

2403A51266

AI ASSISTED CODING

Question: Objective: To explore and compare Zero-shot, One-shot, and Few-shot prompting techniques for classifying emails into predefined categories using a large language model (LLM).

Suppose that you work for a company that receives hundreds of customer emails daily. Management wants to automatically classify emails into categories like "Billing", "Technical Support", "Feedback", and "Others" before assigning them to appropriate departments. Instead of training a new model, your task is to use prompt engineering techniques with an existing LLM to handle the classification.

Tasks to be completed are as below

1. Prepare Sample Data:

- Create or collect 10 short email samples, each belonging to one of the 4 categories.

2. Zero-shot Prompting:

- Design a prompt that asks the LLM to classify a single email without providing any examples.
- Example prompt:

“Classify the following email into one of the following categories: Billing, Technical Support, Feedback, Others. Email: ‘I have not received my invoice or last month.’”

3. One-shot Prompting:

- Add one labeled example before asking the model to classify a new email.

4. Few-shot Prompting:

- Use 3–5 labeled examples in your prompt before asking the model to classify a new email.

5. Evaluation:

- Run all three techniques on the same set of 5 test emails.
- Compare and document the accuracy and clarity of responses.

Requirements:

- VS Code with Github Copilot or Cursor IDE and/or Google Colab with Gemini

Deliverables:

- A .txt or .md file showing prompts and model responses.
- A comparison table showing classification accuracy for each technique.
- A short reflection on which method was most effective and why

* 10 E-mail samples with categories:

1. "I haven't received my invoice for last month." – Billing
2. "My internet connection keeps dropping randomly." – Technical Support
3. "Great service! I'm very happy with your team." – Feedback
4. "Can you help me reset my password?" – Technical Support
5. "There's an unexpected charge on my bill." – Billing
6. "Just wanted to say thanks for the prompt response." – Feedback
7. "What are your working hours during the holidays?" – Others
8. "The app crashes every time I try to open it." – Technical Support
9. "Why was I charged twice this month?" – Billing
10. "Do you offer student discounts?" – Others

* Zero short prompting:

Classify the following email into one of the following categories: Billing, Technical Support, Feedback, Others.

Email: "I haven't received my invoice for last month."

- One-shot prompting:

Classify the following email into one of the following categories: Billing, Technical Support, Feedback, Others.

Example:

Email: "There's an unexpected charge on my bill."

Category: Billing

Now classify this email:

Email: "I haven't received my invoice for last month."

- Few-shot prompting:

Classify the following email into one of the following categories: Billing, Technical Support, Feedback, Others.

Examples:

Email: "There's an unexpected charge on my bill."

Category: Billing

Email: "My internet connection keeps dropping randomly."

Category: Technical Support

Email: "Great service! I'm very happy with your team."

Category: Feedback

Email: "Do you offer student discounts?"

Category: Others

Now classify this email:

Email: "I haven't received my invoice for last month."

- **Evaluation setup:**


Email (Test Set)	True Category
1. "I haven't received my invoice for last month."	Billing
2. "The app crashes every time I try to open it."	Technical Support
3. "Just wanted to say thanks for the prompt response."	Feedback
4. "Why was I charged twice this month?"	Billing

5. "What are your working hours during the holidays?"	Others
---	--------

• Output:

Zero-shot Prompting Results:

Email: "I haven't received my invoice for last month."

Response: Billing 
















Email: "The app crashes every time I try to open it."

Response: Technical Support 

Email: "Just wanted to say thanks for the prompt response."

Response: Others  (Expected: Feedback)

Comparison Table

	Email	True Category	Zero-shot	One-shot	Few-shot
1	Billing				
2	Technical Support				
3	Feedback				
4	Billing				
5	Others				

Summary

- **Zero-shot Accuracy**: 3/5

- **One-shot Accuracy**: 4/5

- **Few-shot Accuracy**: 5/5

Reflection

Few-shot prompting was the most effective in this experiment. By showing the model several labeled examples, it better understood the context and nuances of the categories. Zero-shot prompting performed decently but struggled with more subtle distinctions (e.g., between "Feedback" and "Others"). One-shot was better, but still missed edge cases.