In-Context Learning for Tabular Data

A Comparative Study on the Titanic Dataset

by @LLMsLab

Agenda

- Introduction
- Dataset Analysis
- Zero-Shot In-Context Learning
- Few-Shot In-Context Learning
- Supervised Learning
- Comparative Analysis
- Improvement Strategies
- Conclusion
- Questions and Discussions

Introduction

- Objective: To compare in-context and supervised learning approaches using the Titanic dataset.
- Dataset: Contains passenger information and survival status from the Titanic disaster.
- In-Context Learning: A method where a model leverages contextual information to generate predictions.
    - Zero-Shot Learning: Predictions on unseen data categories.
    - Few-Shot Learning: Predictions after seeing only a few examples.
- Supervised Learning: A method where the model is explicitly trained on a labeled dataset.

Dataset Analysis

- Dataset: Consists of passenger information from the Titanic disaster.
- Features: Include socio-economic class, sex, age, and family relations among others.
- Target Variable: 'Survived' indicating whether a passenger survived (1) or not (0).
- Balance: Dataset is balanced with approximately equal numbers of survived and not survived instances.

Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.25	NaN	S

Zero-Shot In-Context Learning

- Concept: Model makes predictions on unseen data categories without explicit training.
- Implementation: Utilized GPT-4 model with specific prompts based on passenger information.
- Balancing Strategy: Randomly sampled equal number of survived and not survived instances for the test set.

	Precision	Recall	F1 Score
Result	0.50	1.00	0.67

Zero-Shot In-Context Learning (Prompt Example)

The prompt structure used for the GPT-4 model was:

```
Based on the passenger information, is the passenger most likely to have survived or not survived the Titanic disaster?
Passenger Information:
{{feature1_name}}: {{feature1_value}}
{{feature2_name}}: {{feature2_value}}
...
{{featureN_name}}: {{featureN_value}}
Answer Choices: (A) Survived, (B) Not Survived.
```

Each `{{feature_name}}` was replaced by the feature's name (e.g., "Age", "Sex", "Pclass"), and each `{{feature_value}}` was replaced by the value of that feature for a specific passenger.

Few-Shot In-Context Learning

- Concept: Model makes predictions after seeing only a few examples.
- Implementation: Utilized GPT-4 model with specific prompts and a few examples of the task in the conversation history.
- Balancing Strategy: Similar to zero-shot learning, equal number of survived and not survived instances sampled for the test set.

	Precision	Recall	F1 Score
Result	-	-	-

Few-Shot In-Context Learning (Prompt Example)

The prompt structure used for the GPT-4 model was:

```
I am a model trained to predict passenger survival on the Titanic. Here are some examples of my training:
Example 1:
Question: Given the following passenger information, is the passenger likely to have survived or not survived the Titanic disaster?
Passenger Information: 'Class: 3, Sex: male, Age: 22'
Answer: Not Survived
...
Now, a new example to classify:
Question: Given the following passenger information, is the passenger likely to have survived or not survived the Titanic disaster?
Passenger Information: {{data_row}}
```

Each `{{data_row}}` was replaced by the features of a specific passenger.

Supervised Learning

- Concept: Model is explicitly trained on a labeled dataset.
- Implementation: Utilized XGBoost model with GridSearchCV for hyperparameter tuning and cross-validation.
- Feature Engineering: Created new features and transformed existing ones to enhance the model's predictive capability.

	Precision	Recall	F1 Score
Result	0.80	0.82	0.81

Comparative Analysis

- Objective: Evaluate and compare the efficacy of in-context learning and supervised learning techniques.

	Precision	**Recall**	**F1 Score**
Zero-Shot	0.50	1.00	0.67
Few-Shot	-	-	-
Supervised	0.80	0.82	0.81

Improvement Strategies

- Prompts: Experimenting with the wording and structure of prompts can lead to better results.
- Fine-Tuning: Fine-tuning the GPT-4 model on specific tasks can increase its performance.
- Data Preprocessing: Feature engineering and selection can improve the model's ability to extract useful patterns.
- Hyperparameter Tuning: Optimizing the hyperparameters of the model can enhance its predictive capability.

Conclusion

- All methods have their strengths and challenges.
- Zero-shot learning is quick to deploy but may not deliver the best accuracy.
- Few-shot learning attempts to improve accuracy by providing a few examples.
- Supervised learning, while requiring more setup and computation, can often provide the highest accuracy.
- The best approach depends on the specific use-case, available data, and resources.

Questions and Discussions

Thank you for your attention! Now, let's open the floor for any questions or discussions.