Problem Statement

Business Context

In today's dynamic business landscape, organizations are increasingly recognizing the pivotal role customer feedback plays in shaping the trajectory of their products and services. The ability to swiftly and effectively respond to customer input not only fosters enhanced customer experiences but also serves as a catalyst for growth, prolonged customer engagement, and the nurturing of lifetime value relationships. As a dedicated Product Manager or Product Analyst, staying attuned to the voice of your customers is not just a best practice; it's a strategic imperative.

While your organization may be inundated with a wealth of customer-generated feedback and support tickets, your role entails much more than just processing these inputs. To make your efforts in managing customer experience and expectations truly impactful, you need a structured approach – a method that allows you to discern the most pressing issues, set priorities, and allocate resources judiciously. One of the most effective strategies at your disposal is to harness the power of Support Ticket Categorization.

Objective

Develop an advanced support ticket categorization system that accurately classifies incoming tickets, assigns relevant tags based on their content, implements mechanisms and generate the first response based on the sentiment for prioritizing tickets for prompt resolution.

Sample of Expected Output

support_tick_id	support_ticket_text	Category	Tags	Priority	ETA	Response
ST2023-007	Urgent help required! My laptop refuses to sta...	Hardware Issues	Laptop, Restart, Hardware Issue	High	24 hours	I apologize for the inconvenience. I'll assist...
ST2023-008	I've accidentally deleted essential work docum...	Data Recovery	data loss, essential documents, data recovery ...	High	Immediate	Please do not access or use your device furthe...
ST2023-009	Despite being in close proximity to my Wi-Fi r...	Technical Issues	[Wi-Fi signal strength, weak Wi-Fi signal]	High	24 hours	I understand that you are experiencing a weak ...
ST2023-010	My smartphone battery is draining rapidly, eve...	Technical issues	battery issue	High	2-3 business days	I apologize for the inconvenience you're exper...

Please read the instructions carefully before starting the project.

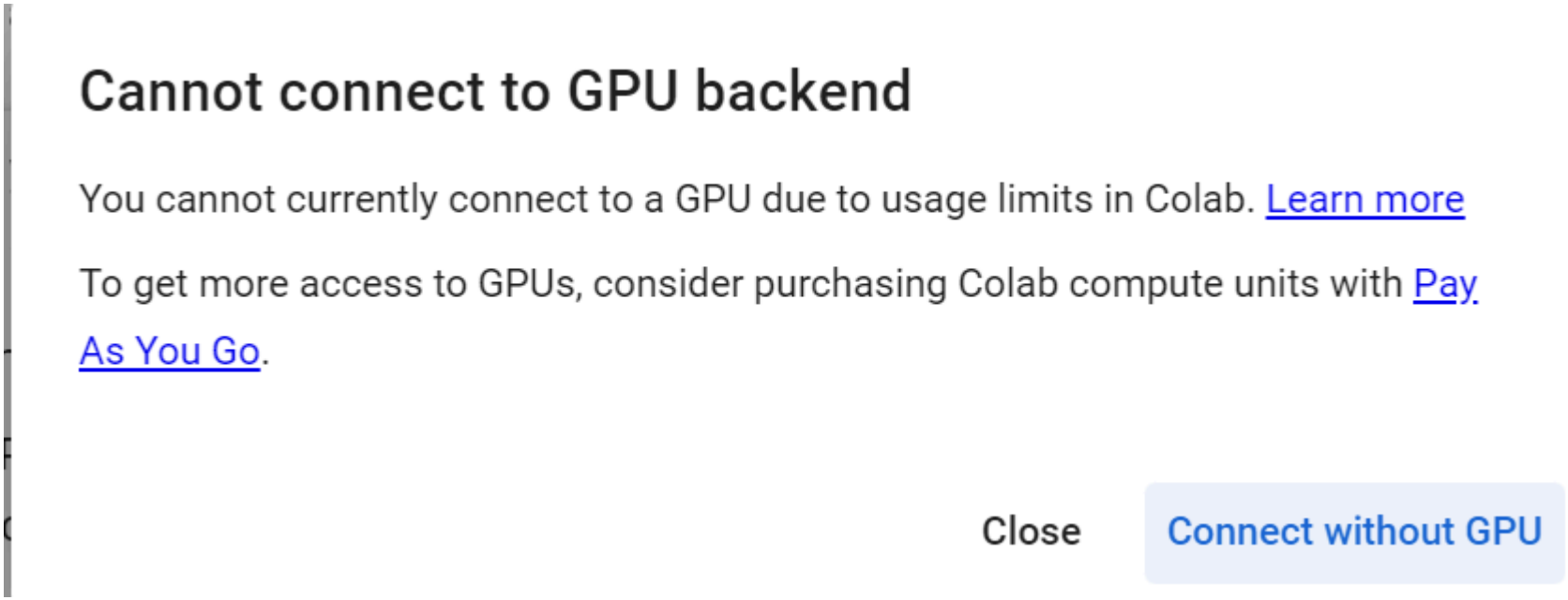
This is a commented Python Notebook file in which all the instructions and tasks to be performed are mentioned.

- Blanks '___' are provided in the notebook that needs to be filled with an appropriate code to get the correct result. With every '___' blank, there is a comment that briefly describes what needs to be filled in the blank space.
- Identify the task to be performed correctly, and only then proceed to write the required code.
- Please run the codes in a sequential manner from the beginning to avoid any unnecessary errors.
- Add the results/observations (wherever mentioned) derived from the analysis in the presentation and submit the same. Any mathematical or computational details which are a graded part of the project can be included in the Appendix section of the presentation.

Note: If the free-tier GPU of Google Colab is not accessible (due to unavailability or exhaustion of daily limit or other reasons), the following steps can be taken:

1. Wait for 12-24 hours until the GPU is accessible again or the daily usage limits are reset.
2. Switch to a different Google account and resume working on the project from there.
3. Try using the CPU runtime:
    - To use the CPU runtime, click on *Runtime => Change runtime type => CPU => Save*
    - One can also click on the *Continue without GPU* option to switch to a CPU runtime (kindly refer to the snapshot below)

- The instructions for running the code on the CPU are provided in the respective sections of the notebook.



