Prompt Studio

status indicator
(offline, online, saving)

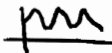## Error Messages

Prompt Designer

Experiment Configurator

Analysis

Open   Export   Settings

Control Panel
DB-Status :
Available LLMs :

Navbar

- Prompt Designer
- Prompt Template
- Test Prompt
- Open Prompt
- Export Config

EXPERIMENT

ANALYZER

| Error Msg |

Create prompt from existing version

Prompt name		version
A	▷	1
B	▷	2
C	▷	3
D	▷	4

Template preview

Create new prompt

Prompt name

[Must be unique]

Create

| ◁ Home |

| Template ▷ | EXPAND |

Navbar

- Prompt Designer
- Prompt Template
- Test Prompt
- Open Prompt
- Export Config
- EXPERIMENT
- ANALYZER

Error Msg

Prompt name: Label

Prompt Template

Marker Index

| Verify | Save | Save as new |

◁ Prompt Designer

| Test D | EXP D |

Navbar

- Prompt Designer
- Prompt Template
- Test Prompt
- Open Prompt
- Export config

EXPERIMENT

ANALYZER

Error Msg

Sample Input

[   ]   [Browse]   [Load]

Preview data

Prompt Preview

[   ]

[Save]

[◁◁ Prompt Designer]   [◁ Template]   [EXPAD]

Navbar

- Prompt Design
- **EXPERIMENT**
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Error Msg

Create from existing experiment

title	▽
A	STATUS
▷ B	complete
▷ C D	in progress
D	draft

[Create New]   [Edit]   [Duplicate]

[◁ Home]   [Prompt ▷]   [ANALYSIS ▷]

Error Msg

Navbar

- Prompt Designer
- EXPERIMENTS
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Select a prompt

Prompt name		Version
A	▷	1
B	▷	2
C	▷	3

Prompt Preview

☑ color markers

Assign Markers

Persona [SZ] ▽
num-shots [1] ▽
feature [Tn A x] ▽

Shot markers

☑ shot   ☑ scope
☐ feature ☐ topic

[Save]

Shot dataset

[_____] [Browse] [Load]

Preview data

O Fixed number   Number of shots
O First
O Random   0 —— 2 —— Max

[◁ Exp Configuration]

[Input ▷]   [ANALYSIS ▷▷]

Error Msg

Sample Input

[   ] [Browse]   [Load]

Navbar

- Prompt Designer
- EXPERIMENTS
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Preview data

Prompt Preview

[Save]

[◁◁ Exp Configuration]   [Model ▷]   [ANALYSIS ▷▷]

Error Msg

Navbar

- Prompt Designer
- EXPERIMENT
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Model Family

- OpenAI ▷
- GoogleAI ▷
- Mozilla AI ▷
- HuggingFaceHub
- Sci-kit-learn ▷
- Random ▷

Model Version

- gpt-3.5-turbo
- gpt-4o

Key

[   ]

Set Hyperparameters

○ Simple

Temperature

Max tokens

0.0 —— 1.5 —— 2.0

|——— 2000 ———— Max|

○ Advanced

Custom parameters

[key] [value] ⊕

[Save]

[◁ Exp Configuration]

[Evaluation ▷]   [ANALYSIS ▷]

Error Msg

Navbar

- Prompt Design
- **EXPERIMENT**
- Prompt Config
- Input dataset
- **Model Selection**
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Model Family

- Open AI ▷
- Google AI ▷
- Mozilla AI ▷
- HuggingFaceHub
- Sci-Kit-learn ▷
- Random ▷

Model Category

- Linear ▷
- Naive Bayes ▷
- SupportVec ▷
- Ensemble ▷

Model Name

- Complaint NB
- Multinomial NB

Set Hyperparameters

[Key] [Value] ⊕

Training Method

- Sample ratio
- Vectorization
- Cross Validation
- Folding

0%   70% 100%

[▽]
[◁]
[▽]

Custom training

[Save]

