

AI Workshop - Lab 2-2: Intent Classification

In this lab, we'll build a system to classify customer text messages into different categories (called **intents**) using a powerful type of AI model called a transformer. Transformers are a key technology behind tools like ChatGPT and other modern language systems.

Data Overview

We're working with a dataset of customer text messages that has already been labeled with their intent (e.g., "Order Status", "Product Inquiry", "Account Help"). The goal is to teach the model to recognize these patterns so it can classify new messages correctly.

- **Number of Categories:** 27 different intents.

What You'll Learn

- **Transformers:** Get an introduction to these models and why they're so powerful for language tasks.
- **Model Evaluation:** Understand how to measure a model's performance and interpret its predictions.

```
In [1]: !pip install -Uq datasets transformers accelerate evaluate sentencepiece sacrebleu
```

For this lab, it's essential that we have a GPU available to speed up training. On Google Colab, you can enable a GPU by going to **Runtime > Change runtime type > Hardware accelerator > GPU**.

The following line of code will check if a GPU is available:

```
In [2]: import torch
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
if device.type == 'cuda':
    print('GPU is available!')
else:
    print('GPU is not available. Enable a GPU runtime in Colab under "Runtime"
```

GPU is available!

Loading the Dataset

Now that we've set up our environment and imported the necessary packages, let's begin by loading our dataset.

In this lab, we'll work with a dataset of **customer text messages** that have been labeled with their **intent**. Each sample in the dataset includes a text message and a corresponding label indicating the intent behind the message (e.g., inquiry, complaint, order request). This dataset will allow us to build and evaluate models for intent classification.

Steps:

1. Load the Dataset:

- Use the `load_dataset` function from the `datasets` library to download and load the dataset.
- The dataset we're using is hosted at `"alexwaolson/customer-intents"`.

2. Inspect the Dataset:

- After loading, examine the training split (`intents['train']`) to understand its structure and the data it contains.

```
In [3]: from datasets import load_dataset
import pandas as pd

# Load the customer intents dataset
intents = load_dataset("alexwaolson/customer-intents")

# Display the training split
pd.DataFrame(intents['train'])
```

```
Out[3]:
```

	message	label
0	acn uhelp me to download my bill from Anna Fre...	get invoice
1	I need to notify of payment errors	payment issue
2	I'm trying to open an standard account for my ...	create account
3	checking invoices from Mr. Jones	check invoice
4	I am waiting for a restitution of 1200 dollars	track refund
...
1550	mail me my invoice from Anna Freeman	get invoice
1551	I have problems setting up another delivery ad...	set up shipping address
1552	is it possible to check what payment options y...	check payment methods
1553	have a question about a lost padsword	recover password
1554	getting refund of money	get refund

1555 rows × 2 columns

The dataset consists of two key columns:

- **message** : Contains the text of the customer message.
- **label** : Contains the intent category for each message.

There are **27 possible intent categories** in this dataset. To understand the distribution of these categories, we can count the number of examples for each intent. This helps us determine whether the dataset is balanced (i.e., whether all categories have similar representation) or imbalanced (some categories have significantly more or fewer samples than others).

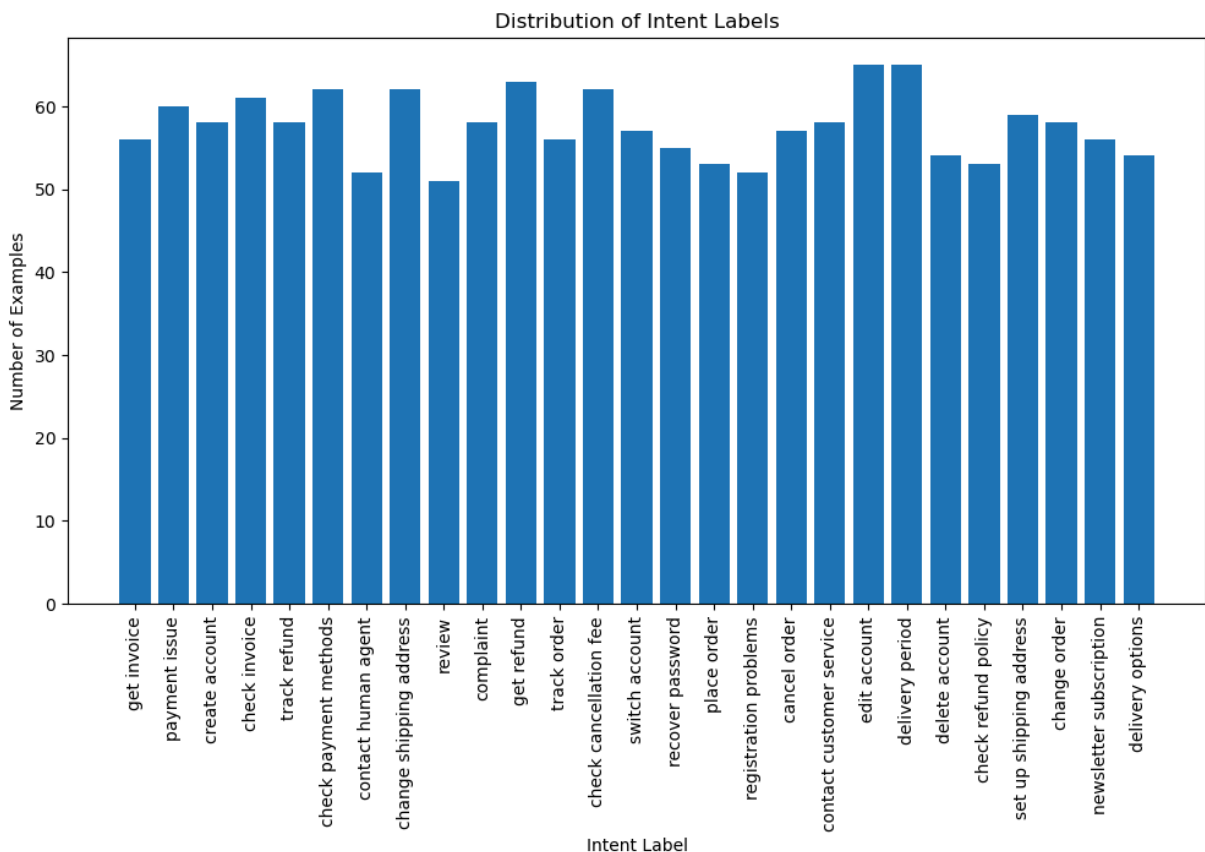
Run the code below to calculate the distribution of intent labels:

```
In [4]: from collections import Counter
import matplotlib.pyplot as plt

# Count the occurrences of each intent label in the training data
label_counts = Counter(intents['train']['label'])
print(f'Number of unique intents: {len(label_counts)}')

# Plot the distribution of intent labels
plt.figure(figsize=(12, 6))
plt.bar(label_counts.keys(), label_counts.values())
plt.xlabel('Intent Label')
plt.ylabel('Number of Examples')
plt.title('Distribution of Intent Labels')
plt.xticks(rotation=90)
plt.show()
```