Installing and Importing Necessary Libraries and Dependencies

In [ ]:
```
# Installation for GPU llama-cpp-python
# uncomment and run the following code in case GPU is being used

!CMAKE_ARGS="-DLLAMA_CUBLAS=on" FORCE_CMAKE=1 pip install llama-cpp-python==0.1.85 --force-reinstall --no-cache-dir -q
```

```
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.8/1.8 MB 127.3 MB/s eta 0:00:00
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 60.9/60.9 kB 84.3 MB/s eta 0:00:00
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 45.5/45.5 kB 254.3 MB/s eta 0:00:00
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 16.3/16.3 MB 277.2 MB/s eta 0:00:00
  Building wheel for llama-cpp-python (pyproject.toml) ... done
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the
source of the following dependency conflicts.
accelerate 0.32.1 requires numpy<2.0.0,>=1.17, but you have numpy 2.1.0 which is incompatible.
albucore 0.0.13 requires numpy<2,>=1.24.4, but you have numpy 2.1.0 which is incompatible.
arviz 0.18.0 requires numpy<2.0,>=1.23.0, but you have numpy 2.1.0 which is incompatible.
cudf-cu12 24.4.1 requires numpy<2.0a0,>=1.23, but you have numpy 2.1.0 which is incompatible.
cupy-cuda12x 12.2.0 requires numpy<1.27,>=1.20, but you have numpy 2.1.0 which is incompatible.
gensim 4.3.3 requires numpy<2.0,>=1.18.5, but you have numpy 2.1.0 which is incompatible.
ibis-framework 8.0.0 requires numpy<2,>=1, but you have numpy 2.1.0 which is incompatible.
numba 0.60.0 requires numpy<2.1,>=1.22, but you have numpy 2.1.0 which is incompatible.
pandas 2.1.4 requires numpy<2,>=1.22.4; python_version < "3.11", but you have numpy 2.1.0 which is incompatible.
rmm-cu12 24.4.0 requires numpy<2.0a0,>=1.23, but you have numpy 2.1.0 which is incompatible.
scikit-learn 1.3.2 requires numpy<2.0,>=1.17.3, but you have numpy 2.1.0 which is incompatible.
tensorflow 2.17.0 requires numpy<2.0.0,>=1.23.5; python_version <= "3.11", but you have numpy 2.1.0 which is incompatible.
thinc 8.2.5 requires numpy<2.0.0,>=1.19.0; python_version >= "3.9", but you have numpy 2.1.0 which is incompatible.
transformers 4.42.4 requires numpy<2.0,>=1.17, but you have numpy 2.1.0 which is incompatible.
```

In [ ]:

```
# Installation for CPU llama-cpp-python
# uncomment and run the following code in case GPU is not being used

# !CMAKE_ARGS="-DLLAMA_CUBLAS=off" FORCE_CMAKE=1 pip install llama-cpp-python==0.1.85 --force-reinstall --no-cache-dir -q
```

Note : There may be an error related to a dependency issue thrown by the pip package. This can be ignored as it will not impact the execution of the code.

In [ ]:

```
# For downloading the models from HF Hub
!pip install huggingface_hub==0.20.3 pandas==1.5.3 -q
```

```
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 330.1/330.1 kB 19.4 MB/s eta 0:00:00
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.1/12.1 MB 106.8 MB/s eta 0:00:00
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the
source of the following dependency conflicts.
accelerate 0.32.1 requires numpy<2.0.0,>=1.17, but you have numpy 2.1.0 which is incompatible.
arviz 0.18.0 requires numpy<2.0,>=1.23.0, but you have numpy 2.1.0 which is incompatible.
cudf-cu12 24.4.1 requires numpy<2.0a0,>=1.23, but you have numpy 2.1.0 which is incompatible.
cudf-cu12 24.4.1 requires pandas<2.2.2dev0,>=2.0, but you have pandas 1.5.3 which is incompatible.
google-colab 1.0.0 requires pandas==2.1.4, but you have pandas 1.5.3 which is incompatible.
ibis-framework 8.0.0 requires numpy<2,>=1, but you have numpy 2.1.0 which is incompatible.
transformers 4.42.4 requires huggingface-hub<1.0,>=0.23.2, but you have huggingface-hub 0.20.3 which is incompatible.
transformers 4.42.4 requires numpy<2.0,>=1.17, but you have numpy 2.1.0 which is incompatible.
xarray 2024.6.0 requires pandas>=2.0, but you have pandas 1.5.3 which is incompatible.
```

In [ ]:
```python
# Function to download the model from the Hugging Face model hub
from huggingface_hub import hf_hub_download

# Importing the Llama class from the llama_cpp module
from llama_cpp import Llama

# Importing the json module
import json

# for loading and manipulating data
import pandas as pd

# for time computations
import time
```

Loading the Data

In [70]:
```python
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

In [ ]:
```python
# Complete the code to read the CSV file.
data = pd.read_csv(r"/content/Support_ticket_text_data_mid_term.csv")
```

Data Overview

Checking the first 5 rows of the data

In [ ]:
```
# Complete the code to check the first 5 rows of the data
data.head()
```

Out[ ]:

	support_tick_id	support_ticket_text
0	ST2023-006	My internet connection has significantly slowe...
1	ST2023-007	Urgent help required! My laptop refuses to sta...
2	ST2023-008	I've accidentally deleted essential work docum...
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...
4	ST2023-010	My smartphone battery is draining rapidly, eve...

Checking the shape of the data

In [ ]:
```
# Complete the code to check the shape of the data
data.shape
```

Out[ ]:
```
(21, 2)
```

Checking the missing values in the data

In [ ]:
```
# Complete the code to check for missing values in the data
data.isnull().sum()
```

Out[ ]:

	0
support_tick_id	0
support_ticket_text	0

dtype: int64

Model Building

Loading the model

In [ ]:
```
model_name_or_path = "TheBloke/Mistral-7B-Instruct-v0.2-GGUF"
model_basename = "mistral-7b-instruct-v0.2.Q6_K.gguf"
```

In [ ]:
```
model_path = hf_hub_download(
    repo_id=model_name_or_path, # Complete the code to mention the repo_id
    filename=model_basename  # Complete the code to mention the filename
)
```

```
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:88: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secret in your Google Colab and restart your session.
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
mistral-7b-instruct-v0.2.Q6_K.gguf:   0%|          | 0.00/5.94G [00:00<?, ?B/s]
```

In [ ]:
```
# uncomment and run the following code in case GPU is being used

llm = Llama(
    model_path=model_path,
    n_ctx=1024, # Context window
 )
```

```
AVX = 1 | AVX2 = 1 | AVX512 = 0 | AVX512_VBMI = 0 | AVX512_VNNI = 0 | FMA = 1 | NEON = 0 | ARM_FMA = 0 | F16C = 1 | FP16_VA = 0 | WASM_SIMD = 0 | BLAS = 1 | SSE3 = 1 | SSSE3 = 1 | VSX = 0 |
```

In [ ]:
```
# uncomment and run the following code in case GPU is not being used

llm = Llama(
     model_path=model_path,
     n_ctx=1024, # Context window
     #n_cores=2 # Number of CPU cores to use
 )
```