◁ Exp Configuration

Evaluation ▷

ANALYSIS ▷

Navbar

- Prompt Design
- **EXPERIMENT**
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- **ANALYZER**

Error Msg

[Test]

Output

Data Pre-processing
Explanation

```
def preprocess(data frame):
    pass
```

Output post-processing
Explanation

```
def postprocess(output, result):
    return output
```

Result evaluation
Explanation

```
def evaluate(output, oracle):
    return output == oracle
```

[Save]

[<< Exp Configuration]

[Validation >]   [ANALYSIS >>]

Error Msg

Navbar

- Prompt Design
- EXPERIMENTS
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- Setup
- Export Config
- ANALYZER

Self Validation Prompt ☑ Default

Default

Marker index ☑ Default

Default

Number of iterations [1] [▲▼]

Save

Choose the Model

☐ Same as experiment

Family   Model

Key [_____]

[◀ Exp Configuration]   [Setup ▷]   [ANALYSIS ▷▷]

Error Msg

Navbar

- Prompt Designer
- **EXPERIMENT**
- Prompt Config
- Input dataset
- Model Selection
- Evaluation
- Validation
- **Setup**
- Export Config
- ANALYZER

Configuration tag

Experiment title

Number of iterations	1 ⬆⬇
Self-validation	☑
Synchronous execution	☑
Number of threads	1 ⬆⬇
Batch mode	☑
Stop upon error	☑
Retry upon error	☑
Max retries	1 ⬆⬇
Tokens per day	1 ⬆⬇

RUN

◁ Exp Configuration

ANALYSIS ▷

Error Msg

Navbar

- Prompt Design
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Experiment status ☑ completed ☐ in progress

▽ Filter

⊕ Criterion ▽   OPERATOR ▽   Value ▽

Sort by Title   Sort by Date

Experiment titles
title A
title B
title C

Analyze

◁ Home   Dashboard ▷

Error Msg

Navbar

- Prompt Design
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Interpretation Prompt ☑ Default

Default

Marker index   ☐ Default

Default

Choose the Model

☐ Same as experiment

Family   Model

Key

Save

◁◁ Exp Configuration

Dashboard ▷   ANALYSIS ▷▷

Error Msg

Navbar

- Prompt Design
- EXPERIMENT
- **ANALYSIS**
- Interpretation
- **Dashboard**
- Results
- Statistics
- Compare
- Query
- Export CSV

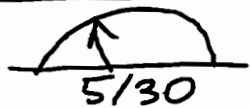Experiment title [status]   ☑ Live Monitor

Execution time Dd Hh ms
YY-MM-DD H:m:s ——>YY-MM-DD H:m:s
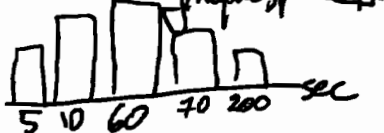
Data processed
652/1000

Iterations completed

5/30

Tokens used
4096

Requests
Total: 1002
Ratio: 1.05
Success ful: 702
Errors: 300

Latency/Request

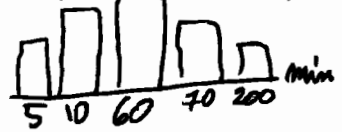5 10 60 70 200 sec

Tokens/Request

10 20 100 400 1000

Retries 62.4

5 10 60 70 200

Requests/min 62.4

5 10 60 70 200 min

Tokens/min 62.4

5 10 60 70 200 min

Error messages

Msg1	73
Msg2	51
Msg3	17

◁ Analysis   Results ▷

Error Msg

Navbar

- Prompt Designer
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Experiment title [status]

Processed data

652/1000

Errors

348/1000

Results

A B C D E
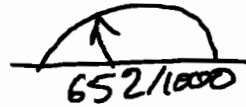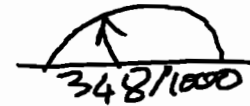
Confusion Matrix

	T	F
P	100	53
N	252	17

◁◁ Analysis

Statistics ▷

Error Msg

Navbar

- Prompt Designer
- EXPERIMENT
- **ANALYSIS**
- Interpretation
- Dashboard
- Results
- **Statistics**
- Compare
- Query
- Export CSV