Number of unique intents: 27



Zero-Shot Learning

One of the most powerful features of large language models is their ability to perform **zero-shot learning**. Unlike traditional models that require task-specific training, a zero-shot learning model can classify text based on its general understanding of language, even if it hasn't been explicitly trained on that specific task.

How It Works:

- Instead of fine-tuning the model, you provide it with a **prompt** that describes the task and possible labels (e.g., "What is the intent of this message?").
- The model uses its pre-trained knowledge to predict the most appropriate label.

This approach leverages the model's extensive training on a wide variety of text, making it flexible for many tasks.

Why Use Zero-Shot Learning?

- **Quick Prototyping:** No need to preprocess or fine-tune the model for every new task.
- **Versatility:** Works for tasks the model wasn't explicitly trained on, as long as the task can be described in a prompt.

Model Selection:

For zero-shot classification, we'll use the `flan-t5-large` model, which is well suited for this task due to its size and broad understanding of language. Since this model doesn't require fine-tuning, we can focus on testing its performance directly.

Zero-Shot Intent Classification with Flan-T5

We'll now use the **Flan-T5 large** model to classify intents via zero-shot learning. This approach involves crafting a **prompt** that describes the task and provides the model with the possible labels. The model then uses its language understanding to predict the intent without task-specific training.

Prompt Construction

The prompt is key to zero-shot learning. For our task:

1. The prompt begins by instructing the model to classify the intent of the message.
2. It lists the available intent categories.
3. Finally, it appends the message to classify.

```
In [5]: # Define the prompt
prompt = "Classify the intent of the following message using these categories"
```

```

for label in label_counts.keys():
    prompt += f"- {label}\n"
prompt += "Message: "

print(prompt)

```

Classify the intent of the following message using these categories:

- get invoice
- payment issue
- create account
- check invoice
- track refund
- check payment methods
- contact human agent
- change shipping address
- review
- complaint
- get refund
- track order
- check cancellation fee
- switch account
- recover password
- place order
- registration problems
- cancel order
- contact customer service
- edit account
- delivery period
- delete account
- check refund policy
- set up shipping address
- change order
- newsletter subscription
- delivery options

Message:

```

In [6]: from transformers import T5Tokenizer, T5ForConditionalGeneration

# Load the tokenizer and model
tokenizer = T5Tokenizer.from_pretrained("google/flan-t5-large")
model = T5ForConditionalGeneration.from_pretrained("google/flan-t5-large", c

```

You are using the default legacy behaviour of the `<class 'transformers.models.t5.tokenization_t5.T5Tokenizer'>`. This is expected, and simply means that the `legacy` (previous) behavior will be used so nothing changes for you. If you want to use the new behaviour, set `legacy=False`. This should only be set if you understand what it means, and thoroughly read the reason why this was added as explained in <https://github.com/huggingface/transformers/pull/24565>

```

In [7]: # Function for zero-shot classification
def zero_shot_intent_classification(model, prompt, message):
    # Combine the prompt and the message
    input_text = prompt + message
    # Tokenize the input text
    input_ids = tokenizer(input_text, return_tensors="pt").input_ids.to(device)
    # Generate a prediction

```

```

    output = model.generate(input_ids, max_length=50, num_beams=5, early_stop)
    # Decode the prediction into text
    return tokenizer.decode(output[0], skip_special_tokens=True)

# Test the function
zero_shot_intent_classification(model, prompt, "I need to cancel my order")

```

Out[7]: 'cancel order'

