compliance_email_classifier

July 4, 2025

[1]: `pip install pandas scikit-learn shap matplotlib`

```
Requirement already satisfied: pandas in /opt/conda/lib/python3.12/site-packages
(2.2.3)
Requirement already satisfied: scikit-learn in /opt/conda/lib/python3.12/site-
packages (1.6.0)
Collecting shap
  Downloading
shap-0.48.0-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(25 kB)
Requirement already satisfied: matplotlib in /opt/conda/lib/python3.12/site-
packages (3.10.0)
Requirement already satisfied: numpy>=1.26.0 in ./.local/lib/python3.12/site-
packages (from pandas) (1.26.4)
Requirement already satisfied: python-dateutil>=2.8.2 in
/opt/conda/lib/python3.12/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.12/site-
packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in /opt/conda/lib/python3.12/site-
packages (from pandas) (2024.2)
Requirement already satisfied: scipy>=1.6.0 in /opt/conda/lib/python3.12/site-
packages (from scikit-learn) (1.14.1)
Requirement already satisfied: joblib>=1.2.0 in /opt/conda/lib/python3.12/site-
packages (from scikit-learn) (1.4.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/opt/conda/lib/python3.12/site-packages (from scikit-learn) (3.5.0)
Requirement already satisfied: tqdm>=4.27.0 in /opt/conda/lib/python3.12/site-
packages (from shap) (4.67.1)
Requirement already satisfied: packaging>20.9 in /opt/conda/lib/python3.12/site-
packages (from shap) (24.2)
Collecting slicer==0.0.8 (from shap)
  Downloading slicer-0.0.8-py3-none-any.whl.metadata (4.0 kB)
Requirement already satisfied: numba>=0.54 in /opt/conda/lib/python3.12/site-
packages (from shap) (0.60.0)
Requirement already satisfied: cloudpickle in /opt/conda/lib/python3.12/site-
packages (from shap) (3.1.0)
Requirement already satisfied: typing-extensions in
/opt/conda/lib/python3.12/site-packages (from shap) (4.12.2)
```

Requirement already satisfied: contourpy>=1.0.1 in
/opt/conda/lib/python3.12/site-packages (from matplotlib) (1.3.1)
Requirement already satisfied: cycler>=0.10 in /opt/conda/lib/python3.12/site-
packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
/opt/conda/lib/python3.12/site-packages (from matplotlib) (4.55.3)
Requirement already satisfied: kiwisolver>=1.3.1 in
/opt/conda/lib/python3.12/site-packages (from matplotlib) (1.4.7)
Requirement already satisfied: pillow>=8 in /opt/conda/lib/python3.12/site-
packages (from matplotlib) (11.0.0)
Requirement already satisfied: pyparsing>=2.3.1 in
/opt/conda/lib/python3.12/site-packages (from matplotlib) (3.2.0)
Requirement already satisfied: llvmlite<0.44,>=0.43.0dev0 in
/opt/conda/lib/python3.12/site-packages (from numba>=0.54->shap) (0.43.0)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.12/site-
packages (from python-dateutil>=2.8.2->pandas) (1.17.0)
Downloading
shap-0.48.0-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.0 MB)
                             1.0/1.0 MB
33.4 MB/s eta 0:00:00
Downloading slicer-0.0.8-py3-none-any.whl (15 kB)
Installing collected packages: slicer, shap
Successfully installed shap-0.48.0 slicer-0.0.8
Note: you may need to restart the kernel to use updated packages.

[2]: pip install notebook

Requirement already satisfied: notebook in /opt/conda/lib/python3.12/site-
packages (7.3.2)
Requirement already satisfied: jupyter-server<3,>=2.4.0 in
/opt/conda/lib/python3.12/site-packages (from notebook) (2.15.0)
Requirement already satisfied: jupyterlab-server<3,>=2.27.1 in
/opt/conda/lib/python3.12/site-packages (from notebook) (2.27.3)
Requirement already satisfied: jupyterlab<4.4,>=4.3.4 in
/opt/conda/lib/python3.12/site-packages (from notebook) (4.3.4)
Requirement already satisfied: notebook-shim<0.3,>=0.2 in
/opt/conda/lib/python3.12/site-packages (from notebook) (0.2.4)
Requirement already satisfied: tornado>=6.2.0 in /opt/conda/lib/python3.12/site-
packages (from notebook) (6.4.2)
Requirement already satisfied: anyio>=3.1.0 in /opt/conda/lib/python3.12/site-
packages (from jupyter-server<3,>=2.4.0->notebook) (4.7.0)
Requirement already satisfied: argon2-cffi>=21.1 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (23.1.0)
Requirement already satisfied: jinja2>=3.0.3 in /opt/conda/lib/python3.12/site-
packages (from jupyter-server<3,>=2.4.0->notebook) (3.1.5)
Requirement already satisfied: jupyter-client>=7.4.4 in
/opt/conda/lib/python3.12/site-packages (from jupyter-

server<3,>=2.4.0->notebook) (8.6.3)
Requirement already satisfied: jupyter-core!=5.0.*,>=4.12 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (5.7.2)
Requirement already satisfied: jupyter-events>=0.11.0 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (0.11.0)
Requirement already satisfied: jupyter-server-terminals>=0.4.4 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (0.5.3)
Requirement already satisfied: nbconvert>=6.4.4 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (7.16.4)
Requirement already satisfied: nbformat>=5.3.0 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (5.10.4)
Requirement already satisfied: overrides>=5.0 in /opt/conda/lib/python3.12/site-
packages (from jupyter-server<3,>=2.4.0->notebook) (7.7.0)
Requirement already satisfied: packaging>=22.0 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (24.2)
Requirement already satisfied: prometheus-client>=0.9 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (0.21.1)
Requirement already satisfied: pyzmq>=24 in /opt/conda/lib/python3.12/site-
packages (from jupyter-server<3,>=2.4.0->notebook) (26.2.0)
Requirement already satisfied: send2trash>=1.8.2 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (1.8.3)
Requirement already satisfied: terminado>=0.8.3 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (0.18.1)
Requirement already satisfied: traitlets>=5.6.0 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (5.14.3)
Requirement already satisfied: websocket-client>=1.7 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
server<3,>=2.4.0->notebook) (1.8.0)
Requirement already satisfied: async-lru>=1.0.0 in
/opt/conda/lib/python3.12/site-packages (from jupyterlab<4.4,>=4.3.4->notebook)
(2.0.4)
Requirement already satisfied: httpx>=0.25.0 in /opt/conda/lib/python3.12/site-
packages (from jupyterlab<4.4,>=4.3.4->notebook) (0.28.1)
Requirement already satisfied: ipykernel>=6.5.0 in
/opt/conda/lib/python3.12/site-packages (from jupyterlab<4.4,>=4.3.4->notebook)
(6.29.5)
Requirement already satisfied: jupyter-lsp>=2.0.0 in
/opt/conda/lib/python3.12/site-packages (from jupyterlab<4.4,>=4.3.4->notebook)