```
AVX = 1 | AVX2 = 1 | AVX512 = 0 | AVX512_VBMI = 0 | AVX512_VNNI = 0 | FMA = 1 | NEON = 0 | ARM_FMA = 0 | F16C = 1 | FP16_VA = 0
| WASM_SIMD = 0 | BLAS = 1 | SSE3 = 1 | SSSE3 = 1 | VSX = 0 |
```

Utility functions

In [ ]:
```python
# defining a function to parse the JSON output from the model
def extract_json_data(json_str):
    try:
        # Find the indices of the opening and closing curly braces
        json_start = json_str.find('{')
        json_end = json_str.rfind('}')

        if json_start != -1 and json_end != -1:
            extracted_category = json_str[json_start:json_end + 1]  # Extract the JSON object
            data_dict = json.loads(extracted_category)
            return data_dict
        else:
            print(f"Warning: JSON object not found in response: {json_str}")
            return {}
    except json.JSONDecodeError as e:
        print(f"Error parsing JSON: {e}")
        return {}
```

Task 1: Ticket Categorization and Returning Structured Output

In [ ]:
```python
# creating a copy of the data
data_1 = data.copy()
```

In [ ]:
```python
#Defining the response funciton for Task 1.
def response_1(prompt,ticket):
    model_output = llm(
      f"""
      Q: {prompt}
      Support ticket: {ticket}
      A:
      """,
      max_tokens=150, #Complete the code to set the maximum number of tokens the model should generate for this task.
      stop=["Response:"],
      temperature=0.5, #Complete the code to set the value for temperature.
      echo=False,
```

```
    )

    temp_output = model_output["choices"][0]["text"]
    final_output = temp_output[temp_output.index('{'):]

    return final_output
```

In [ ]:
```
prompt_1 = """

You are a support assistant designed to handle customer support tickets. Your task is to analyze the provided support ticket and

Here is the support ticket:

Ticket: {ticket}

Please provide your response in the following JSON format:
{
  "Category": "<"type of the issue">",
}

Response:
"""
```

Note: The output of the model should be in a structured format (JSON format).

In [ ]:
```
start = time.time()
data_1['model_response'] = data_1['support_ticket_text'].apply(lambda x: response_1(prompt_1, x))
end = time.time()
```

```
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
```

In [ ]:

```
print("Time taken ",(end-start))
```

```
Time taken  716.3667306900024
```

In [ ]:

```
# Write the code to check the first five rows of the data to confirm whether the new column has been added
data_1.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	model_response
0	ST2023-006	My internet connection has significantly slowe...	{\n "Category": "Internet Connectivity...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "Category": "Hardware Issue"\n }
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "Category": "Data Recovery"\n }
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "Category": "Network Connectivity ...
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "Category": "Hardware Issue",\n ...

In [ ]:

```
i = 2
print(data_1.loc[i, 'support_ticket_text'])
```

```
I've accidentally deleted essential work documents, causing substantial data loss. I understand the need to avoid further action
s on my device. Can you please prioritize the data recovery process and guide me through it?
```

In [ ]:
```
print(data_1.loc[i, 'model_response'])
```

```
{
        "Category": "Data Recovery"
      }
```

In [ ]:
```
# applying the function to the model response
data_1['model_response_parsed'] = data_1["model_response"].apply(extract_json_data)
data_1['model_response_parsed'].head()
```

```
Error parsing JSON: Expecting property name enclosed in double quotes: line 3 column 8 (char 55)
Error parsing JSON: Extra data: line 6 column 1 (char 91)
```

Out[ ]:

	model_response_parsed
0	{'Category': 'Internet Connectivity'}
1	{'Category': 'Hardware Issue'}
2	{'Category': 'Data Recovery'}
3	{'Category': 'Network Connectivity Issues'}
4	{'Category': 'Hardware Issue', 'SolutionSteps'...

dtype: object

In [ ]:
```
data_1['model_response_parsed'].value_counts()
```

Out[ ]:

model_response_parsed	
{'Category': 'Data Recovery'}	4
{'Category': 'Hardware'}	4
{'Category': 'Network Connectivity Issues'}	2
{}	2
{'Category': 'Internet Connectivity'}	1
{'Category': 'Hardware Issue', 'SolutionSteps': ['Check the battery health status in phone settings.', 'Perform a soft reset by removing the SIM card, battery, then press and hold the power button for 30 seconds before reinserting them.', 'Update your phone's software to the latest version.', 'Consider having your device checked by a professional technician if the issue persists.']}	1
{'Category': 'Hardware Issue'}	1
{'Category': 'Performance'}	1
{'Category': 'Display Issues'}	1
{'Category': 'Network Connectivity'}	1
{'Category': 'Hardware - External Drives'}	1
{'Category': 'Internet Connection Issue'}	1
{'Category': 'Software Compatibility/System Issue'}	1

dtype: int64

In [ ]:
```
# Normalizing the model_response_parsed column
model_response_parsed_df_1 = pd.json_normalize(data_1['model_response_parsed'])
model_response_parsed_df_1.head()
```

Out[ ]:

	Category	SolutionSteps
0	Internet Connectivity	NaN
1	Hardware Issue	NaN
2	Data Recovery	NaN
3	Network Connectivity Issues	NaN
4	Hardware Issue	[Check the battery health status in phone sett...

In [ ]:
```
# Concatinating two dataframes
data_with_parsed_model_output_1 = pd.concat([data_1, model_response_parsed_df_1], axis=1)
data_with_parsed_model_output_1.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	model_response	model_response_parsed	Category	SolutionSteps
0	ST2023-006	My internet connection has significantly slowe...	{\n "Category": "Internet Connectivity...	{'Category': 'Internet Connectivity'}	Internet Connectivity	NaN
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "Category": "Hardware Issue"\n }	{'Category': 'Hardware Issue'}	Hardware Issue	NaN
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "Category": "Data Recovery"\n }	{'Category': 'Data Recovery'}	Data Recovery	NaN
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "Category": "Network Connectivity ...	{'Category': 'Network Connectivity Issues'}	Network Connectivity Issues	NaN
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "Category": "Hardware Issue",\n ...	{'Category': 'Hardware Issue', 'SolutionSteps'...	Hardware Issue	[Check the battery health status in phone sett...

In [ ]:
```
# Dropping model_response and model_response_parsed columns
final_data_1 = data_with_parsed_model_output_1.drop(['model_response','model_response_parsed','SolutionSteps'], axis=1)
final_data_1.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	Category
0	ST2023-006	My internet connection has significantly slowe...	Internet Connectivity
1	ST2023-007	Urgent help required! My laptop refuses to sta...	Hardware Issue
2	ST2023-008	I've accidentally deleted essential work docum...	Data Recovery
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	Network Connectivity Issues
4	ST2023-010	My smartphone battery is draining rapidly, eve...	Hardware Issue