Testing Zero-Shot Intent Classification

You can now test the zero-shot classification capabilities of the `flan-t5-large` model on a subset of messages from the test set. This will provide a sense of how well the model performs without task-specific training.

```

In [8]: for message in intents['test']['message'][25:35]:
        print(f"Message: {message}")
        print(f"Predicted Intent: {zero_shot_intent_classification(model, prompt)}")
        print()

```

Message: delete Gold account
Predicted Intent: delete account

Message: I am trying to unsubscribe to the newsletter
Predicted Intent: newsletter subscription

Message: what do I need to do to change to the free account?
Predicted Intent: switch account

Message: open another standard account
Predicted Intent: create account

Message: I'd like to switch to the damn Premium account how to do it
Predicted Intent: switch account

Message: editing standard account
Predicted Intent: edit account

Message: wanna order several items help me
Predicted Intent: place order

Message: problems with standard account terminations
Predicted Intent: contact customer service

Message: want help to earn some of ur product
Predicted Intent: place order

Message: I need help notifying of a trouble with online payment
Predicted Intent: payment issue

Predicting Intent At Scale

To evaluate the performance of the `flan-t5-large` zero-shot model on the entire test dataset, we'll:

1. **Generate Predictions:** Use the `zero_shot_intent_classification` function to predict intents for all test messages.
2. **Compare Predictions:** Compare the zero-shot predictions to the true labels in the test set.
3. **Examine mis-classified text:** Look at incorrectly classified examples to see if we can understand what went wrong.

```
In [9]: from tqdm import tqdm

zero_shot_predictions = [zero_shot_intent_classification(model, prompt, message) for message, prompt in zip(messages, prompts)]
true_labels_text = intents['test']['label']
```

```
0%|
| 0/389 [00:00<?, ?it/s]
0%|█
| 1/389 [00:00<00:41, 9.33it/s]
1%|█
| 2/389 [00:00<00:41, 9.34it/s]
1%|██
| 3/389 [00:00<00:41, 9.34it/s]
1%|███
| 4/389 [00:00<00:41, 9.34it/s]
1%|████
| 5/389 [00:00<00:41, 9.35it/s]
2%|█████
| 6/389 [00:00<00:40, 9.40it/s]
2%|██████
| 7/389 [00:00<00:40, 9.36it/s]
2%|███████
| 8/389 [00:00<00:40, 9.41it/s]
2%|████████
| 9/389 [00:00<00:40, 9.43it/s]
3%|█████████
| 11/389 [00:01<00:38, 9.77it/s]
3%|██████████
| 12/389 [00:01<00:39, 9.62it/s]
3%|███████████
| 13/389 [00:01<00:39, 9.51it/s]
4%|███████████
| 14/389 [00:01<00:39, 9.47it/s]
4%|████████████
| 15/389 [00:01<00:39, 9.43it/s]
4%|█████████████
| 16/389 [00:01<00:39, 9.44it/s]
4%|█████████████
| 17/389 [00:01<00:39, 9.44it/s]
```

5%|██████████
| 18/389 [00:01<00:39, 9.37it/s]
5%|██████████
| 19/389 [00:02<00:39, 9.36it/s]
5%|██████████
| 20/389 [00:02<00:39, 9.36it/s]
6%|██████████
| 22/389 [00:02<00:37, 9.77it/s]
6%|██████████
| 23/389 [00:02<00:37, 9.67it/s]
6%|██████████
| 24/389 [00:02<00:38, 9.57it/s]
6%|██████████
| 25/389 [00:02<00:38, 9.52it/s]
7%|██████████
| 26/389 [00:02<00:38, 9.51it/s]
7%|██████████
| 28/389 [00:02<00:36, 9.86it/s]
7%|██████████
| 29/389 [00:03<00:36, 9.78it/s]
8%|██████████
| 30/389 [00:03<00:37, 9.64it/s]
8%|██████████
| 31/389 [00:03<00:37, 9.62it/s]
8%|██████████
| 32/389 [00:03<00:37, 9.57it/s]
8%|██████████
| 33/389 [00:03<00:37, 9.54it/s]
9%|██████████
| 34/389 [00:03<00:37, 9.50it/s]
9%|██████████
| 35/389 [00:03<00:37, 9.46it/s]
9%|██████████
| 36/389 [00:03<00:37, 9.47it/s]
10%|██████████
| 37/389 [00:03<00:37, 9.47it/s]
10%|██████████
| 38/389 [00:03<00:37, 9.47it/s]
10%|██████████
| 39/389 [00:04<00:36, 9.46it/s]
10%|██████████
| 40/389 [00:04<00:37, 9.42it/s]
11%|██████████
| 41/389 [00:04<00:36, 9.43it/s]
11%|██████████
| 42/389 [00:04<00:36, 9.45it/s]
11%|██████████
| 43/389 [00:04<00:36, 9.41it/s]