(2.2.5)
Requirement already satisfied: setuptools>=40.8.0 in
/opt/conda/lib/python3.12/site-packages (from jupyterlab<4.4,>=4.3.4->notebook)
(75.6.0)
Requirement already satisfied: babel>=2.10 in /opt/conda/lib/python3.12/site-
packages (from jupyterlab-server<3,>=2.27.1->notebook) (2.16.0)
Requirement already satisfied: json5>=0.9.0 in /opt/conda/lib/python3.12/site-
packages (from jupyterlab-server<3,>=2.27.1->notebook) (0.10.0)
Requirement already satisfied: jsonschema>=4.18.0 in
/opt/conda/lib/python3.12/site-packages (from jupyterlab-
server<3,>=2.27.1->notebook) (4.23.0)
Requirement already satisfied: requests>=2.31 in /opt/conda/lib/python3.12/site-
packages (from jupyterlab-server<3,>=2.27.1->notebook) (2.32.3)
Requirement already satisfied: idna>=2.8 in /opt/conda/lib/python3.12/site-
packages (from anyio>=3.1.0->jupyter-server<3,>=2.4.0->notebook) (3.10)
Requirement already satisfied: sniffio>=1.1 in /opt/conda/lib/python3.12/site-
packages (from anyio>=3.1.0->jupyter-server<3,>=2.4.0->notebook) (1.3.1)
Requirement already satisfied: typing_extensions>=4.5 in
/opt/conda/lib/python3.12/site-packages (from anyio>=3.1.0->jupyter-
server<3,>=2.4.0->notebook) (4.12.2)
Requirement already satisfied: argon2-cffi-bindings in
/opt/conda/lib/python3.12/site-packages (from argon2-cffi>=21.1->jupyter-
server<3,>=2.4.0->notebook) (21.2.0)
Requirement already satisfied: certifi in /opt/conda/lib/python3.12/site-
packages (from httpx>=0.25.0->jupyterlab<4.4,>=4.3.4->notebook) (2024.12.14)
Requirement already satisfied: httpcore==1.* in /opt/conda/lib/python3.12/site-
packages (from httpx>=0.25.0->jupyterlab<4.4,>=4.3.4->notebook) (1.0.7)
Requirement already satisfied: h11<0.15,>=0.13 in
/opt/conda/lib/python3.12/site-packages (from
httpcore==1.*->httpx>=0.25.0->jupyterlab<4.4,>=4.3.4->notebook) (0.14.0)
Requirement already satisfied: comm>=0.1.1 in /opt/conda/lib/python3.12/site-
packages (from ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (0.2.2)
Requirement already satisfied: debugpy>=1.6.5 in /opt/conda/lib/python3.12/site-
packages (from ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (1.8.11)
Requirement already satisfied: ipython>=7.23.1 in
/opt/conda/lib/python3.12/site-packages (from
ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (8.31.0)
Requirement already satisfied: matplotlib-inline>=0.1 in
/opt/conda/lib/python3.12/site-packages (from
ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (0.1.7)
Requirement already satisfied: nest-asyncio in /opt/conda/lib/python3.12/site-
packages (from ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (1.6.0)
Requirement already satisfied: psutil in /opt/conda/lib/python3.12/site-packages
(from ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (6.1.1)
Requirement already satisfied: MarkupSafe>=2.0 in
/opt/conda/lib/python3.12/site-packages (from jinja2>=3.0.3->jupyter-
server<3,>=2.4.0->notebook) (3.0.2)
Requirement already satisfied: attrs>=22.2.0 in /opt/conda/lib/python3.12/site-

packages (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.27.1->notebook)
(24.3.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in
/opt/conda/lib/python3.12/site-packages (from jsonschema>=4.18.0->jupyterlab-
server<3,>=2.27.1->notebook) (2024.10.1)
Requirement already satisfied: referencing>=0.28.4 in
/opt/conda/lib/python3.12/site-packages (from jsonschema>=4.18.0->jupyterlab-
server<3,>=2.27.1->notebook) (0.35.1)
Requirement already satisfied: rpds-py>=0.7.1 in /opt/conda/lib/python3.12/site-
packages (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.27.1->notebook)
(0.22.3)
Requirement already satisfied: python-dateutil>=2.8.2 in
/opt/conda/lib/python3.12/site-packages (from jupyter-client>=7.4.4->jupyter-
server<3,>=2.4.0->notebook) (2.9.0.post0)
Requirement already satisfied: platformdirs>=2.5 in
/opt/conda/lib/python3.12/site-packages (from jupyter-
core!=5.0.*,>=4.12->jupyter-server<3,>=2.4.0->notebook) (4.3.6)
Requirement already satisfied: python-json-logger>=2.0.4 in
/opt/conda/lib/python3.12/site-packages (from jupyter-events>=0.11.0->jupyter-
server<3,>=2.4.0->notebook) (2.0.7)
Requirement already satisfied: pyyaml>=5.3 in /opt/conda/lib/python3.12/site-
packages (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
(6.0.2)
Requirement already satisfied: rfc3339-validator in
/opt/conda/lib/python3.12/site-packages (from jupyter-events>=0.11.0->jupyter-
server<3,>=2.4.0->notebook) (0.1.4)
Requirement already satisfied: rfc3986-validator>=0.1.1 in
/opt/conda/lib/python3.12/site-packages (from jupyter-events>=0.11.0->jupyter-
server<3,>=2.4.0->notebook) (0.1.1)
Requirement already satisfied: beautifulsoup4 in /opt/conda/lib/python3.12/site-
packages (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (4.12.3)
Requirement already satisfied: bleach!=5.0.0 in /opt/conda/lib/python3.12/site-
packages (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (6.2.0)
Requirement already satisfied: defusedxml in /opt/conda/lib/python3.12/site-
packages (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (0.7.1)
Requirement already satisfied: jupyterlab-pygments in
/opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-
server<3,>=2.4.0->notebook) (0.3.0)
Requirement already satisfied: mistune<4,>=2.0.3 in
/opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-
server<3,>=2.4.0->notebook) (3.0.2)
Requirement already satisfied: nbclient>=0.5.0 in
/opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-
server<3,>=2.4.0->notebook) (0.10.2)
Requirement already satisfied: pandocfilters>=1.4.1 in
/opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-
server<3,>=2.4.0->notebook) (1.5.0)
Requirement already satisfied: pygments>=2.4.1 in

/opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (2.18.0)
Requirement already satisfied: tinycss2 in /opt/conda/lib/python3.12/site-packages (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (1.4.0)
Requirement already satisfied: fastjsonschema>=2.15 in /opt/conda/lib/python3.12/site-packages (from nbformat>=5.3.0->jupyter-server<3,>=2.4.0->notebook) (2.21.1)
Requirement already satisfied: charset_normalizer<4,>=2 in /opt/conda/lib/python3.12/site-packages (from requests>=2.31->jupyterlab-server<3,>=2.27.1->notebook) (3.4.0)
Requirement already satisfied: urllib3<3,>=1.21.1 in /opt/conda/lib/python3.12/site-packages (from requests>=2.31->jupyterlab-server<3,>=2.27.1->notebook) (2.3.0)
Requirement already satisfied: ptyprocess in /opt/conda/lib/python3.12/site-packages (from terminado>=0.8.3->jupyter-server<3,>=2.4.0->notebook) (0.7.0)
Requirement already satisfied: webencodings in /opt/conda/lib/python3.12/site-packages (from bleach!=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook) (0.5.1)
Requirement already satisfied: decorator in /opt/conda/lib/python3.12/site-packages (from ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (5.1.1)
Requirement already satisfied: jedi>=0.16 in /opt/conda/lib/python3.12/site-packages (from ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (0.19.2)
Requirement already satisfied: pexpect>4.3 in /opt/conda/lib/python3.12/site-packages (from ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (4.9.0)
Requirement already satisfied: prompt_toolkit<3.1.0,>=3.0.41 in /opt/conda/lib/python3.12/site-packages (from ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (3.0.48)
Requirement already satisfied: stack_data in /opt/conda/lib/python3.12/site-packages (from ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook) (0.6.3)
Requirement already satisfied: fqdn in /opt/conda/lib/python3.12/site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (1.5.1)
Requirement already satisfied: isoduration in /opt/conda/lib/python3.12/site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (20.11.0)
Requirement already satisfied: jsonpointer>1.13 in /opt/conda/lib/python3.12/site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (3.0.0)
Requirement already satisfied: uri-template in /opt/conda/lib/python3.12/site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (1.3.0)
Requirement already satisfied: webcolors>=24.6.0 in /opt/conda/lib/python3.12/site-packages (from jsonschema[format-

```
nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
(24.11.1)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.12/site-
packages (from python-dateutil>=2.8.2->jupyter-client>=7.4.4->jupyter-
server<3,>=2.4.0->notebook) (1.17.0)
Requirement already satisfied: cffi>=1.0.1 in /opt/conda/lib/python3.12/site-
packages (from argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-
server<3,>=2.4.0->notebook) (1.17.1)
Requirement already satisfied: soupsieve>1.2 in /opt/conda/lib/python3.12/site-
packages (from beautifulsoup4->nbconvert>=6.4.4->jupyter-
server<3,>=2.4.0->notebook) (2.5)
Requirement already satisfied: pycparser in /opt/conda/lib/python3.12/site-
packages (from cffi>=1.0.1->argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-
server<3,>=2.4.0->notebook) (2.22)
Requirement already satisfied: parso<0.9.0,>=0.8.4 in
/opt/conda/lib/python3.12/site-packages (from
jedi>=0.16->ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook)
(0.8.4)
Requirement already satisfied: wcwidth in /opt/conda/lib/python3.12/site-
packages (from prompt_toolkit<3.1.0,>=3.0.41->ipython>=7.23.1->ipykernel>=6.5.0-
>jupyterlab<4.4,>=4.3.4->notebook) (0.2.13)
Requirement already satisfied: arrow>=0.15.0 in /opt/conda/lib/python3.12/site-
packages (from isoduration->jsonschema[format-nongpl]>=4.18.0->jupyter-
events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (1.3.0)
Requirement already satisfied: executing>=1.2.0 in
/opt/conda/lib/python3.12/site-packages (from
stack_data->ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook)
(2.1.0)
Requirement already satisfied: asttokens>=2.1.0 in
/opt/conda/lib/python3.12/site-packages (from
stack_data->ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook)
(3.0.0)
Requirement already satisfied: pure_eval in /opt/conda/lib/python3.12/site-
packages (from
stack_data->ipython>=7.23.1->ipykernel>=6.5.0->jupyterlab<4.4,>=4.3.4->notebook)
(0.2.3)
Requirement already satisfied: types-python-dateutil>=2.8.10 in
/opt/conda/lib/python3.12/site-packages (from
arrow>=0.15.0->isoduration->jsonschema[format-nongpl]>=4.18.0->jupyter-
events>=0.11.0->jupyter-server<3,>=2.4.0->notebook) (2.9.0.20241206)
Note: you may need to restart the kernel to use updated packages.
```

```python
[4]: import pandas as pd

# Load the CSV file
df = pd.read_csv("sample_emails.csv")
```

```
# Show the first 5 emails
df.head()
```

[4]:
```
   email_id                                         email_text          label
0         1  The client was verified, and all KYC documents…     Compliant
1         2  We pushed this trade without the compliance of…      Escalate
2         3  Can you confirm if this falls under internal a…  Needs Review
3         4  All invoices were cleared through the approved…     Compliant
4         5  This payment exceeded the threshold but was st…      Escalate
```

[6]:
```python
# Simple rule-based classifier (like how GPT prompt logic works)
def classify_email(text):
    text = text.lower()
    if any(keyword in text for keyword in ["skipped", "bypassed", "without␣
 ↪approval", "ignored", "shortcut", "overrode"]):
        return "Escalate"
    elif any(keyword in text for keyword in ["confirm", "check", "unsure",␣
 ↪"verify", "clarify", "guidance"]):
        return "Needs Review"
    else:
        return "Compliant"
```

[7]:
```python
# Apply the function to all rows
df["predicted_label"] = df["email_text"].apply(classify_email)

# Show a few rows with original and predicted labels
df[["email_text", "label", "predicted_label"]].head(10)
```

[7]:
```
                                         email_text          label \
0  The client was verified, and all KYC documents…     Compliant
1  We pushed this trade without the compliance of…      Escalate
2  Can you confirm if this falls under internal a…  Needs Review
3  All invoices were cleared through the approved…     Compliant
4  This payment exceeded the threshold but was st…      Escalate
5  Awaiting confirmation from legal if this contr…  Needs Review
6  Documentation uploaded to the compliance drive…     Compliant
7  Let's process it even though it hasn't been si…      Escalate
8     This needs compliance input before we proceed.  Needs Review
9   Filed all reports before deadline as per policy.     Compliant

  predicted_label
0       Compliant
1       Compliant
2    Needs Review
3       Compliant
4       Compliant
5    Needs Review
```

```
6        Compliant
7        Compliant
8        Compliant
9        Compliant
```

[8]:
```python
# Accuracy of rule-based classifier
accuracy = (df["label"] == df["predicted_label"]).mean()
print(f"Accuracy: {accuracy:.2%}")
```

Accuracy: 52.94%

[9]:
```python
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, classification_report

# Convert text to numeric features
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(df["email_text"])
y = df["label"]

# Split into train and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
 ↪random_state=42)

# Train model
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Predict on test set
y_pred = model.predict(X_test)

# Accuracy
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
```

Accuracy: 0.6363636363636364

```
Classification Report:
                precision    recall  f1-score   support

   Compliant        0.57      1.00      0.73         4
    Escalate        1.00      0.20      0.33         5
Needs Review        0.67      1.00      0.80         2

    accuracy                            0.64        11
   macro avg        0.75      0.73      0.62        11
weighted avg        0.78      0.64      0.56        11
```

```
[10]: !pip install shap
```

Requirement already satisfied: shap in /opt/conda/lib/python3.12/site-packages
(0.48.0)
Requirement already satisfied: numpy in ./.local/lib/python3.12/site-packages
(from shap) (1.26.4)
Requirement already satisfied: scipy in /opt/conda/lib/python3.12/site-packages
(from shap) (1.14.1)
Requirement already satisfied: scikit-learn in /opt/conda/lib/python3.12/site-
packages (from shap) (1.6.0)
Requirement already satisfied: pandas in /opt/conda/lib/python3.12/site-packages
(from shap) (2.2.3)
Requirement already satisfied: tqdm>=4.27.0 in /opt/conda/lib/python3.12/site-
packages (from shap) (4.67.1)
Requirement already satisfied: packaging>20.9 in /opt/conda/lib/python3.12/site-
packages (from shap) (24.2)
Requirement already satisfied: slicer==0.0.8 in /opt/conda/lib/python3.12/site-
packages (from shap) (0.0.8)
Requirement already satisfied: numba>=0.54 in /opt/conda/lib/python3.12/site-
packages (from shap) (0.60.0)
Requirement already satisfied: cloudpickle in /opt/conda/lib/python3.12/site-
packages (from shap) (3.1.0)
Requirement already satisfied: typing-extensions in
/opt/conda/lib/python3.12/site-packages (from shap) (4.12.2)
Requirement already satisfied: llvmlite<0.44,>=0.43.0dev0 in
/opt/conda/lib/python3.12/site-packages (from numba>=0.54->shap) (0.43.0)
Requirement already satisfied: python-dateutil>=2.8.2 in
/opt/conda/lib/python3.12/site-packages (from pandas->shap) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /opt/conda/lib/python3.12/site-
packages (from pandas->shap) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in /opt/conda/lib/python3.12/site-
packages (from pandas->shap) (2024.2)
Requirement already satisfied: joblib>=1.2.0 in /opt/conda/lib/python3.12/site-
packages (from scikit-learn->shap) (1.4.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/opt/conda/lib/python3.12/site-packages (from scikit-learn->shap) (3.5.0)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.12/site-
packages (from python-dateutil>=2.8.2->pandas->shap) (1.17.0)

```python
[16]: import shap

      # Convert to dense matrix for SHAP
      X_test_dense = X_test.toarray()

      # Use the NEW API (recommended for SHAP > 0.40)
      explainer = shap.Explainer(model, X_train.toarray())
```
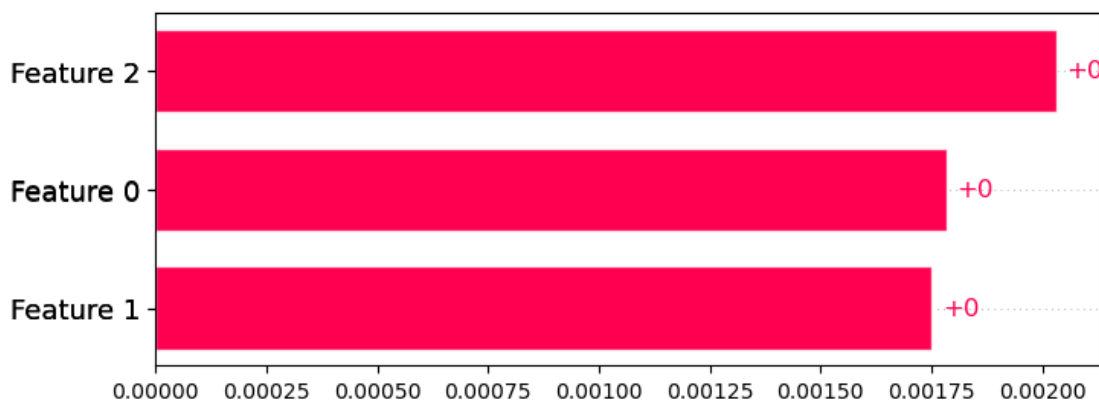
```python
shap_values = explainer(X_test_dense)

# Choose a test sample (e.g., first email)
i = 0   # Change to 1, 2, etc. to test other samples

# Plot the SHAP bar chart for this sample
shap.plots.bar(shap_values[i])
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
Cell In[16], line 14
     11 i = 0  # Change to 1, 2, etc. to test other samples
     13 # Plot the SHAP bar chart for this sample
---> 14 shap.plots.bar(shap_values[i])

File /opt/conda/lib/python3.12/site-packages/shap/plots/_bar.py:315, in
↪bar(shap_values, max_display, order, clustering, clustering_cutoff, show_data
↪ax, show)
    312        ax.axhline(i + 1, color="#888888", lw=0.5, dashes=(1, 5), zorder=-1
    314 if features is not None:
--> 315        features = list(features)
    317        # try and round off any trailing zeros after the decimal point in
↪the feature values
    318        for i in range(len(features)):

TypeError: 'numpy.float64' object is not iterable
```