10.1145/3366423.3380289