FinTech Industrial Project Planning Form

Student Name & ID: DU Jiaxing, 1155204280

Project title: Graph Neural Network Implementation

Sponsoring Organisation: CEFAR

Industrial Supervisor: Cobi XU, A.I. Phoenix Technology Co

Instructions: This form is to specify the tasks the student plans to undertake for the FinTech Project at the beginning of a semester. Please clearly specify the nature of the task and the expected achievements up to the midterm and the final. Please complete the form using Microsoft Word and add rows if necessary.

|  |  |  |
| --- | --- | --- |
| Task | Expected achievements by the midterm | Expected achievements by the final |
| Understand the framework of GNN | 1. Grasp standard computational modules (e.g., conv/recurrent, sampling, pooling operators) 2. Identify applicable scenarios of models, their advantages and disadvantages 3. Understand heterogeneous graph classifications 4. Able to perform different classification tasks (nodes, edges, graphs classifiers) by Pytorch | 1. Read related literature based on the problems encountered in practical operations, attempt to find motivations and inspiration 2. Extend to more complicated framework (e.g., hypergraph classification) |
| Implement on real datasets | 1. Perform different classification tasks using existing GNN models, focusing on *torch\_geometric.datasets* 2. Evaluate the effectiveness of popular GNN variants on real datasets | 1. Able to handle raw datasets (e.g., pre-process and cleaning, encoding into numeric matrix) 2. Develop proper GNN models, focus on heterogeneous graph tasks 3. Formulate new challenge problems in real-world applications (e.g., Fintech) |
| Improve model performance and efficiency | 1. Learn common tricks for handling large-scale datasets (e.g., subsampling, sparse computation) 2. Understand particular layers (e.g., aggregation, pooling) effect on performance 3. Understand regularization techniques (e.g., dropout, normalization) to mitigate overfitting | 1. Perform experimental analysis for a better understanding of the applicable scenarios 2. Able to tune hyperparameters for model adjustment 3. Able to employ techniques like parallel processing and federated learning for large-scale datasets |
| Concepts of Federate learning | 1. Understand challenges faced when handling non-IID and unbalanced distributions 2. Gain knowledge on federate optimization algorithms in deep learning and their benefits | 1. Able to apply it to deep learning (e.g.. FedAvg algorithm), visualize the effect on communication costs, and model accuracy 2. Explore the application in Fintech area and design possible solutions for handling datasets that are privately sensitive and large-scale |