Task 2: Creating Tags

In [ ]:
```
# creating a copy of the data
data_2 = data.copy()
```

In [ ]:
```
def response_2(prompt,ticket,category):
    model_output = llm(
      f"""
      Q: {prompt}
      Support ticket: {ticket}
      Category: {category}
      A:
      """,
      max_tokens=200, #Complete the code to set the maximum number of tokens the model should generate for this task.
      stop=["Response:"],
      temperature=0.7, #Complete the code to set the value for temperature.
      echo=False,
    )

    temp_output = model_output["choices"][0]["text"]
    final_output = temp_output[temp_output.index('{'):]

    return final_output
```

In [ ]:
```
prompt_2 = """
You are a support assistant designed to handle customer support tickets. Your task is to analyze the provided support ticket and

Here is the support ticket:
```

```
Ticket: {ticket}

Please provide your response in the following JSON format:
{
  "Tags": "<specific problem about the issue (e.g., laptop, restart, hardware issue, data loss, etc.)>",
}

Response:
"""
```

Note: The output of the model should be in a structured format (JSON format).

In [ ]:

```python
start = time.time()
data_2["model_response"]=final_data_1[['support_ticket_text','Category']].apply(lambda x: response_2(prompt_2, x[0],x[1]),axis =1
end = time.time()
```

```
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
```

In [ ]:

```python
print("Time taken ",end-start)
```

```
Time taken  446.70553255081177
```

In [ ]:
```
# Write the code to check the first five rows of the data to confirm whether the new column has been added
data_2.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	model_response
0	ST2023-006	My internet connection has significantly slowe...	{\n "Tags": "internet connectivity, sl...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "Tags": ["laptop", "hardware issue...
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "Tags": ["data loss"]\n }
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "Tags": ["Wi-Fi", "Signal Strengt...
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "Tags": ["smartphone", "battery", ...

In [ ]:
```
i = 2
print(data_2.loc[i, 'support_ticket_text'])
```

```
I've accidentally deleted essential work documents, causing substantial data loss. I understand the need to avoid further action
s on my device. Can you please prioritize the data recovery process and guide me through it?
```

In [ ]:
```
print(data_2.loc[i, 'model_response'])
```

```
{
        "Tags": ["data loss"]
      }
```

In [ ]:
```
# Applying the function to the model response
data_2['model_response_parsed'] = data_2['model_response'].apply(extract_json_data)
```

In [ ]:
```
data_2["model_response_parsed"]
```

Out[ ]:

	model_response_parsed
0	{'Tags': 'internet connectivity, slow connecti...
1	{'Tags': ['laptop', 'hardware issue']}
2	{'Tags': ['data loss']}
3	{'Tags': ['Wi-Fi', 'Signal Strength', 'Network...
4	{'Tags': ['smartphone', 'battery', 'hardware i...
5	{'Tags': ['online banking', 'password reset']}
6	{'Tags': 'performance optimization'}
7	{'Tags': ['hardware issue', 'blue screen error']}
8	{'Tags': ['hardware issue', 'external hard dri...
9	{'Tags': 'laptop, graphics card'}
10	{'Tags': ['Data Loss']}
11	{'Tags': ['laptop', 'display issues']}
12	{'Tags': ['laptop', 'hardware issue']}
13	{'Tags': ['data loss', 'hardware issue']}
14	{'Tags': ['laptop', 'hardware issue']}
15	{'Tags': ['internet', 'connection', 'dropping']}
16	{'Tags': ['Wi-Fi', 'network connectivity']}
17	{'Tags': ['data loss', 'usb drive']}
18	{'Tags': ['Hardware issue', 'External drives',...
19	{'Tags': 'internet connection, slow speed, fre...
20	{'Tags': ['software compatibility', 'system is...

dtype: object

In [ ]:
```
# Normalizing the model_response_parsed column
model_response_parsed_df_2 = pd.json_normalize(data_2['model_response_parsed'])
model_response_parsed_df_2.head()
```

Out[ ]:

	Tags
0	internet connectivity, slow connection, discon...
1	[laptop, hardware issue]
2	[data loss]
3	[Wi-Fi, Signal Strength, Network Connectivity]
4	[smartphone, battery, hardware issue]

In [ ]:
```
# Concatinating two dataframes
data_with_parsed_model_output_2 = pd.concat([data_2, model_response_parsed_df_2], axis=1)
data_with_parsed_model_output_2.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	model_response	model_response_parsed	Tags
0	ST2023-006	My internet connection has significantly slowe...	{\n "Tags": "internet connectivity, sl...	{'Tags': 'internet connectivity, slow connecti...	internet connectivity, slow connection, discon...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "Tags": ["laptop", "hardware issue...	{'Tags': ['laptop', 'hardware issue']}	[laptop, hardware issue]
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "Tags": ["data loss"]\n }	{'Tags': ['data loss']}	[data loss]
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "Tags": ["Wi-Fi", "Signal Strengt...	{'Tags': ['Wi-Fi', 'Signal Strength', 'Network...	[Wi-Fi, Signal Strength, Network Connectivity]
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "Tags": ["smartphone", "battery", ...	{'Tags': ['smartphone', 'battery', 'hardware i...	[smartphone, battery, hardware issue]

In [ ]:
```
# Dropping model_response and model_response_parsed columns
final_data_2 = data_with_parsed_model_output_2.drop(['model_response','model_response_parsed'], axis=1)
final_data_2.head()
```

Out[ ]:

	support_tick_id	support_ticket_text	Tags
0	ST2023-006	My internet connection has significantly slowe...	internet connectivity, slow connection, discon...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	[laptop, hardware issue]
2	ST2023-008	I've accidentally deleted essential work docum...	[data loss]
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	[Wi-Fi, Signal Strength, Network Connectivity]
4	ST2023-010	My smartphone battery is draining rapidly, eve...	[smartphone, battery, hardware issue]

In [ ]:

```
# Checking the value counts of Category column
final_data_2['Tags'].value_counts()
```

Out[ ]:

