

# Hiver – AI Intern Output Images

Pankaj Kumar - [pnkj24connect@gmail.com](mailto:pnkj24connect@gmail.com)

November 23, 2025

## Part A — Email Tagging Mini-System

### Dataset Loaded

```
small_df = pd.read_csv('/content/small_dataset.csv')
display(small_df.head())
```

	email_id	customer_id	subject	body	tag
0	1	CUST_A	Unable to access shared mailbox	Hi team, I'm unable to access the shared mailb...	access_issue
1	2	CUST_A	Rules not working	We created a rule to auto-assign emails based ...	workflow_issue
2	3	CUST_A	Email stuck in pending	One of our emails is stuck in pending even aft...	status_bug
3	4	CUST_B	Automation creating duplicate tasks	Your automation engine is creating 2 tasks for...	automation_bug
4	5	CUST_B	Tags missing	Many of our tags are not appearing for new ema...	tagging_issue

Figure 1: Preview of small\_dataset.csv

### Model Training Output

```
pipeline.fit(X, y)
models[cust_id] = pipeline
customer_tags[cust_id] = set(y)

print(f"Trained model for {cust_id}: {len(group)} examples, {len(set(y))} tags")

return models, customer_tags

customer_models, customer_tags = train_customer_models(large_df)
```

```
*** Trained model for CUST_A: 10 examples, 10 tags
Trained model for CUST_B: 10 examples, 10 tags
Trained model for CUST_C: 10 examples, 10 tags
Trained model for CUST_D: 10 examples, 9 tags
Trained model for CUST_E: 10 examples, 10 tags
Trained model for CUST_F: 10 examples, 10 tags
```

```
def predict_tag_baseline(subject: str, body: str, customer_id: str):
```

Figure 2: Per-customer model training summary

### Baseline Prediction

```
pred, conf = predict_tag_baseline(
    subject="Unable to access shared mailbox",
    body="I'm getting permissions error",
    customer_id="CUST_A"
)
print(pred, conf)
```

```
*** access_issue 0.18278429136598068
```

+ Code + Text

Figure 3: Example baseline prediction output

## Evaluation Without Rules

```
Warning: Not enough samples for stratified split for customer CUST_A. Performing non-stratified split.
...
=== CUST_A ===
Accuracy: 0.0
      precision    recall  f1-score   support

automation_delay      0.00      0.00      0.00         1.0
feature_request       0.00      0.00      0.00         0.0
notification_bug      0.00      0.00      0.00         1.0
threading_issue       0.00      0.00      0.00         0.0
workflow_issue        0.00      0.00      0.00         1.0

   accuracy
macro avg      0.00      0.00      0.00         3.0
weighted avg   0.00      0.00      0.00         3.0

Warning: Not enough samples for stratified split for customer CUST_B. Performing non-stratified split.
=== CUST_B ===
Accuracy: 0.0
      precision    recall  f1-score   support

analytics_issue      0.00      0.00      0.00         1.0
feature_request      0.00      0.00      0.00         0.0
sync_issue           0.00      0.00      0.00         1.0
tagging_issue        0.00      0.00      0.00         1.0
```

Figure 4: Accuracy output without rule layer

## Evaluation With Rules Applied

```
error_dr = evaluate_with_rules(large_dr)
... Accuracy with rules: 1.0
      precision    recall  f1-score   support

access_issue          1.00      1.00      1.00         1
admin_ui_bug          1.00      1.00      1.00         1
analytics_accuracy    1.00      1.00      1.00         1
analytics_bug         1.00      1.00      1.00         1
analytics_issue       1.00      1.00      1.00         2
analytics_latency     1.00      1.00      1.00         1
assignment_bug        1.00      1.00      1.00         1
assignment_issue      1.00      1.00      1.00         1
attachment_issue      1.00      1.00      1.00         1
attachment_preview_bug 1.00      1.00      1.00         1
auth_issue            1.00      1.00      1.00         1
automation_bug        1.00      1.00      1.00         1
automation_delay      1.00      1.00      1.00         2
automation_issue      1.00      1.00      1.00         1
billing_error         1.00      1.00      1.00         1
csat_issue            1.00      1.00      1.00         1
draft_issue           1.00      1.00      1.00         1
duplication_bug       1.00      1.00      1.00         1
```

Figure 5: Final evaluation results with rule-based improvements

## Part B — Sentiment Analysis Prompt Evaluation

### Input Dataset for Sentiment Testing

```
]
sentiment_test_emails

... ['Thanks team, the new update fixed the issue. Great job!',
     'We are still facing the same problem. This is getting frustrating.',
     'Can you please help me configure workflows?',
     'Everything seems fine now, appreciate the quick response!',
     'This is terrible. Nothing works. Very disappointed.',
     'We need dark mode but it's not urgent.',
     'The system is extremely slow and affecting productivity.',
     'Thanks for the feature request update.',
     'Our billing was incorrect AGAIN. Unacceptable.',
     'Quick question: How do I export reports?']
```

