

## 6 APPENDIX A

For the few-shot setting, following the FedBPT approach, we randomly selected 100 samples ( $k=100$ ) for each class in the datasets. For datasets with a large original validation set, such as QQP, following BDPL, we sampled 1,000 instances for validation.

For the federated learning setting, under Independent and Identically Distributed (IID) data conditions, we ensured uniform distribution of both the quantity and categories of data across each client. For Non-Independent and Identically Distributed (Non-IID) data conditions, we applied the Dirichlet distribution with hyperparameter  $\alpha$  to control the data sparsity across different clients. The smaller the value of  $\alpha$ , the more imbalanced the data across clients becomes. In this study, we set  $\alpha = 0.1$ .

For prompt tuning, we followed commonly used manual templates and set the number of prompt tokens to 4. The number of candidate prompt tokens was set to 100.

## 7 APPENDIX B

The three types of prompts—concise, moderate, and detailed formats—across each dataset:

SST-2:

- "judge the sentiment."
- "judge the sentiment, negative or positive."
- "Please judge the sentiment of the User's words, your answer should be negative or positive."

RTE:

- "Does sentence1 entail sentence2?"
- "Evaluate the relationship between these sentences. Reply with either entailment or not\_entailment."
- "Consider the logical connection between these two statements. Does the information presented in the first statement lead to the conclusion of the second statement? If so, answer entailment. If not, respond with not\_entailment."

QNLI:

- "Does the sentence entail the question?"
- "Can the information in the sentence provide an answer to the question? Respond with entailment or not\_entailment."
- "Please analyze the relationship between the question and the sentence. Does the content of the sentence provide a valid answer to the question posed? If yes, respond with entailment. If not, respond with not\_entailment."

CoLA:

- "Is the sentence grammatically acceptable?"
- "Evaluate the grammatical correctness of the following sentence. Respond with acceptable if it is grammatically correct, otherwise respond with unacceptable."
- "Please analyze the following sentence for grammatical acceptability. If the sentence is grammatically well-formed and makes sense, answer acceptable. If the sentence is not grammatically correct or does not make sense, respond with unacceptable."

WNLI:

- "Does the first sentence entail the second?"

- 702
- 703 • "Read the two sentences and determine whether the second sentence logically follows from  
704 the first. Respond with entailment or not\_entailment ."
  - 705 • "Analyze the relationship between the two sentences provided. Does the meaning of the  
706 first sentence lead to the meaning of the second sentence? If so, respond with entailment .  
707 If not, respond with not\_entailment ."

708 QQP:

- 709
- 710 • "Are these two questions semantically the same?"
  - 711 • "Compare the two questions and decide if they are asking about the same thing in different  
712 words. Answer with duplicate if they are, otherwise respond with not\_duplicate ."
  - 713 • "Evaluate the meaning of the following two questions. If both questions essentially ask the  
714 same thing but are worded differently, reply with duplicate . If they ask different things,  
715 reply with not\_duplicate ."
- 716

717 MRPC:

- 718
- 719 • "Are the two sentences paraphrases of each other?"
  - 720 • "Read the two sentences. Do they convey the same meaning in different words? Respond  
721 with paraphrase or not\_paraphrase ."
  - 722 • "Please analyze the following two sentences and determine their semantic similarity. If the  
723 sentences have the same meaning but are worded differently, answer with paraphrase . If  
724 they differ in meaning, respond with not\_paraphrase ."
- 725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755