Tags	
[laptop, hardware issue]	3
internet connectivity, slow connection, disconnection	1
[data loss]	1
[Wi-Fi, Signal Strength, Network Connectivity]	1
[smartphone, battery, hardware issue]	1
[online banking, password reset]	1
performance optimization	1
[hardware issue, blue screen error]	1
[hardware issue, external hard drive]	1
laptop, graphics card	1
[Data Loss]	1
[laptop, display issues]	1
[data loss, hardware issue]	1
[internet, connection, dropping]	1
[Wi-Fi, network connectivity]	1
[data loss, usb drive]	1
[Hardware issue, External drives, Data recovery]	1
internet connection, slow speed, frequent disconnections	1
[software compatibility, system issue]	1

dtype: int64

In [ ]:
```
final_data_2 = pd.concat([final_data_2,final_data_1["Category"]],axis=1)
```

In [ ]:
```
final_data_2 = final_data_2[["support_tick_id","support_ticket_text","Category","Tags"]]
final_data_2
```

Out[ ]:

	support_tick_id	support_ticket_text	Category	Tags
0	ST2023-006	My internet connection has significantly slowe...	Internet Connectivity	internet connectivity, slow connection, discon...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	Hardware Issue	[laptop, hardware issue]
2	ST2023-008	I've accidentally deleted essential work docum...	Data Recovery	[data loss]
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	Network Connectivity Issues	[Wi-Fi, Signal Strength, Network Connectivity]
4	ST2023-010	My smartphone battery is draining rapidly, eve...	Hardware Issue	[smartphone, battery, hardware issue]
5	ST2023-011	I'm locked out of my online banking account an...	NaN	[online banking, password reset]
6	ST2023-012	My computer's performance is sluggish, severel...	Performance	performance optimization
7	ST2023-013	I'm experiencing a recurring blue screen error...	Hardware	[hardware issue, blue screen error]
8	ST2023-014	My external hard drive isn't being recognized ...	Hardware	[hardware issue, external hard drive]
9	ST2023-015	The graphics card in my gaming laptop seems to...	Hardware	laptop, graphics card
10	ST2023-016	I accidentally formatted my USB drive with cri...	Data Recovery	[Data Loss]
11	ST2023-017	My computer's screen has gone black, and I can...	Display Issues	[laptop, display issues]
12	ST2023-018	I accidentally spilled water on my laptop, and...	NaN	[laptop, hardware issue]
13	ST2023-019	My USB flash drive is physically damaged, and ...	Data Recovery	[data loss, hardware issue]
14	ST2023-020	The touchpad on my laptop has stopped working,...	Hardware	[laptop, hardware issue]
15	ST2023-021	My internet connection is frequently dropping,...	Network Connectivity Issues	[internet, connection, dropping]
16	ST2023-022	Wi-Fi is inconsistent despite proximity to the...	Network Connectivity	[Wi-Fi, network connectivity]
17	ST2023-023	I accidentally formatted my USB drive with cru...	Data Recovery	[data loss, usb drive]
18	ST2023-024	My external hard drive isn't being recognized,...	Hardware - External Drives	[Hardware issue, External drives, Data recovery]
19	ST2023-025	I am experiencing a critical problem with my i...	Internet Connection Issue	internet connection, slow speed, frequent disc...
20	ST2023-026	I hope this message finds you well. I am writi...	Software Compatibility/System Issue	[software compatibility, system issue]

Task 3: Assigning Priority and ETA

In [ ]:
```
# creating a copy of the data
data_3 = data.copy()
```

In [93]:
```
def response_3(prompt,ticket,category,tags):
    model_output = llm(
      f"""
      Q: {prompt}
      Support ticket: {ticket}
      Category: {category}
      Tags: {tags}
      A:
      """,
      max_tokens=200,  #Complete the code to set the maximum number of tokens the model should generate for this task.
      stop=["Response:"],
      temperature=0.7, #Complete the code to set the value for temperature.
      echo=False,
    )

    temp_output = model_output["choices"][0]["text"]
    final_output = temp_output[temp_output.index('{'):]

    return final_output
```

In [94]:
```
prompt_3 = """
 You are a support assistant designed to handle customer support tickets. Your task is to analyze the provided support ticket and
"ETA": The estimated time to resolve the issue (e.g., 2 hours, 1 day, etc.)
"Priority": should be The priority level of the issue
Here is the support ticket:

Ticket: {ticket}

Please provide your response in the following JSON format:
{
  "ETA": "<The estimated time to resolve the issue>",
  "Priority": "<The priority level of the issues).>"
}

Response:
"""
```

Note: The output of the model should be in a structured format (JSON format).

In [95]:
```
# Applying generate_llama_response function on support_ticket_text column
start = time.time()
data_3['model_response'] = final_data_2[['support_ticket_text','Category','Tags']].apply(lambda x: response_3(prompt_3, x[0],x[1]
end = time.time()
```

```
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
```

In [96]:
```
print("Time taken ",(end-start))
```

```
Time taken  1561.8319518566132
```

In [97]:
```
# Write the code to check the first five rows of the data to confirm whether the new column has been added
data_3.head()
```

Out[97]:

	support_tick_id	support_ticket_text	model_response	model_response_parsed
0	ST2023-006	My internet connection has significantly slowe...	{\n "ETA": "Within 4 hours",\n "...	{'ETA': 'Within 4 hours', 'Priority': 'High'}
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "ETA": "Within 4 hours",\n ...	{'ETA': 'Within 4 hours', 'Priority': 'High'}
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "ETA": "1-2 business days, dependi...	{'ETA': '1-2 business days', 'Priority': 'high'}
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "ETA": "1 business day",\n ...	{'ETA': '1 day', 'Priority': 'high'}
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "ETA": "2 hours",\n "Priorit...	{'ETA': '2 hours', 'Priority': 'high'}

In [98]:
```
i = 2
print(data_3.loc[i, 'support_ticket_text'])
```

```
I've accidentally deleted essential work documents, causing substantial data loss. I understand the need to avoid further action
s on my device. Can you please prioritize the data recovery process and guide me through it?
```

In [99]:
```
print(data_3.loc[i, 'model_response'])
```

```
{
        "ETA": "1-2 business days, depending on the extent of data loss and the availability of backup resources",
        "Priority": "High"
      }
```

In [100…
```
# Applying the function to the model response
data_3['model_response_parsed'] = data_3['model_response'].apply(extract_json_data)
data_3['model_response_parsed'].head()
```

Out[100]:

	model_response_parsed
0	{'ETA': 'Within 4 hours', 'Priority': 'High'}
1	{'ETA': 'Within 4 hours', 'Priority': 'High'}
2	{'ETA': '1-2 business days, depending on the e...
3	{'ETA': '1 business day', 'Priority': 'Medium'}
4	{'ETA': '2 hours', 'Priority': 'Medium'}

dtype: object

In [101…

```
# Normalizing the model_response_parsed column
model_response_parsed_df_3 = pd.json_normalize(data_3['model_response_parsed'])
model_response_parsed_df_3.head(21)
```