Iteration	Time (s)	Throughput (it/s)
45/389	00:04:00.35	9.82it/s
46/389	00:04:00.35	9.72it/s
47/389	00:04:00.35	9.63it/s
48/389	00:05:00.35	9.56it/s
49/389	00:05:00.35	9.49it/s
50/389	00:05:00.35	9.45it/s
51/389	00:05:00.35	9.45it/s
52/389	00:05:00.35	9.40it/s
53/389	00:05:00.35	9.39it/s
54/389	00:05:00.35	9.41it/s
55/389	00:05:00.35	9.42it/s
57/389	00:05:00.32	10.28it/s
59/389	00:06:00.32	10.30it/s
61/389	00:06:00.32	9.97it/s
62/389	00:06:00.33	9.83it/s
63/389	00:06:00.33	9.75it/s
64/389	00:06:00.33	9.66it/s
65/389	00:06:00.33	9.57it/s
67/389	00:06:00.32	9.88it/s
68/389	00:07:00.32	9.76it/s
69/389	00:07:00.33	9.68it/s
70/389	00:07:00.33	9.62it/s
72/389	00:07:00.31	9.93it/s
73/389	00:07:00.32	9.81it/s

[illegible]

[illegible]

[illegible]

[illegible]

```

42%|██████████          | 165/389 [00:17<00:24,  9.20it/s]
43%|██████████          | 166/389 [00:17<00:24,  9.27it/s]
43%|██████████          | 167/389 [00:17<00:23,  9.33it/s]
43%|██████████          | 168/389 [00:17<00:23,  9.35it/s]
43%|██████████          | 169/389 [00:17<00:23,  9.35it/s]
44%|██████████          | 170/389 [00:17<00:23,  9.35it/s]
44%|██████████          | 171/389 [00:17<00:23,  9.39it/s]
44%|██████████          | 172/389 [00:18<00:23,  9.43it/s]
44%|██████████          | 173/389 [00:18<00:22,  9.41it/s]
45%|██████████          | 175/389 [00:18<00:21,  9.84it/s]
45%|██████████          | 176/389 [00:18<00:21,  9.72it/s]
46%|██████████          | 178/389 [00:18<00:21, 10.02it/s]
46%|██████████          | 179/389 [00:18<00:21,  9.89it/s]
46%|██████████          | 180/389 [00:18<00:21,  9.75it/s]
47%|██████████          | 181/389 [00:18<00:22,  9.25it/s]
47%|██████████          | 182/389 [00:19<00:22,  9.27it/s]
47%|██████████          | 183/389 [00:19<00:22,  9.29it/s]

```

[illegible]

[illegible]


```
56%|███████████| 219/389 [00:22<00:17, 9.46it/s]
57%|███████████| 220/389 [00:23<00:17, 9.44it/s]
57%|███████████| 221/389 [00:23<00:17, 9.45it/s]
57%|███████████| 222/389 [00:23<00:17, 9.42it/s]
57%|███████████| 223/389 [00:23<00:17, 9.46it/s]
58%|███████████| 224/389 [00:23<00:17, 9.48it/s]
58%|███████████| 225/389 [00:23<00:17, 9.48it/s]
58%|███████████| 226/389 [00:23<00:17, 9.43it/s]
58%|███████████| 227/389 [00:23<00:17, 9.40it/s]
59%|███████████| 229/389 [00:23<00:16, 9.83it/s]
59%|███████████| 230/389 [00:24<00:16, 9.69it/s]
59%|███████████| 231/389 [00:24<00:16, 9.61it/s]
60%|███████████| 232/389 [00:24<00:16, 9.58it/s]
60%|███████████| 233/389 [00:24<00:16, 9.52it/s]
60%|███████████| 234/389 [00:24<00:16, 9.48it/s]
60%|███████████| 235/389 [00:24<00:16, 9.49it/s]
61%|███████████| 236/389 [00:24<00:16, 9.46it/s]
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
81%|███████████ | 317/389 [00:  
33<00:07, 9.62it/s]  
82%|███████████ | 318/389 [00:  
33<00:07, 9.57it/s]  
82%|███████████ | 319/389 [00:  
33<00:07, 9.54it/s]  
82%|███████████ | 320/389 [00:  
33<00:07, 9.49it/s]  
83%|███████████ | 321/389 [00:  
33<00:07, 9.47it/s]  
83%|███████████ | 322/389 [00:  
33<00:07, 9.47it/s]  
83%|███████████ | 324/389 [00:  
33<00:06, 10.27it/s]  
84%|███████████ | 326/389 [00:  
34<00:06, 9.62it/s]  
84%|███████████ | 328/389 [00:  
34<00:06, 9.86it/s]  
85%|███████████ | 329/389 [00:  
34<00:06, 9.75it/s]  
85%|███████████ | 330/389 [00:  
34<00:06, 9.69it/s]  
85%|███████████ | 331/389 [00:  
34<00:06, 9.60it/s]  
85%|███████████ | 332/389 [00:  
34<00:05, 9.54it/s]
```

[illegible]


```
89%|███████████| 347/389 [00:36<00:04, 9.35it/s]
90%|███████████| 349/389 [00:36<00:04, 9.78it/s]
90%|███████████| 350/389 [00:36<00:04, 9.65it/s]
90%|███████████| 351/389 [00:36<00:03, 9.62it/s]
90%|███████████| 352/389 [00:36<00:03, 9.55it/s]
91%|███████████| 353/389 [00:36<00:03, 9.46it/s]
91%|███████████| 354/389 [00:37<00:03, 9.46it/s]
91%|███████████| 355/389 [00:37<00:03, 9.42it/s]
92%|███████████| 356/389 [00:37<00:03, 9.38it/s]
92%|███████████| 357/389 [00:37<00:03, 9.38it/s]
92%|███████████| 358/389 [00:37<00:03, 9.38it/s]
92%|███████████| 359/389 [00:37<00:03, 9.38it/s]
93%|███████████| 360/389 [00:37<00:03, 9.38it/s]
```

