

# Hesitant Fuzzy Graph Neural Network-Based Prototypical Network

Subjects: [Computer Science](#), [Artificial Intelligence](#)

Contributor: Xinyu Guo , Bingjie Tian , Xuedong Tian

Few-shot text classification aims to recognize new classes with only a few labeled text instances. Previous studies mainly utilized text semantic features to model the instance-level relation among partial samples. However, the single relation information makes it difficult for many models to address complicated natural language tasks. A novel hesitant fuzzy graph neural network (HFGNN) model that explores the multi-attribute relations between samples is proposed. HFGNN is combined with the Prototypical Network (HFGNN-Proto) to achieve few-shot text classification.

few-shot text classification

multi-attribute relations

graph neural networks

hesitant fuzzy set

prototypical networks

## 1. Introduction

In recent years, the great success of deep learning has promoted the development of multitudinous fields such as computer vision and natural language processing [\[1\]\[2\]\[3\]](#), but the effectiveness of deep learning models relies on a large amount of labeled data. The generalization ability of deep learning models is severely limited when labeled data are scarce. Humans, on the other hand, have the ability to learn quickly and can easily build awareness of new things with just a few examples. This significant gap in learning between machine learning models and humans inspires researchers to explore few-shot learning (FSL) [\[4\]](#).

Inspired by the human-learning process, researchers proposed a meta-learning strategy for FSL, which utilizes the distribution of similar tasks to learn how to identify unseen classes accurately and efficiently with a small amount of training data. A cross-task meta-learner learns from multiple similar tasks and provides better initialization for unseen classes based on the knowledge acquired from prior experience. One of the typical meta-learning methods is the Prototypical Network [\[5\]](#), which computes the class prototypical representation of the support set and classifies the query sample to the nearest prototype.

Prototypical Network [\[5\]](#) and its variants [\[6\]\[7\]\[8\]\[9\]](#) have been widely used for few-shot text classification tasks. Different from the Prototypical Network that computes class prototypes and query sample embeddings separately, MLMAN [\[6\]](#) interactively encodes text based on the matching information between the query and support sets at the local and instance levels. Gao et al. [\[7\]](#) proposed a hybrid attention-based prototypical network that employs instance-level attention and feature-level attention to highlight important instances and features, respectively. Sun

et al. [8] improved the Prototypical Network by using feature level, word level, and instance level multi cross attention. Geng et al. [9] proposed the Induction Network that induces a better class-level representation using a dynamic routing algorithm [10]. Although the above methods have considered the intra-class similarity of the support set and the relations between support and query, they ignore inter-class dissimilarity and the relations between query samples. Furthermore, these methods only measure instance-level relations while neglecting other substantive relations. Due to the complexity and diversity of textual forms, simple relations have difficulty describing the true connection between texts and even introduce additional noise to the model.

Instead, researchers measure the commonalities and differences among all samples in the task from multiple aspects. Inspired by the Distribution Propagation Graph Neural Network [11], researchers introduce the distribution-level information of one sample to all other support samples into few-shot text classification. To better explore relations information, a dual-graph structure consisting of instance graphs and distribution graphs is adopted in HFGNN. The instance graph models instance-level relations and directional relations based on instance features. The distribution graph aggregates instance-level relations to model distribution-level relations and the distance relations. The relations in researchers' model represents the similarity between samples at multiple levels. However, similarity is a fuzzy concept, which is not clearly defined. To solve the problem of multiple fuzzy relations, researchers introduce HFS [12] theory, which can handle multi-attribute decision-making problems well for the HFGNN model. Researchers design the membership functions corresponding to the multi-attribute relations for the comprehensive evaluation to avoid the loss of relations information. In addition, researchers use a linear function that can further enhance stronger relations and weaken weaker relations to help the model generate more reasonable graph structures and provide inductive biases for the enhancement of instance features. Finally, a prototypical network takes the instance features generated by HFGNN as input and quickly classifies query samples. HFGNN-Proto adopts an episodic strategy [13] for meta-training in an end-to-end manner. It has a stronger generalization ability and can adapt to new classes without retraining.