Out[101]:

	ETA	Priority
0	Within 4 hours	High
1	Within 4 hours	High
2	1-2 business days, depending on the extent of ...	High
3	1 business day	Medium
4	2 hours	Medium
5	Within 1 hour	High
6	4 hours	High
7	Within 2 working days	High
8	1-2 business days	High
9	2-3 hours	High
10	1-2 days	High
11	1 hour	High
12	1-2 business days for assessment, longer if pa...	High
13	2-3 business days	High
14	1-2 business days (due to the need for replace...	High
15	Within 4 hours	High
16	Within 4 hours	High
17	1-3 days depending on the extent of the data loss	High
18	2-3 business days for a thorough diagnostic an...	High
19	1 day	High
20	1 day	High

In [102…
```
# Concatinating two dataframes
data_with_parsed_model_output_3 = pd.concat([data_3, model_response_parsed_df_3], axis=1)
data_with_parsed_model_output_3.head()
```

Out[102]:

	support_tick_id	support_ticket_text	model_response	model_response_parsed	ETA	Priority
0	ST2023-006	My internet connection has significantly slowe...	{\n "ETA": "Within 4 hours",\n "...	{'ETA': 'Within 4 hours', 'Priority': 'High'}	Within 4 hours	High
1	ST2023-007	Urgent help required! My laptop refuses to sta...	{\n "ETA": "Within 4 hours",\n ...	{'ETA': 'Within 4 hours', 'Priority': 'High'}	Within 4 hours	High
2	ST2023-008	I've accidentally deleted essential work docum...	{\n "ETA": "1-2 business days, dependi...	{'ETA': '1-2 business days, depending on the e...	1-2 business days, depending on the extent of ...	High
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	{\n "ETA": "1 business day",\n ...	{'ETA': '1 business day', 'Priority': 'Medium'}	1 business day	Medium
4	ST2023-010	My smartphone battery is draining rapidly, eve...	{\n "ETA": "2 hours",\n "Priorit...	{'ETA': '2 hours', 'Priority': 'Medium'}	2 hours	Medium

In [103…
```
# Dropping model_response and model_response_parsed columns
final_data_3 = data_with_parsed_model_output_3.drop(['model_response','model_response_parsed'], axis=1)
final_data_3.head()
```

Out[103]:

	support_tick_id	support_ticket_text	ETA	Priority
0	ST2023-006	My internet connection has significantly slowe...	Within 4 hours	High
1	ST2023-007	Urgent help required! My laptop refuses to sta...	Within 4 hours	High
2	ST2023-008	I've accidentally deleted essential work docum...	1-2 business days, depending on the extent of ...	High
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	1 business day	Medium
4	ST2023-010	My smartphone battery is draining rapidly, eve...	2 hours	Medium

In [104…
```
final_data_3 = pd.concat([final_data_3,final_data_2[["Category","Tags"]]],axis=1)
```

In [105…
```
final_data_3 = final_data_3[["support_tick_id","support_ticket_text","Category","Priority","Tags","ETA"]]
```

In [106…

```
final_data_3
```

Out[106]:

	support_tick_id	support_ticket_text	Category	Priority	Tags	ETA
0	ST2023-006	My internet connection has significantly slowe...	Internet Connectivity	High	internet connectivity, slow connection, discon...	Within 4 hours
1	ST2023-007	Urgent help required! My laptop refuses to sta...	Hardware Issue	High	[laptop, hardware issue]	Within 4 hours
2	ST2023-008	I've accidentally deleted essential work docum...	Data Recovery	High	[data loss]	1-2 business days, depending on the extent of ...
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	Network Connectivity Issues	Medium	[Wi-Fi, Signal Strength, Network Connectivity]	1 business day
4	ST2023-010	My smartphone battery is draining rapidly, eve...	Hardware Issue	Medium	[smartphone, battery, hardware issue]	2 hours
5	ST2023-011	I'm locked out of my online banking account an...	NaN	High	[online banking, password reset]	Within 1 hour
6	ST2023-012	My computer's performance is sluggish, severel...	Performance	High	performance optimization	4 hours
7	ST2023-013	I'm experiencing a recurring blue screen error...	Hardware	High	[hardware issue, blue screen error]	Within 2 working days
8	ST2023-014	My external hard drive isn't being recognized ...	Hardware	High	[hardware issue, external hard drive]	1-2 business days
9	ST2023-015	The graphics card in my gaming laptop seems to...	Hardware	High	laptop, graphics card	2-3 hours
10	ST2023-016	I accidentally formatted my USB drive with cri...	Data Recovery	High	[Data Loss]	1-2 days
11	ST2023-017	My computer's screen has gone black, and I can...	Display Issues	High	[laptop, display issues]	1 hour
12	ST2023-018	I accidentally spilled water on my laptop, and...	NaN	High	[laptop, hardware issue]	1-2 business days for assessment, longer if pa...
13	ST2023-019	My USB flash drive is physically damaged, and ...	Data Recovery	High	[data loss, hardware issue]	2-3 business days
14	ST2023-020	The touchpad on my laptop has stopped working,...	Hardware	High	[laptop, hardware issue]	1-2 business days (due to the need for replace...

	support_tick_id	support_ticket_text	Category	Priority	Tags	ETA
15	ST2023-021	My internet connection is frequently dropping,...	Network Connectivity Issues	High	[internet, connection, dropping]	Within 4 hours
16	ST2023-022	Wi-Fi is inconsistent despite proximity to the...	Network Connectivity	High	[Wi-Fi, network connectivity]	Within 4 hours
17	ST2023-023	I accidentally formatted my USB drive with cru...	Data Recovery	High	[data loss, usb drive]	1-3 days depending on the extent of the data loss
18	ST2023-024	My external hard drive isn't being recognized,...	Hardware - External Drives	High	[Hardware issue, External drives, Data recovery]	2-3 business days for a thorough diagnostic an...
19	ST2023-025	I am experiencing a critical problem with my i...	Internet Connection Issue	High	internet connection, slow speed, frequent disc...	1 day
20	ST2023-026	I hope this message finds you well. I am writi...	Software Compatibility/System Issue	High	[software compatibility, system issue]	1 day

Task 4 - Creating a Draft Response

In [ ]:
```
# creating a copy of the data
data_4 = data.copy()
```

In [ ]:
```
def response_4(prompt,ticket,category,tags,priority,eta):
    model_output = llm(
      f"""
      Q: {prompt}
      Support ticket: {ticket}
      Category : {category}
      Tags : {tags}
      Priority: {priority}
      ETA: {eta}
      A:
      """,
      max_tokens=200, #Complete the code to set the maximum number of tokens the model should generate for this task.
      stop=["Response:"],
      temperature=0.5, #Complete the code to set the value for temperature.
      echo=False,
    )
```

```
    temp_output = model_output["choices"][0]["text"]


    return temp_output
```

In [ ]:
```
prompt_4 = """
 You are a support assistant designed to handle customer support tickets. Your task is to analyze the provided support ticket and

Here is the support ticket:

Ticket: {ticket}

Please provide your response in  necessary precaution or suggestions in string format

Response:
"""
```

Note : For this task, we will not be using the `extract_json_data` function. Hence, the output from the model should be a plain string and not a JSON object.