[illegible]

[illegible]

```
In [10]: from sklearn.metrics import accuracy_score

print(f'Accuracy: {accuracy_score(true_labels_text, zero_shot_predictions)}')
```

Accuracy: 0.712082262210797

Incredibly, our accuracy using zero-shot learning is around **70%**, even without training on the categories first! Let's take a look at accuracy by category to see if there are any that the model struggles on. The **classification report** will break down the performance of the model by category, allowing us to understand if some categories are less well supported by the model than others. It provides us with the following information:

- **Precision:** This measures the proportion of correctly predicted positive observations to the total predicted positives. High precision indicates that the model makes few false positive errors.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

- **Example:** If the task is to classify emails as "spam," a **true positive** is an email correctly classified as spam, while a **false positive** is a legitimate email incorrectly classified as spam.
- **Recall:** This measures the proportion of correctly predicted positive observations to all the actual positives. High recall indicates that the model captures most of the true positive cases.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- **Example:** In the same email classification task, a **true positive** is an email correctly classified as spam, while a **false negative** is a spam email incorrectly classified as legitimate.
- **F1 Score:** This is the harmonic mean of precision and recall, balancing the two metrics. A high F1 score indicates a good trade-off between precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Support:** This refers to the number of actual occurrences of each category in the dataset. It helps us understand the distribution of the categories and whether any are underrepresented, which can impact performance metrics.

```
In [11]: from sklearn.metrics import classification_report

print(classification_report(true_labels_text, zero_shot_predictions, zero_di
```

	precision	recall	f1-score	support
cancel order	0.75	0.64	0.69	14
change order	1.00	0.50	0.67	18
change shipping address	0.28	1.00	0.43	8
check cancellation fee	1.00	0.73	0.84	11
check invoice	1.00	0.46	0.63	13
check payment methods	1.00	1.00	1.00	13
check refund policy	0.79	0.94	0.86	16
complaint	1.00	0.54	0.70	13
contact customer service	0.16	0.73	0.26	11
contact human agent	0.74	1.00	0.85	20
create account	0.73	1.00	0.84	16
delete account	1.00	0.78	0.88	18
delivery options	1.00	0.18	0.30	17
delivery period	1.00	0.60	0.75	10
edit account	0.83	1.00	0.91	10
get invoice	0.93	0.93	0.93	14
get refund	0.42	1.00	0.59	10
newsletter subscription	1.00	1.00	1.00	11
payment issue	0.73	1.00	0.85	11
place order	0.76	0.94	0.84	17
recover password	1.00	0.65	0.79	20
registration problems	1.00	0.65	0.79	17
review	0.95	0.95	0.95	19
set up shipping address	0.00	0.00	0.00	11
switch account	1.00	0.37	0.54	19
track order	0.87	0.81	0.84	16
track refund	1.00	0.12	0.22	16
accuracy			0.71	389
macro avg	0.81	0.72	0.70	389
weighted avg	0.84	0.71	0.71	389

Let's now look at some mis-classified examples to see if we can understand why they were not classified correctly.

```
In [12]: # Display mis-classified examples
misclassified_examples = [(message, true_label, pred) for message, true_label, pred in
pd.DataFrame(misclassified_examples, columns=['Message', 'True Label', 'Pred
```

Out[12]:

	Message	True Label	Predicted Label
0	I have to see how long it takes for the packag...	delivery period	track order
1	how do I file a customer claim against your or...	complaint	contact customer service
2	question about cancelling purchase 732201349959	cancel order	contact customer service
3	I xepect a compensation of \$1499	track refund	get refund
4	problems with standard account terminations	delete account	contact customer service
...
107	reporting issues creating user	registration problems	create account
108	problems with canceling purchase 370795561790	cancel order	contact customer service
109	can you help me to use the Pro profile?	switch account	contact customer service
110	what do I have to do to file a claim?	complaint	contact customer service
111	help to locate bill #85632	check invoice	payment issue

112 rows × 3 columns

Flagging Abuse

One of the challenges in customer service is identifying and handling abusive messages. Even in this dataset there are examples where customers have used inappropriate language in their requests.

```
In [13]: # Display abusive examples
pd.DataFrame([(message, label) for message, label in zip(intents['test']['me
```

Out[13]:

	Message	Label
0	i have got to request goddamn erimbursements o...	get refund
1	I'd like to switch to the damn Premium account...	switch account
2	i cannot sign up where to notify of goddamn si...	registration problems
3	I don't know how to demand a goddamn restituti...	get refund
4	can you help me informing of a goddamn signup ...	registration problems
5	I need a damn Platinum account, could I open it?	create account
6	could you help me to close the damn free account?	delete account
7	I want to delete a goddamn item from order 732...	change order
8	I call to check what damn shipping methods are...	delivery options
9	where can I see what damn shipment options you...	delivery options