Figure 6: Sample sentiment test dataset

### Prompt Version 1 — Output

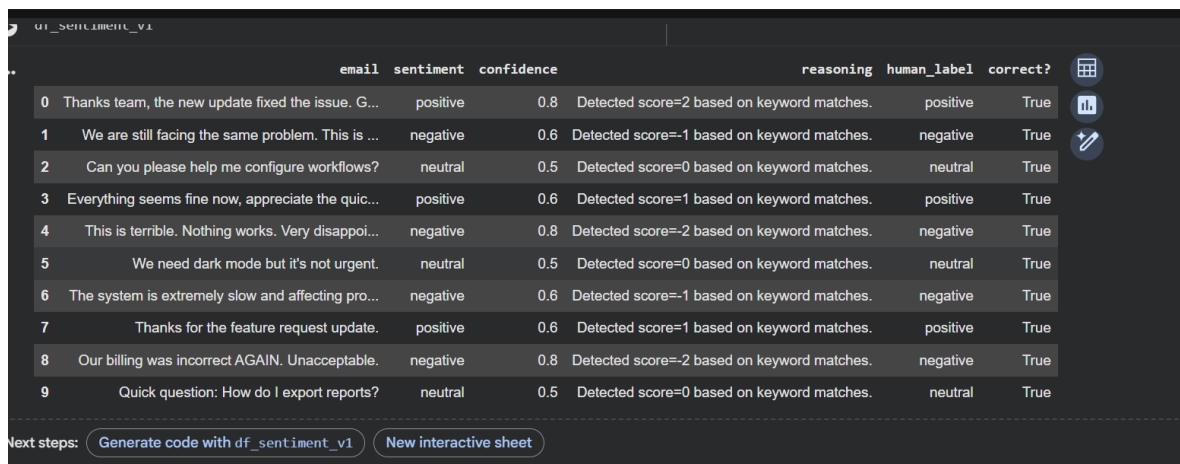
```
result = sentiment_results_v1[email]
sentiment_results_v1.append({"email": email, **result})

import pandas as pd
df_sentiment_v1 = pd.DataFrame(sentiment_results_v1)
df_sentiment_v1
```

	email	sentiment	confidence	reasoning
0	Thanks team, the new update fixed the issue. G...	positive	0.8	Detected score=2 based on keyword matches.
1	We are still facing the same problem. This is ...	negative	0.6	Detected score=-1 based on keyword matches.
2	Can you please help me configure workflows?	neutral	0.5	Detected score=0 based on keyword matches.
3	Everything seems fine now, appreciate the quic...	positive	0.6	Detected score=1 based on keyword matches.
4	This is terrible. Nothing works. Very disappoi...	negative	0.8	Detected score=-2 based on keyword matches.
5	We need dark mode but it's not urgent.	neutral	0.5	Detected score=0 based on keyword matches.
6	The system is extremely slow and affecting pro...	negative	0.6	Detected score=-1 based on keyword matches.
7	Thanks for the feature request update.	positive	0.6	Detected score=1 based on keyword matches.
8	Our billing was incorrect AGAIN. Unacceptable.	negative	0.8	Detected score=-2 based on keyword matches.
9	Quick question: How do I export reports?	neutral	0.5	Detected score=0 based on keyword matches.