## 2. Few Shot Learning

Early studies mainly applied fine-tuning [14] and data augmentation [15] to alleviate the overfitting problem caused by insufficient training data but achieved unsatisfactory results. In the meta-learning strategy [16][17][18][19], transferable knowledge that can guide the learning of models are extracted from various tasks, so that the model has the ability of learning to learn. Current meta-learning methods mainly include optimization-based methods [18][19][20] and metric learning [5][13][21]. Some representative few-shot learning methods and their corresponding descriptions are listed in **Table 1**.

**Table 1.** Representative few-shot learning methods and their descriptions.

Type	Core Idea	Method	Description
Optimization-based Methods	This type of approach aims to learn to optimize the model parameters given the gradients	MAML [18]	MAML trains with a set of initialization parameters, and on the basis of the initial parameters, one or more steps of gradient adjustment can achieve the

Type	Core Idea	Method	Description
	computed from the few-shot examples.		purpose of quickly adapting the model to new tasks with only a small amount of data.
		SNAIL <a href="#">[20]</a>	A novel combination of temporal convolution and soft attention to learn the optimal optimization strategy.
		ATAML <a href="#">[22]</a>	ATAML facilitates task-agnostic representation learning through task-agnostic parameterization and enhances the adaptability of the model to specific tasks through attention mechanism.
Metric Learning	In metric-based methods, instances are mapped into the feature space, the distance between the query and support sets are measured, and the classification is completed using the nearest neighbors concept.	Siamese Network <a href="#">[23]</a>	Siamese network contains two parallel neural networks that are trained to extract pair-wise sample features, and the Euclidean distance between features are measured.
		Matching Network <a href="#">[13]</a>	It generates a weighted K-nearest neighbor classifier based on the cosine distance between sample features.
		Relation Network <a href="#">[21]</a>	Different from the Siamese network and Matching network which adopt a single and fixed metric, Relation network compares relations with a nonlinear metric learned by a neural network.
		MsKPRN <a href="#">[24]</a>	MsKPRN extends the Relation Network to be position-aware and integrates multi-scale features.
		MSFN <a href="#">[25]</a>	MSFN learns a multi-scale feature space and similarities between the multi-scale and class representation are computed.
		Adaptive Metric Learning Model <a href="#">[26]</a>	Yu et al. <a href="#">[26]</a> proposed an adaptive metric learning model that is able to automatically determine the best weighted combination for emerging few-shot tasks from a set of metrics obtained by meta-learning.
		Knowledge-Guided Relation Network <a href="#">[27]</a>	Sui et al. <a href="#">[27]</a> proposed a knowledge-guided metric model that uses external knowledge to imitate human knowledge and generate relational networks that

Type	Core Idea	Method	Description
Other Methods	There is no unified core idea for these methods. They solve few-shot tasks in different ways, but all achieve competitive results.		can apply different metrics to different tasks.
		BERT-PAIR [28]	BERT-PAIR combines query and each support sample into a sequence and utilizes BERT [29] to predict whether each pair expresses the same class.
		LsSAML [30]	It utilizes the information implied by class labels to assist pretrained language models extracting more discriminative features.
		SALNet [31]	This method trains a classifier from labeled data through an attention mechanism and collects lexicons containing important words for each category, and then uses new data labeled by the combination of classifiers and lexicons to guide the learning of the classifier.

efficiently  
with the  
ole 2.

Model	Description
Simple GNN [32]	Garcia et al. [32] constructed a graph model in which the query and all support samples are closely connected and used a node-focused GNN to transfer instance-level relations and label information.
TPN [33]	This method further considers the relations among query samples.
EGNN [34]	It adopts an edge-labeling framework to explicitly model the intra-class similarity and inter-class dissimilarity of samples, and dynamically update node and edge features to achieve complex information interactions.

However, these models are all designed for image classification tasks, and these methods only transfer instance-level relations in GNNs, which make it difficult to handle elusive NLP tasks. In contrast, the HFGNN model proposed in this entry considers relations between samples from multiple perspectives, and the accurate and sufficient relations help the model construct more discriminative features.