Experiment title [status]

Metrics

- ☑ Accuracy
- ☑ Recall
- ☑ Precision

Accuracy	98%
AUC	50.01%
bAcc	97%
Kappa	
MCC	
Pre	
Rec	

Radar chart axes: Acc, bacc, AUC, K, MCC, Pre, Rec, Spec
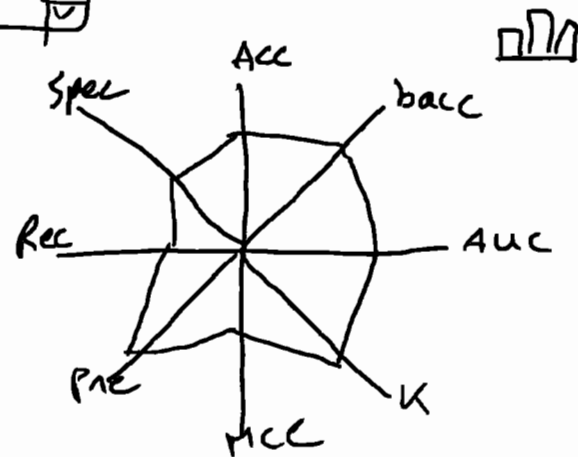
◁ Analysis

Compare ▷

Error Msg

Experiment title [status]

Navbar

- Prompt Designer
- **EXPERIMENT**
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Metrics

- ☑ Accuracy
- ☑ Recall
- ☐ Precision

Aggregator

Mean
Std

Metric	Mean	Std	Med	IQR	Sk	Kurt
Acc						
AUC						
Rec						
Spec						

Fleiss Kappa 75%

◁ Analysis

Compare ▷

Error Msg

Navbar

- Prompt Designer
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Experiments

▽ Filter

⊕ | Criteria ▽ | OPERATOR ▽ | Value ▽ |

| Sort by Title | Sort by Date |

Experiment titles
☐ Title A
☐ Title B
☐ Title C

Metrics

☐ Accuracy
☐ Recall
☐ Precision

Series

| Dataset ▽ | ▽ |
| Model |
| Prompt |

| ◁ Analysis |

Series	Accuracy	Recall
A	98%	85%
B	96%	87%
C	91%	90%
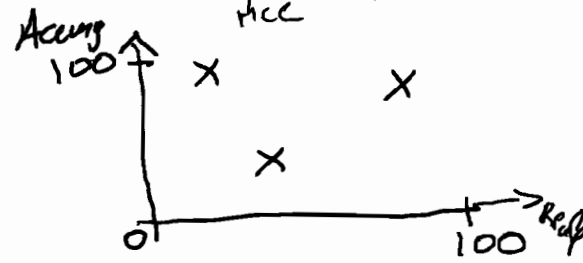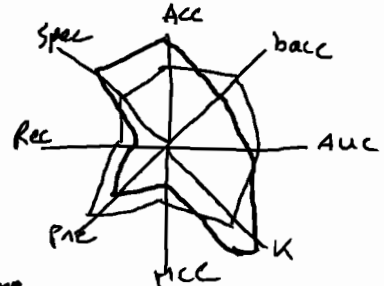
Radar chart with axes: Acc, bacc, auc, K, mcc, pre, Rec, spe

Accur 100 ↑ ... → Recl 100 (scatter plot with X marks)

Compare

Query ▷

Error Msg

Experiment title [status]

Query the Result table using MySQL syntax

```
Select *
From Result
Where error is null
```

Navbar

- Prompt Design
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

Exp	It	Dataset	Prompt	Model	Oracle ---

◁ Analysis's

Export ▷

Error Msg

Experiment title [status]

Navbar

- Prompt Design
- EXPERIMENT
- ANALYSIS
- Interpretation
- Dashboard
- Results
- Statistics
- Compare
- Query
- Export CSV

◁ Analysis