Let's say that we want to introduce a new task to classify messages as abusive or not. We can use the same zero-shot approach to classify messages as abusive or not abusive. The prompt will be similar to the previous one, but with the new task and labels.

```
In [14]: # Define the prompt for abusive language classification
abuse_prompt = "Classify the following as inappropriate if it contains cursi

# Load the tokenizer and model
abuse_tokenizer = T5Tokenizer.from_pretrained("google/flan-t5-large")
abuse_model = T5ForConditionalGeneration.from_pretrained("google/flan-t5-lar

# Function for zero-shot classification of abusive language
def zero_shot_abuse_classification(model, prompt, message):
    # Combine the prompt and the message
    input_text = prompt + message
    # Tokenize the input text
    input_ids = abuse_tokenizer(input_text, return_tensors="pt").input_ids.t
    # Generate a prediction
    output = model.generate(input_ids, max_length=50, num_beams=5, early_sto
    # Decode the prediction into text
    return abuse_tokenizer.decode(output[0], skip_special_tokens=True)

# Test the function
zero_shot_abuse_classification(abuse_model, abuse_prompt, "I'm so damn frust
```

Out[14]: 'Inappropriate'

Now we can use the model to flag messages as inappropriate or not inappropriate:

```
In [15]: # Predict for samples that contain the word "damn"
for message in intents['test']['message']:
```

```

if 'damn' in message.lower():
    print(f"Message: {message}")
    print(f"Predicted Intent: {zero_shot_abuse_classification(abuse_model(message))}")
    print()

```

Message: i have got to request goddamn erimbursements of my money
Predicted Intent: Inappropriate

Message: I'd like to switch to the damn Premium account how to do it
Predicted Intent: Inappropriate

Message: i cannot sign up where to notify of goddamn signup problems
Predicted Intent: Inappropriate

Message: I don't know how to demand a goddamn restitution of my money
Predicted Intent: Inappropriate

Message: can you help me informing of a goddamn signup issue?
Predicted Intent: Inappropriate

Message: I need a damn Platinum account, could I open it?
Predicted Intent: Inappropriate

Message: could you help me to close the damn free account?
Predicted Intent: Inappropriate

Message: I want to delete a goddamn item from order 732201349959
Predicted Intent: Inappropriate

Message: I call to check what damn shipping methods are available
Predicted Intent: Inappropriate

Message: where can I see what damn shipment options you offer?
Predicted Intent: Inappropriate

Translation

Finally, let's explore how we can use transformers for translation tasks. We'll start by translating English messages to French using the `translation_en_to_fr` pipeline. Then we'll look at translating to Japanese using a different model.

```

In [16]: from transformers import pipeline

en_fr_translator = pipeline("translation_en_to_fr", device='cuda')

# Translate a sample message
en_message = "I need help with my order."

en_fr_translator(en_message)

```

No model was supplied, defaulted to google-t5/t5-base and revision a9723ea (<https://huggingface.co/google-t5/t5-base>).
Using a pipeline without specifying a model name and revision in production is not recommended.


```
Out[16]: [{'translation_text': "J'ai besoin d'aide pour faire ma commande."}]
```

```
In [17]: # Translate first 50 messages in the test set
en_messages = intents['test']['message'][:50]
fr_messages = en_fr_translator(en_messages)

# Display the translations
for en_message, fr_message in zip(en_messages, fr_messages):
    print(f"EN: {en_message}\nFR: {fr_message['translation_text']}\n")
```

EN: I do not know how I can get the bill from Anna Freeman
FR: Je ne sais pas comment obtenir la facture d'Anna Freeman

EN: help me to check how long it takes for my item to arrive
FR: Aidez-moi à vérifier combien de temps il faut pour que mon article arrive

EN: I have to see how long it takes for the package to arrive
FR: Je dois voir combien de temps il faut pour que le colis arrive.

EN: how do I file a customer claim against your organization?
FR: Comment puis-je déposer une réclamation à l'encontre de votre organisation?

EN: I need help to see in what cases can I ask for refunds
FR: J'ai besoin d'aide pour savoir dans quels cas je peux demander des remboursements

EN: i try to talk to an agent
FR: j'essaie de parler à un agent

EN: i dont know what i have to do to buy some of ur product
FR: Je ne sais pas ce que je dois faire pour acheter une partie de votre produit.

EN: help to notify of an payment error
FR: aider à notifier une erreur de paiement

EN: chat wth somebody
FR: chat avec quelqu'un

EN: i have got to request goddamn erimbursements of my money
FR: i have to request goddamn erimbursements of my money

EN: how could i talk with an assistant
FR: Comment puis-je parler à un assistant?

EN: wanna switch an item of order 113542617735902 ohw to do it
FR: wanna switch an item of order 113542617735902 ohw to do it

EN: i do not know how i could cancel order 00123842
FR: Je ne sais pas comment je pourrais annuler la commande 00123842

EN: want help tracking order 113542617735902
FR: Vous voulez obtenir de l'aide pour suivre votre commande 113542617735902

EN: question about cancelling purchase 732201349959
FR: question sur l'annulation de l'achat 732201349959

EN: help me modify order 370795561790
FR: Aidez-moi à modifier la commande 370795561790

EN: where to retrieve the pass of my account?
FR: où récupérer la carte de crédit de mon compte ?

EN: I do not have a fucking Freemium account, could I register?

FR: Je n'ai pas de compte Freemium, puis-je m'inscrire ?