### 3. Multi-Criteria Decision-Making

Zadeh [35] proposed fuzzy set theory to address problems related to fuzzy, subjective, and imprecise judgments. However, this theory lacks the ability to solve the problem of multi-criteria decision-making(MCDM). In this regard, Torra [12] proposed HFS, which determines the corresponding evaluation index and membership function according to the different attributes of the elements in the universe. HFS is a powerful tool for solving problems involving many uncertainties.

In recent years, some more efficient MCDM methods have been proposed. Deveci et al. [36] explored a novel approach that integrates Combined Compromise Solution (CoCoSo) with the context of type-2 neutrosophic

numbers to overcome the challenging decision process in Urban freight transportation tasks. Pamucar et al. [37] developed a novel integrated decision-making model which is based on Measuring Attractiveness by a Categorical Based Evaluation TecHnique (MACBETH) for calculating the criteria weights and Weight Aggregated Sum Product ASsessment (WASPAS) methods under the fuzzy environment with Dombi norms.

Considering the operating efficiency of the graph neural network model and the simplicity and effectiveness of the HFS theory, researchers introduce the HFS theory instead of other complex MCDM methods into the dual graph neural networks to fuse the relations between few-shot examples.

## References

1. Johnson, R.; Zhang, T. Deep pyramid convolutional neural networks for text categorization. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Vancouver, BC, Canada, 30 July–4 August 2017; pp. 562–570.
2. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
3. Kuang, S.; Li, J.; Branco, A.; Luo, W.; Xiong, D. Attention Focusing for Neural Machine Translation by Bridging Source and Target Embeddings. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia, 15–20 July 2018; pp. 1767–1776.
4. Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a few examples: A survey on few-shot learning. *ACM Comput. Surv.* 2020, 53, 1–34.
5. Snell, J.; Swersky, K.; Zemel, R. Prototypical networks for few-shot learning. *Adv. Neural Inf. Process. Syst.* 2017, 30, 2153.
6. Ye, Z.X.; Ling, Z.H. Multi-Level Matching and Aggregation Network for Few-Shot Relation Classification. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 2019, 28 July–2 August; pp. 2872–2881.
7. Gao, T.; Han, X.; Liu, Z.; Sun, M. Hybrid attention-based prototypical networks for noisy few-shot relation classification. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 6407–6414.
8. Sun, S.; Sun, Q.; Zhou, K.; Lv, T. Hierarchical attention prototypical networks for few-shot text classification. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 476–485.

9. Geng, R.; Li, B.; Li, Y.; Zhu, X.; Jian, P.; Sun, J. Induction Networks for Few-Shot Text Classification. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 3904–3913.
10. Sabour, S.; Frosst, N.; Hinton, G.E. Dynamic routing between capsules. *Adv. Neural Inf. Process. Syst.* 2017, 30, 2100.
11. Yang, L.; Li, L.; Zhang, Z.; Zhou, X.; Zhou, E.; Liu, Y. Dpgn: Distribution propagation graph network for few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 13390–13399.
12. Torra, V. Hesitant fuzzy sets. *Int. J. Intell. Syst.* 2010, 25, 529–539.
13. Vinyals, O.; Blundell, C.; Lillicrap, T.; Wierstra, D. Matching networks for one shot learning. *Adv. Neural Inf. Process. Syst.* 2016, 29, 1804.
14. Donahue, J.; Jia, Y.; Vinyals, O.; Hoffman, J.; Zhang, N.; Tzeng, E.; Darrell, T. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 21–26 June 2014; pp. 647–655.
15. Salamon, J.; Bello, J.P. Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Process. Lett.* 2017, 24, 279–283.
16. Zhang, R.; Che, T.; Ghahramani, Z.; Bengio, Y.; Song, Y. Metagan: An adversarial approach to few-shot learning. *Adv. Neural Inf. Process. Syst.* 2018, 31, 1207.
17. Sun, Q.; Liu, Y.; Chua, T.-S.; Schiele, B. Meta-transfer learning for few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 403–412.
18. Finn, C.; Abbeel, P.; Levine, S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. In Proceedings of the 34th International Conference on Machine Learning, Sydney, NSW, Australia, 6–11 August 2017; pp. 1126–1135.
19. Ravi, S.; Larochelle, H. Optimization as a model for few-shot learning. In Proceedings of the 5th International Conference on Learning Representation (ICLR 2017), Toulon, French, 24–26 April 2017; pp. 1–11.
20. Mishra, N.; Rohaninejad, M.; Chen, X.; Abbeel, P. A Simple Neural Attentive Meta-Learner. In Proceedings of the 6th International Conference on Learning Representations (ICLR 2018), Vancouver, BC, Canada, 30 April–3 May 2018; pp. 1–17.
21. Sung, F.; Yang, Y.; Zhang, L.; Xiang, T.; Torr, P.H.; Hospedales, T.M. Learning to compare: Relation network for few-shot learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake, UT, USA, 18–23 June 2018; pp. 1199–1208.