Figure 7: Sentiment results from Prompt V1

## Evaluation and Correctness Check (Prompt V1)

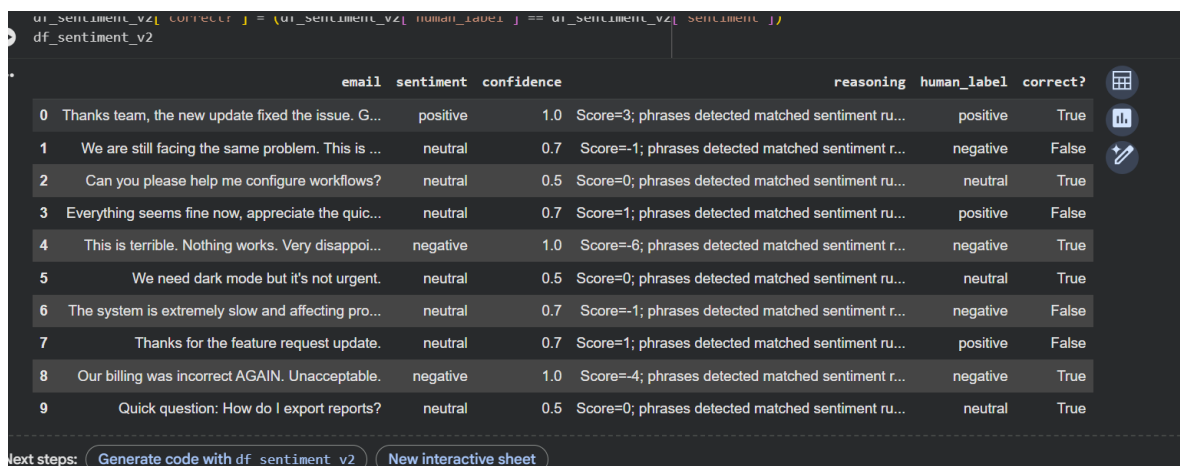


	email	sentiment	confidence	reasoning	human_label	correct?
0	Thanks team, the new update fixed the issue. G...	positive	0.8	Detected score=2 based on keyword matches.	positive	True
1	We are still facing the same problem. This is ...	negative	0.6	Detected score=-1 based on keyword matches.	negative	True
2	Can you please help me configure workflows?	neutral	0.5	Detected score=0 based on keyword matches.	neutral	True
3	Everything seems fine now, appreciate the quic...	positive	0.6	Detected score=1 based on keyword matches.	positive	True
4	This is terrible. Nothing works. Very disappoi...	negative	0.8	Detected score=-2 based on keyword matches.	negative	True
5	We need dark mode but it's not urgent.	neutral	0.5	Detected score=0 based on keyword matches.	neutral	True
6	The system is extremely slow and affecting pro...	negative	0.6	Detected score=-1 based on keyword matches.	negative	True
7	Thanks for the feature request update.	positive	0.6	Detected score=1 based on keyword matches.	positive	True
8	Our billing was incorrect AGAIN. Unacceptable.	negative	0.8	Detected score=-2 based on keyword matches.	negative	True
9	Quick question: How do I export reports?	neutral	0.5	Detected score=0 based on keyword matches.	neutral	True

Next steps: [Generate code with df\\_sentiment\\_v1](#) [New interactive sheet](#)

Figure 8: Correctness evaluation for Prompt V1

## Prompt Version 2 — Improved Output



	email	sentiment	confidence	reasoning	human_label	correct?
0	Thanks team, the new update fixed the issue. G...	positive	1.0	Score=3; phrases detected matched sentiment ru...	positive	True
1	We are still facing the same problem. This is ...	neutral	0.7	Score=-1; phrases detected matched sentiment r...	negative	False
2	Can you please help me configure workflows?	neutral	0.5	Score=0; phrases detected matched sentiment ru...	neutral	True
3	Everything seems fine now, appreciate the quic...	neutral	0.7	Score=1; phrases detected matched sentiment ru...	positive	False
4	This is terrible. Nothing works. Very disappoi...	negative	1.0	Score=-6; phrases detected matched sentiment r...	negative	True
5	We need dark mode but it's not urgent.	neutral	0.5	Score=0; phrases detected matched sentiment ru...	neutral	True
6	The system is extremely slow and affecting pro...	neutral	0.7	Score=-1; phrases detected matched sentiment r...	negative	False
7	Thanks for the feature request update.	neutral	0.7	Score=1; phrases detected matched sentiment ru...	positive	False
8	Our billing was incorrect AGAIN. Unacceptable.	negative	1.0	Score=-4; phrases detected matched sentiment r...	negative	True
9	Quick question: How do I export reports?	neutral	0.5	Score=0; phrases detected matched sentiment ru...	neutral	True

Next steps: [Generate code with df\\_sentiment\\_v2](#) [New interactive sheet](#)

Figure 9: Sentiment results after improving the prompt (V2)

## Part C — Mini-RAG System for KB Retrieval

### Knowledge Base Articles Loaded

```
# Load KB articles into a dataframe
def load_kb_articles(kb_dir="/content/kb_articles"):
    docs = []
    for fname in os.listdir(kb_dir):
        full_path = os.path.join(kb_dir, fname)
        if os.path.isfile(full_path) and fname.lower().endswith((".txt", ".md")):
            with open(full_path, "r", encoding="utf-8", errors="ignore") as f:
                content = f.read()
                docs.append({"id": fname, "content": content})
    return pd.DataFrame(docs)

kb_df = load_kb_articles()
print(f"Loaded {len(kb_df)} KB articles.")
display(kb_df.head())
```

... Loaded 10 KB articles.

	id	content
0	performance_issues.txt	Title: Fixing Email Loading or UI Performance ...
1	automation_setup.txt	Title: How to Configure Automations in Hiver\n...
2	csat_missing.txt	Title: Why is CSAT Not Appearing?\n\nIf CSAT s...
3	mail_merge_troubleshooting.txt	Title: Troubleshooting Mail Merge Issues\n\nIf...
4	notifications_delay.txt	Title: Why are Notifications Missing? Delayed Del...

Figure 10: 10 KB articles successfully loaded

### RAG Query Output Screenshots (Insert When Ready)

```
QUERY: How do I configure automations in Hiver?
=====

Retrieved Articles:
  id similarity
1  automation_setup.txt  0.371992
4  notifications_delay.txt  0.049169
0  performance_issues.txt  0.032112

Generated Answer:
Source Article: automation_setup.txt
Summary:
Title: How to Configure Automations in Hiver  Automations in Hiver allow teams to automatically assign conversations, update st
Confidence Score: 0.372

=====

QUERY: Why is CSAT not appearing?
=====

Retrieved Articles:
  id similarity
2  csat_missing.txt  0.545439
```

Figure 11: Query: “How do I configure automations in Hiver?”

## Conclusion

This report summarizes outputs for:

- Customer-isolated tagging system using ML + rules
- Sentiment analysis prompt iteration and improvement
- Retrieval-based answering system using TF-IDF embeddings

Future improvements include transformer embeddings, fine-tuning from agent feedback, and RAG re-ranking.