EN: I expect a compensation of \$1499

FR: I expect une indemnité de 1499 \$

EN: I have got to call customer service, how to do it?

FR: Je dois appeler le service à la clientèle, comment le faire ?

EN: reimbursing 1499 dollars

FR: remboursement de 1499 dollars

EN: where do i make a consumer claim against ur business

FR: Où puis-je déposer une réclamation de consommateur contre votre entreprise?

EN: can you help me canceling my subscription to the newsletter?

FR: pouvez-vous m'aider à annuler mon abonnement au bulletin ?

EN: i do nolt know how to open a Platinum account for my wife

FR: Je ne sais pas comment ouvrir un compte Platinum pour ma femme.

EN: want help to see in what cases can I ask for reimbursements

FR: Vous voulez obtenir de l'aide pour savoir dans quels cas je peux demander des remboursements

EN: delete Gold account

FR: Supprimer le compte Or

EN: I am trying to unsubscribe to the newsletter

FR: J'essaie de m'abstenir du bulletin d'information

EN: what do I need to do to change to the free account?

FR: Que dois-je faire pour changer de compte gratuit ?

EN: open anotherstandard account

FR: ouvrir un autre compte standard

EN: I'd like to switch to the damn Premium account how to do it

FR: J'aimerais passer au compte Premium. Comment le faire?

EN: editing standard account

FR: édition du compte standard

EN: wanna order several items help me

FR: j'aimerais commander plusieurs articles m'aider

EN: problems with standard account terminations

FR: problèmes liés à la cessation des comptes normalisés

EN: want help ti earn some of ur product

FR: Vous voulez obtenir de l'aide pour gagner une partie de votre produit

EN: I need help notifying of a trouble with online payment

FR: J'ai besoin d'aide pour signaler un problème de paiement en ligne

EN: shipping to Japan

FR: transport vers le Japon

EN: problem with my forgotten password

FR: problème avec mon mot de passe oublié

EN: I have got to speak with a human agent

FR: Je dois parler à un agent humain.

EN: where do i open a Platinum account

FR: Où ouvrir un compte Platinum?

EN: I want to buy several of your product, how can I do it?

FR: Je veux acheter plusieurs de vos produits, comment puis-je le faire?

EN: change purchase 113542617735902

FR: changement d'achat 113542617735902

EN: modifying details on Gold account

FR: Modifier les détails du compte Gold

EN: i dont know how to report an issue with sign-up

FR: Je ne sais pas comment signaler un problème d'inscription

EN: check order 113542617735902 status

FR: Ordonnance de vérification 113542617735902 état

EN: assistance to check the cancellation fee

FR: assistance pour vérifier les frais d'annulation

EN: I cannot check your money back guarantee

FR: Je ne peux pas vérifier votre garantie de remboursement

EN: I expect a reimbursement of \$1499, has it been processed?

FR: Je m'attends à un remboursement de 1499 \$, a-t-il été traité?

EN: wanna delete several items from purchase 00123842 help me

FR: veux supprimer plusieurs articles de l'achat 00123842 aide-moi

EN: want help to see at what time I can call customer support

FR: Vous voulez obtenir de l'aide pour voir à quel moment je peux appeler le support à la clientèle

EN: I don't know what to do to reset my user account PIN code

FR: Je ne sais pas ce qu'il faut faire pour réinitialiser mon NIP pour mon compte utilisateur

We can also try translation to Japanese, although be aware that this model is not as effective as the English to French model above.

```
In [18]: from transformers import AutoModelForSeq2SeqLM, AutoTokenizer

# Load the model and tokenizer
model_name = "Helsinki-NLP/opus-mt-en-jap"
tokenizer = AutoTokenizer.from_pretrained(model_name)
```

```

model = AutoModelForSeq2SeqLM.from_pretrained(model_name)

def translate_message(text):
    inputs = tokenizer(text, return_tensors="pt")
    outputs = model.generate(**inputs)
    return tokenizer.decode(outputs[0], skip_special_tokens=True)

```

In [19]: `translate_message("This is a test message.")`

Out[19]: 'これ が た め に , あ い さ つ で あ る .'

```

In [20]: # Translate first 50 messages in the test set
en_messages = intents['test']['message'][:50]

for en_message in en_messages:
    print(f"EN: {en_message}\nJP: {translate_message(en_message)}\n")

```

EN: I do not know how I can get the bill from Anna Freeman

JP: わたし は 知 ら な い の か , 報 告 を し て み る が よ い .

EN: help me to check how long it takes for my item to arrive

JP: わたし を 久 し く 攻 め る た め に , わたし を 助 け て く だ さ い .

EN: I have to see how long it takes for the package to arrive

JP: わたし は 見 て い る 間 は , ど う し て い る か を 知 ろ う .

EN: how do I file a customer claim against your organization?

JP: わたし は , な ん の た め に , あ な た の 寛 容 を む さ ぼ り わ ず ら っ て い る の か .

EN: I need help to see in what cases can I ask for refunds

JP: わたし は , ど ん な 場 合 に も , 足 る こ と が で き る か を 見 る こ と が で き る .

EN: i try to talk to an agent

JP: 世 に あ る こ と を 語 る た め に 語 る べ き で あ っ て ,

EN: i dont know what i have to do to buy some of ur product

JP: わたし た ち の 中 の 誓 っ て い る こ と を , あ ら か じ め 知 り に し て も ら い た い .