In [69]:
```
#Applying generate_llama_response function on support_ticket_text column
start = time.time()
data_4['model_response'] = final_data_3[['support_ticket_text','Category','Tags','Priority','ETA']].apply(lambda x: response_4(pr
end = time.time()
```

```
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
Llama.generate: prefix-match hit
```

In [71]:

```
print("Time taken",(end-start))
```

```
Time taken 3890.2632644176483
```

In [72]:

```
# Write the code to check the first five rows of the data to confirm whether the new column has been added
data_4.head()
```

Out[72]:

	support_tick_id	support_ticket_text	model_response
0	ST2023-006	My internet connection has significantly slowe...	1. Check your internet connection by running a...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	\n We are sorry to hear that you're exper...
2	ST2023-008	I've accidentally deleted essential work docum...	\n Dear Valued Customer,\n \n W...
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	\n Dear Valued Customer,\n \n T...
4	ST2023-010	My smartphone battery is draining rapidly, eve...	\n We're sorry to hear that your smartpho...

In [73]:

```
i = 2
print(data_4.loc[i, 'support_ticket_text'])
```

I've accidentally deleted essential work documents, causing substantial data loss. I understand the need to avoid further actions on my device. Can you please prioritize the data recovery process and guide me through it?

In [74]:
```
print(data_4.loc[i, 'model_response'])
```

Dear Valued Customer,

We are truly sorry for the inconvenience you've experienced due to accidental data loss. Our team of experts will prioritize your case and work diligently to recover your essential documents as soon as possible.

In the meantime, we strongly advise against using your device for any critical tasks or activities that could potentially overwrite the deleted files. This includes saving new files, installing software, or running system updates. The more data that is written to your device, the lower the chances of successful recovery become.

If you have a backup solution in place such as an external hard drive or cloud storage, please ensure it is connected to your device and let us know so we can guide you through the process of restoring from the backup.

In case you don't have a backup, our data recovery team will use advanced tools and techniques to attempt the recovery. However, there

In [110…
```
final_data_4 = pd.concat([final_data_3,data_4["model_response"]],axis=1)
```

In [111…
```
final_data_4.rename(columns={"model_response":"Response"},inplace=True)
```

In [112…
```
final_data_4
```

Out[112]:

	support_tick_id	support_ticket_text	Category	Priority	Tags	ETA	Response
0	ST2023-006	My internet connection has significantly slowe...	Internet Connectivity	High	internet connectivity, slow connection, discon...	Within 4 hours	1. Check your internet connection by running a...
1	ST2023-007	Urgent help required! My laptop refuses to sta...	Hardware Issue	High	[laptop, hardware issue]	Within 4 hours	\n We are sorry to hear that you're exper...
2	ST2023-008	I've accidentally deleted essential work docum...	Data Recovery	High	[data loss]	1-2 business days, depending on the extent of ...	\n Dear Valued Customer,\n \n W...
3	ST2023-009	Despite being in close proximity to my Wi-Fi r...	Network Connectivity Issues	Medium	[Wi-Fi, Signal Strength, Network Connectivity]	1 business day	\n Dear Valued Customer,\n \n T...
4	ST2023-010	My smartphone battery is draining rapidly, eve...	Hardware Issue	Medium	[smartphone, battery, hardware issue]	2 hours	\n We're sorry to hear that your smartpho...
5	ST2023-011	I'm locked out of my online banking account an...	NaN	High	[online banking, password reset]	Within 1 hour	\n Dear Valued Customer,\n \n W...
6	ST2023-012	My computer's performance is sluggish, severel...	Performance	High	performance optimization	4 hours	1. Check for and remove unnecessary startup pr...
7	ST2023-013	I'm experiencing a recurring blue screen error...	Hardware	High	[hardware issue, blue screen error]	Within 2 working days	1. Ensure that all drivers are up to date, inc...
8	ST2023-014	My external hard drive isn't being recognized ...	Hardware	High	[hardware issue, external hard drive]	1-2 business days	\n Dear Valued Customer,\n \n W...
9	ST2023-015	The graphics card in my gaming laptop seems to...	Hardware	High	laptop, graphics card	2-3 hours	\n Dear Valued Customer,\n\n We apo...
10	ST2023-016	I accidentally formatted my USB drive with cri...	Data Recovery	High	[Data Loss]	1-2 days	1. Do not write any new data on the formatted ...
11	ST2023-017	My computer's screen has gone black, and I can...	Display Issues	High	[laptop, display issues]	1 hour	1. Check if the laptop is plugged in and the p...
12	ST2023-018	I accidentally spilled water on my laptop, and...	NaN	High	[laptop, hardware issue]	1-2 business days for assessment, longer if pa...	1. We're sorry to hear that your laptop has en...

	support_tick_id	support_ticket_text	Category	Priority	Tags	ETA	Response
13	ST2023-019	My USB flash drive is physically damaged, and ...	Data Recovery	High	[data loss, hardware issue]	2-3 business days	\n We're sorry to hear that your USB flas...
14	ST2023-020	The touchpad on my laptop has stopped working,...	Hardware	High	[laptop, hardware issue]	1-2 business days (due to the need for replace...	\n We apologize for any inconvenience cau...
15	ST2023-021	My internet connection is frequently dropping,...	Network Connectivity Issues	High	[internet, connection, dropping]	Within 4 hours	\n We apologize for the inconvenience you...
16	ST2023-022	Wi-Fi is inconsistent despite proximity to the...	Network Connectivity	High	[Wi-Fi, network connectivity]	Within 4 hours	1. Restart your Wi-Fi router and modem by unpl...
17	ST2023-023	I accidentally formatted my USB drive with cru...	Data Recovery	High	[data loss, usb drive]	1-3 days depending on the extent of the data loss	\n We're sorry to hear that you have acci...
18	ST2023-024	My external hard drive isn't being recognized,...	Hardware - External Drives	High	[Hardware issue, External drives, Data recovery]	2-3 business days for a thorough diagnostic an...	1. Ensure the external hard drive is properly ...
19	ST2023-025	I am experiencing a critical problem with my i...	Internet Connection Issue	High	internet connection, slow speed, frequent disc...	1 day	Dear Valued Customer,\n We apologize fo...
20	ST2023-026	I hope this message finds you well. I am writi...	Software Compatibility/System Issue	High	[software compatibility, system issue]	1 day	Dear Valued Customer,\n Thank you for r...