22. Jiang, X.; Havaei, M.; Chartrand, G.; Chouaib, H.; Vincent, T.; Jesson, A.; Chapados, N.; Matwin, S. Attentive task-agnostic meta-learning for few-shot text classification. In Proceedings of the 7th International Conference on Learning Representations (ICLR 2019), New Orleans, LA, USA, 6–9 May 2019; pp. 1–14.
23. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In Proceedings of the Conference and Workshop on the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015.
24. Abdelaziz, M.; Zhang, Z. Multi-scale kronecker-product relation networks for few-shot learning. *Multimed. Tools Appl.* 2022, 81, 6703–6722.
25. Han, M.; Wang, R.; Yang, J.; Xue, L.; Hu, M. Multi-scale feature network for few-shot learning. *Multimed. Tools Appl.* 2020, 79, 11617–11637.
26. Yu, M.; Guo, X.; Yi, J.; Chang, S.; Potdar, S.; Cheng, Y.; Tesauro, G.; Wang, H.; Zhou, B. Diverse Few-Shot Text Classification with Multiple Metrics. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), New Orleans, LA, USA, 1–6 June 2018; pp. 1206–1215.
27. Sui, D.; Chen, Y.; Mao, B.; Qiu, D.; Liu, K.; Zhao, J. Knowledge Guided Metric Learning for Few-Shot Text Classification. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; pp. 3266–3271.
28. Gao, T.; Han, X.; Zhu, H.; Liu, Z.; Li, P.; Sun, M.; Zhou, J. FewRel 2.0: Towards More Challenging Few-Shot Relation Classification. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 6250–6255.
29. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. 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, Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186.
30. Luo, Q.; Liu, L.; Lin, Y.; Zhang, W. Don't miss the labels: Label-semantic augmented meta-learner for few-shot text classification. In Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Online, 1–6 August 2021; pp. 2773–2782.
31. Lee, J.-H.; Ko, S.-K.; Han, Y.-S. Salnet: Semi-supervised few-shot text classification with attention-based lexicon construction. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 2–9 February 2021; pp. 13189–13197.
32. Garcia, V.; Bruna, J. Few-shot learning with graph neural networks. In Proceedings of the 6th International Conference on Learning Representations (ICLR 2018), Vancouver, BC, Canada, 30

April–3 May 2018; pp. 1–13.

33. Liu, Y.; Lee, J.; Park, M.; Kim, S.; Yang, E.; Hwang, S.J.; Yang, Y. Learning to propagate labels: Transductive propagation network for few-shot learning. In Proceedings of the 7th International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019; pp. 1–11.
34. Kim, J.; Kim, T.; Kim, S.; Yoo, C.D. Edge-labeling graph neural network for few-shot learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 11–20.
35. Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst.* 1978, 1, 3–28.
36. Deveci, M.; Pamucar, D.; Gokasar, I.; Delen, D.; Wu, Q.; Simic, V. An analytics approach to decision alternative prioritization for zero-emission zone logistics. *J. Bus. Res.* 2022, 146, 554–570.
37. Pamucar, D.; Torkayesh, A.E.; Deveci, M.; Simic, V. Recovery center selection for end-of-life automotive lithium-ion batteries using an integrated fuzzy WASPAS approach. *Expert Syst. Appl.* 2022, 206, 117827.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/78050>