EN: help to notify of an payment error

JP: 誘 惑 に 陥 ら な い よ う に す る た め で あ る .

EN: chat wth somebody

JP: ア ビ ム は , く れ た 者 は , む ち を も っ て 道 を し て い る .

EN: i have got to request goddamn erimburements of my money

JP: か え っ て 神 に 祈 る こ と を 求 め て , そ の 金 を わ た し の 銀 で じ ゅ う ぶ ん に 与 え た .

EN: how could i talk with an assistant

JP: わたし は , さ と い う こ と が あ る の で , ど う し て 言 葉 を 出 す こ と が で き る か .

EN: wanna switch an item of order 113542617735902 ohw to do it

JP: いの中には"クタ・クミン"というように、クミンの上にあって、これをなし、Jerusalemすことを好まなければならない。

EN: i do not know how i could cancel order 00123842

JP: 知らないのか、どうしてその事が、明らかにすることができないか。

EN: want help tracking order 113542617735902

JP: 慎重深く、寛容であって、慎重深く、寛容を示してください。

EN: question about cancelling purchase 732201349959

JP: ザフカイとじっては、どこでもあるかをみななければならない。

EN: help me modify order 370795561790

JP: わたしを助けて、メソポン・ルコナとして、わたしを助けてください。

EN: where to retrieve the pass of my account?

JP: わたしの計りごとは、どこにあるのか。わたしの言うところはどこにあるか。

EN: I do not have a fucking Freemium account, could I register?

JP: わたしは、自制することをしているのではないか。わたしの誇ることができましょうか。

EN: I xepect a compensation of \$1499

JP: わたしは持っているくびきを摘み、持っているばを"。

EN: I have got to call customer service, how to do it?

JP: わたしは礼拝をするためにきたのであるが、何をしたらよいであろうか"。

EN: reimbursing 1499 dollars

JP: キッコデ・ハッコク、カルクムを生む。

EN: where do i make a consumer claim against ur business

JP: また、ここでは、わたしたちの働きをくつがえしていることを、決定している。

EN: can you help me canceling my subscription to the newsletter?

JP: あなたはわたしを助け、おとずれを伝えることができるか。

EN: i do nolt know how to open a Platinum account for my wife

JP: わたしの妻がどういう場合についても、前を開いてはならない。

EN: want help to see in what cases can I ask for reimbursements

JP: 助けを得るために、わたしはどんなにしたらよければならないかを見て、その勧告を望することができるだろう。

EN: delete Gold account

JP: 気をつけて、それを正しなければならない。

EN: I am trying to unsubscribe to the newsletter

JP: わたしは、福音をゆだねることができないようにと、努めているのである。

EN: what do I need to do to change to the free account?

JP: 自由を得させるために、何をしたらよいでしょうか。

EN: open another standard account

JP: もう一つの開いたことがある。しかし、ほかのどの点を開くこともできる。

EN: I'd like to switch to the damn Premium account how to do it

JP: わたしは、これがためすべきこと、また、どうしてこのような理由をしているかのように、

EN: editing standard account

JP: 次に旗を立てた者、

EN: wanna order several items help me

JP: 貧しい命じて、わたしを助けさせよ。

EN: problems with standard account terminations

JP: 目負われた者たちと共に、あらかじめあかしをしていた者である。

EN: want help ti earn some of ur product

JP: クエキリスポからは、わたしたちの中のある者たちが監督者となっている。

EN: I need help notifying of a trouble with online payment

JP: わたしは、忍耐を弁明のためには、助けを受けることがない。

EN: shipping to Japan

JP: アグライにはベバルに、

EN: problem with my forgotten password

JP: わが心の影を忘れて言いました。

EN: I have got to speak with a human agent

JP: 人間については、前のことを語るべきである。

EN: where do i open a Platinum account

JP: アスントからは、告訴のころがある。

EN: I want to buy several of your product, how can I do it?

JP: わたしはあなたの身から価を買おうとしているので、どうしてこれをするのか。

EN: change purchase 113542617735902

JP: 買収てびとシベノブテイ、シベノン、シャフル、

EN: modifying details on Gold account

JP: それは決定して悪いことであって、慎重深く行動しなければならない。

EN: i dont know how to report an issue with sign-up

JP: かつ、このようなしるしをわきまえているのが、どうしてわかるだろうか、

EN: check order 113542617735902 status

JP: シャフカテ・シクシャフル・シャシャン, シャハル,

EN: assistance to check the cancellation fee

JP: というのは, わたしたちがかたく所に出ることができるのものであっても,

EN: I cannot check your money back guarantee

JP: わたしは, あなたの金を散らしてはいけません.

EN: I expect a reimbursement of \$1499, has it been processed?

JP: わたしは, これがあわてさせるという望んだのであろうか. その勧告をしはじめたので, 今なお忍耐しているであらうか.

EN: wanna delete several items from purchase 00123842 help me

JP: (見よ, 人々がわたしをしえたげるときから, わたしを助けてください.)

EN: want help to see at what time I can cll customer support

JP: わたしは, どうして訓練することができるかを見ることができるか.

EN: I don't know what to do to reset my user account PIN code

JP: わたしは, 自分の推薦をするために, どんな事したらよいか, わかりません.