Model Output Analysis

In [113...
```
# Creating a copy of the dataframe of task-4
final_data = final_data_4.copy()
```

In [114...
```
final_data['Category'].value_counts()    # complete the code with the column name for the column containing ticket categories
```

Out[114]:

	Category
Data Recovery	4
Hardware	4
Hardware Issue	2
Network Connectivity Issues	2
Internet Connectivity	1
Performance	1
Display Issues	1
Network Connectivity	1
Hardware - External Drives	1
Internet Connection Issue	1
Software Compatibility/System Issue	1

dtype: int64

In [115…

```
final_data["Priority"].value_counts() # complete the code with the column name for the column containing the priorities of the ti
```

Out[115]:

	Priority
High	19
Medium	2

dtype: int64

In [116…

```
final_data["ETA"].value_counts()# complete the code with the column name for the column containing ticket resolution ETA.
```

Out[116]:

ETA	
Within 4 hours	4
1 day	2
1-2 business days, depending on the extent of data loss and the availability of backup resources	1
2 hours	1
1 business day	1
4 hours	1
Within 2 working days	1
1-2 business days	1
Within 1 hour	1
2-3 hours	1
1-2 days	1
1-2 business days for assessment, longer if parts need to be ordered	1
1 hour	1
2-3 business days	1
1-2 business days (due to the need for replacement parts)	1
1-3 days depending on the extent of the data loss	1
2-3 business days for a thorough diagnostic and potential data recovery process.	1

dtype: int64

Let's dive in a bit deeper here.

In [117…

```
final_data.groupby(['Category', 'ETA']).support_tick_id.count() # complete the code to group by based on the categories and ETA.
```

Out[117]:

Category	ETA	support_tick_id
Data Recovery	1-2 business days, depending on the extent of data loss and the availability of backup resources	1
	1-2 days	1
	1-3 days depending on the extent of the data loss	1
	2-3 business days	1
Display Issues	1 hour	1
Hardware	1-2 business days	1
	1-2 business days (due to the need for replacement parts)	1
	2-3 hours	1
	Within 2 working days	1
Hardware - External Drives	2-3 business days for a thorough diagnostic and potential data recovery process.	1
Hardware Issue	2 hours	1
	Within 4 hours	1
Internet Connection Issue	1 day	1
Internet Connectivity	Within 4 hours	1
Network Connectivity	Within 4 hours	1
Network Connectivity Issues	1 business day	1
	Within 4 hours	1
Performance	4 hours	1
Software Compatibility/System Issue	1 day	1

dtype: int64

Actionable Insights and Recommendations

Actionable Insights

1. Leverage Natural Language Processing (NLP)

Sentiment Analysis: Use NLP techniques to analyze the sentiment of incoming tickets. This will help in determining the urgency and priority based on customer emotions and language. Text Classification: Implement machine learning models to classify tickets into predefined categories based on their content. This can include issues like "Technical Support," "Billing Inquiry," or "Feature Request."

1. Implement Automated Tagging

Tagging Mechanisms: Develop an automated tagging system that applies relevant tags to tickets based on their content. Tags could include keywords or phrases that indicate the nature of the issue. Dynamic Tagging: Use models that can learn from new data to adapt and apply relevant tags as the nature of issues evolves over time.

1. Create a Prioritization Framework

Priority Levels: Define clear criteria for prioritizing tickets. For example, tickets from high-value customers or those expressing severe frustration might be classified as high priority. Automated Prioritization: Integrate the sentiment analysis and categorization models to automatically assign priority levels to tickets. Ensure the system can handle various priority levels such as "Critical," "High," "Medium," and "Low."

1. Develop a Response Generation System

Template Responses: Create a library of response templates for different types of issues and sentiments. This can include predefined responses for common problems and customizable templates for more complex scenarios. Personalization: Ensure the response system can incorporate customer-specific details to personalize the communication, improving customer satisfaction.

1. Integrate with Existing Systems

CRM Integration: Ensure that the categorization and prioritization system integrates seamlessly with your Customer Relationship Management (CRM) system for unified data management. Support Ticket Platforms: Integrate with existing support ticket platforms (e.g., Zendesk, Freshdesk) to automate the ticket classification and response process.

1. Monitor and Improve

Feedback Loop: Establish a feedback loop to continually assess the effectiveness of the categorization and prioritization system. Use this feedback to make iterative improvements. Performance Metrics: Track key performance metrics such as response time, resolution time, and customer satisfaction scores to evaluate the system's impact and efficiency.

Recommendations

1. Adopt an Agile Approach

Iterative Development: Develop the system in phases, starting with a basic version and gradually adding features based on user feedback and performance data. Continuous Improvement: Regularly update the models and tagging mechanisms based on new insights and changing customer feedback patterns.

1. Invest in Training and Support

Training Programs: Provide training for support staff to effectively use the new system and understand how to leverage automated insights and responses. Support Resources: Ensure ongoing support for maintaining and troubleshooting the system to address any issues promptly.

1. Engage Stakeholders

Collaborate with Teams: Work closely with product managers, support teams, and data scientists to ensure the system aligns with organizational goals and customer needs. Gather Feedback: Regularly solicit feedback from end-users and stakeholders to refine and enhance the system's functionality and usability.

1. Ensure Data Privacy and Security

Compliance: Ensure that the system complies with relevant data privacy regulations and industry standards. Security Measures: Implement robust security measures to protect customer data and maintain the integrity of the support ticket information.

1. Leverage Advanced Analytics

Predictive Analytics: Use predictive analytics to forecast future ticket volumes and trends, allowing for proactive resource allocation and planning. Insight Generation: Extract actionable insights from ticket data to identify recurring issues, customer pain points, and opportunities for product improvement. By following these insights and recommendations, you will be able to create a robust support ticket categorization system that enhances your ability to manage customer feedback effectively, prioritize issues, and deliver timely and personalized responses.

*

Power Ahead

In [2]:

```
/bin/bash: -c: line 1: syntax error near unexpected token `('
/bin/bash: -c: line 1: `jupyter nbconvert --to html /content/NLP_Project_Low_Code_Copy (1).ipynb'
```

In [ ]: