L3 Exercise 3 - Custom Form 1040 Recognizer Exercise

November 26, 2022

0.0.1 Please install the required Python modules/SDKs

```
In [2]: ! activate ai-azure-c1

        import sys

        sys.path.append("/opt/conda/envs/ai-azure-c1/lib/python3.8/site-packages")
```

0.0.2 Importing Azure Form Recognizer Python modules

```
In [3]: import os
        from azure.core.exceptions import ResourceNotFoundError
        from azure.ai.formrecognizer import FormRecognizerClient
        from azure.ai.formrecognizer import FormTrainingClient
        from azure.core.credentials import AzureKeyCredential
```

0.0.3 Train the boarding pass model

```
In [4]: endpoint = "https://udacity-form-recognizer.cognitiveservices.azure.com/"
        key = "62bb6572766845bc8cf5a5c386ae3542"

In [5]: form_training_client = FormTrainingClient(endpoint=endpoint, credential=AzureKeyCredenti

In [6]: trainingDataUrl = "https://udacitystoragetest.blob.core.windows.net/boarding-pass?sp=rac

In [7]: training_process =  form_training_client.begin_training(trainingDataUrl, use_training_la
        boarding_pass_model = training_process.result()

In [8]: boarding_pass_model.status

Out[8]: 'ready'

In [9]: boarding_pass_model.training_started_on

Out[9]: datetime.datetime(2022, 11, 26, 18, 52, 54, tzinfo=<isodate.tzinfo.Utc object at 0x7f423

In [10]: boarding_pass_model.training_documents
```

```
Out[10]: [TrainingDocumentInfo(name=boarding-cedrick-cash.pdf, status=succeeded, page_count=1, e
         TrainingDocumentInfo(name=boarding-charles-clark.pdf, status=succeeded, page_count=1,
         TrainingDocumentInfo(name=boarding-johnny-bishop.pdf, status=succeeded, page_count=1,
         TrainingDocumentInfo(name=boarding-kevin-crim.pdf, status=succeeded, page_count=1, err
         TrainingDocumentInfo(name=boarding-lesa-thomas.pdf, status=succeeded, page_count=1, er
         TrainingDocumentInfo(name=boarding-martha-dillon.pdf, status=succeeded, page_count=1,
         TrainingDocumentInfo(name=boarding-michael-graves.pdf, status=succeeded, page_count=1,
         TrainingDocumentInfo(name=boarding-pete-coover.pdf, status=succeeded, page_count=1, er
         TrainingDocumentInfo(name=boarding-robin-reed.pdf, status=succeeded, page_count=1, err
         TrainingDocumentInfo(name=boarding-wayne-cushing.pdf, status=succeeded, page_count=1,
```

0.0.4 Predict on boarding passes

```
In [11]: new_test_url = ["https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/bo

         test_urls = [
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-avkas
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-james
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-james
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-libby
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-radha
             "https://udacitystoragetest.blob.core.windows.net/test-boarding-pass/boarding-samee
         ]

In [12]: form_recognizer_client = FormRecognizerClient(endpoint=endpoint, credential=AzureKeyCre

In [13]: def recognize_and_display_boarding_passes(client,custom_model,urls):
             for url in urls:
                 boarding_pass_model_results = client.begin_recognize_custom_forms_from_url(mode
                 for content in boarding_pass_model_results:
                     print("Form type: {}".format(content.form_type))
                     for name, field in content.fields.items():
                         print("Label '{}' has value '{}' and a confidence score of {}".format(
                             field.label_data.text if field.label_data else name,
                             field.value,
                             field.confidence
                         ))

In [14]: recognize_and_display_boarding_passes(form_recognizer_client,boarding_pass_model,test_u

Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Class' has value 'E' and a confidence score of 0.994
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'Boarding Pass ID' has value 'ETK-34236751E' and a confidence score of 0.99
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Destination' has value 'Chicago' and a confidence score of 0.99
Label 'PassengerName' has value 'Avkash Chauhan' and a confidence score of 0.99
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
```

Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Seat' has value '20A' and a confidence score of 0.99
Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Boarding Pass ID' has value 'ETK-34236748E' and a confidence score of 0.99
Label 'Seat' has value '1A' and a confidence score of 0.99
Label 'PassengerName' has value 'James Webb' and a confidence score of 0.99
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Class' has value 'E' and a confidence score of 0.994
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
Label 'Destination' has value 'Chicago' and a confidence score of 0.991
Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Seat' has value '25B' and a confidence score of 0.99
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Class' has value 'E' and a confidence score of 0.994
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Destination' has value 'Chicago' and a confidence score of 0.991
Label 'Boarding Pass ID' has value 'ETK-34236750E' and a confidence score of 0.99
Label 'PassengerName' has value 'James Jackson' and a confidence score of 0.991
Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Seat' has value '3D' and a confidence score of 0.99
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'PassengerName' has value 'Libby Herold' and a confidence score of 0.99
Label 'Destination' has value 'Chicago' and a confidence score of 0.991
Label 'Class' has value 'B' and a confidence score of 0.994
Label 'Boarding Pass ID' has value 'ETK-34236749E' and a confidence score of 0.99
Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Boarding Pass ID' has value 'ETK-34236747E' and a confidence score of 0.99
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Destination' has value 'Chicago' and a confidence score of 0.99

```
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Seat' has value '34B' and a confidence score of 0.99
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'Class' has value 'E' and a confidence score of 0.994
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
Label 'PassengerName' has value 'Radha S. Kumar' and a confidence score of 0.99
Form type: custom:3ea1cb18-a0e3-4f4f-866b-720c05c8f581
Label 'Gate' has value 'G1' and a confidence score of 0.99
Label 'Carrier' has value 'UA' and a confidence score of 0.993
Label 'Bag' has value 'YES' and a confidence score of 0.994
Label 'Boarding Pass ID' has value 'ETK-34236748E' and a confidence score of 0.99
Label 'Seat' has value '34A' and a confidence score of 0.99
Label 'PassengerName' has value 'Sameer Kumar' and a confidence score of 0.99
Label 'Flight Number' has value '234' and a confidence score of 0.995
Label 'Departure Location' has value 'San Francisco' and a confidence score of 0.99
Label 'Boarding Time' has value '10:00 AM PST' and a confidence score of 0.99
Label 'Class' has value 'E' and a confidence score of 0.994
Label 'Date' has value 'April 20, 2022' and a confidence score of 0.99
Label 'Destination' has value 'Chicago' and a confidence score of 0.99
```

In [ ]: