Ground TruthとSageMakerで行う
飲食店メニュー画像分類について

株式会社ぐるなび
データインテグレーショングループ
長谷川正彦

自己紹介

- データインテグレーショングループ
  - ぐるなびのビッグデータを使ったサービスの創出・運用
  - 弊社エンジニアブログにも一部掲載
    - https://developers.gnavi.co.jp/entry/datalibrary/

店舗レコメンドシステム

- ユーザーの嗜好に合った店舗をレコメンドするシステム
- アクセスログの共起頻度から算出
- 2018年にレコメンドシステムをリニューアル
  - 表示面の拡充
  - 複数のモデル適用
    - リニューアル前は単一ロジックのみ

店舗レコメンドシステムの課題

- 出面によってCTRにバラツキがある（0.3%～30%）
- アクセス数の少ないエリアや記事コンテンツのCTRが低い

CTR向上のための施策を検討

- レコメンド精度を上げるために複数の施策を検討
  - 共起頻度算出にskip-gramを適用
  - メニューテキストにTF-IDFを適用
  - ユーザーの嗜好に合った画像を優先的に表示したい

ユーザーの嗜好に合った画像を優先的に表示したい

- メニュー画像の中でユーザーが興味を持ちそうな画像を優先的に表示する
  - 閲覧店舗の業態（居酒屋、フレンチなど）の中から閲覧されやすい画像を分類する
  - ユーザーのログと組み合わせ、ユーザーが過去に閲覧した画像に類似した画像を表示する、など

類似度を算出するために画像をベクトル化

メニュー画像をベクトル化して類似度を計算する

1. CNNを使った画像分類器を作成
2. 作成した画像分類器から重みベクトルを抽出する
3. 画像同士の重みベクトルから類似度を算出する

モデル作成

- 類似度算出のために、重みベクトルを出力する必要がある
- AWSのビルトインアルゴリズムでは重み抽出ができない
- Tensorflowの画像分類モデルであるinception v2を使用

教師データの準備

- 教師データの作成を効率的に行いたい
- ぐるなび内のメニュー画像登録数はすごく多い (300万超)
- 複数の作業者で作業を共有したい
- いつでもアクセスできる環境で、空き時間を使ってアノテーションしたい

すべてSageMaker Ground Truthで解決できる

ぐるなび GURUNAVI

SageMaker Ground Truthとは

- アノテーションの一般的なワークフローをサポート
- 4種類の組み込みラベリングツールを提供
- アノテーション作業を行うワーカーとの連携・管理機能を提供
- 大規模データセットに対しては自動ラベリング機能で最大７０%のコスト削減

（AWS公式ドキュメントより引用）

メニュー画像分類タスク

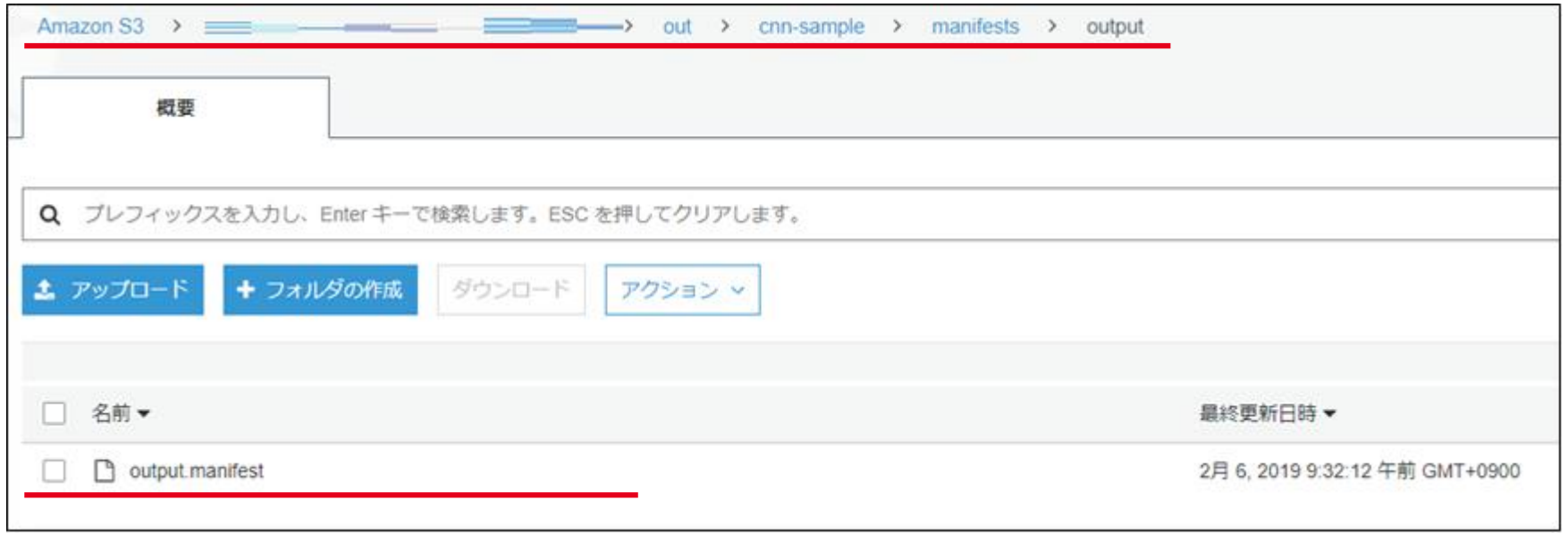
- 飲食店のメニュー画像を分類したい
  - コース料理、ドリンク、スイーツなどのカテゴリ分類を行う
  - Ground Truthのラベリングジョブの「画像分類」を利用

アノテーション作業

- チームメンバーで作業を共有
  - プライベートワーカータイプを選択
  - 登録済みチームメンバー全員に、ラベリングタスク用のURLリンクが送られる

ラベリングジョブ完了

- 結果出力
  - ラベリングジョブの結果は拡張マニフェストファイルとしてS3上に出力される
  - 拡張マニフェストファイルは、そのまま教師データとしてsagemaker.tensorflowのパイプラインに使用できる



拡張マニフェストファイルの出力サンプル


```
{
        "source-ref": "s3://sagemaker-xxxxxxxx/data/train_data/course/XXXXXXXX.jpg",
        "cnn-sample": 0,
        "cnn-sample-metadata": {
                "confidence": 0.68,
                "job-name": "labeling-job/cnn-sample",
                "class-name": "course",
                "human-annotated": "yes",
                "creation-date": "2019-02-06T00:28:45.196810",
                "type": "groundtruth/image-classification"
        }
}
```


source-ref：教師画像のS3上のパス
cnn-sample:トレーニングジョブでアノテーションされた正解ラベル（cnn-sampleはジョブ名）
metadata属性には対応するクラス名などが格納される
カスタムモデルにはsource-refとcnn-sampleをモデルに渡す必要がある。

Tensorflowカスタムトレーニングモデルの作成

- モデルの生成
  - TensorFlowモデルをイニシャライズしてトレーニングジョブを生成する
  - Ground Truthを使用する場合、入力はパイプラインモードを指定
  - input_modeにPipeを指定する

```
import sagemaker
from sagemaker.tensorflow import TensorFlow

model = TensorFlow(
        input_mode = 'Pipe’,
        entry_point='cnn-train.py’,
        role=role,
        train_instance_type='ml.p2.xlarge’,
        framework_version='1.12.0’,
        train_instance_count=1
)
```

Tensorflowカスタムトレーニングモデルの作成

- 教師データの作成
  - パイプラインモードでトレーニングを行うため、session.s3_inputクラスを使って教師データを初期化する
  - その際の引数にはパイプラインモード用の設定を指定する。
  - 拡張マニフェストファイルのどの属性を教師データとして使うかをattribute_namesに指定する。

```
train_data = sagemaker.session.s3_input(
        'path/to/manifest/file’,
        distribution='FullyReplicated’,
        content_type='image/jpeg’,
        record_wrapping='RecordIO’,
        s3_data_type='AugmentedManifestFile’,
        attribute_names=['source-ref','cnn-sample’]
)
data_channels = {'train': train_data,'eval':eval_data}
model.fit(data_channels)
```

カスタムモデルスクリプト

- SageMakerでTensorFlowのモデルをトレーニングする場合、カスタムモデルスクリプトを記述する必要がある。
- TensorFlowのカスタムスクリプトを構築する場合、最低限以下の関数を実装する必要がある。
  - train_input_fn（トレーニングデータをモデルに渡す際に使用）
  - eval_input_fn（検証データをモデルに渡す際に使用）
  - model_fn（構築するカスタムモデル）

train_input_fn

- train_input_fnはモデルのトレーニングデータ入力を受け持つ
  - カスタムモデルの入力に、Ground Truthでアノテーションした結果を渡す
- パイプラインモードでは、データはRecordIOフォーマットで渡ってくるため、PipeModeDataSetとしてmodel_fnに引き渡す必要がある。

train_input_fnの記述例

```
def train_input_fn ('train',num_epochs,batch_size):
        ds = PipeModeDataset('train')
        def combine(records):
                return (records[0], records[1])
        ds = ds.batch(2)
        ds = ds.map(combine)
        def parse(data, label):
                parsed0 = tf.image.decode_jpeg(data, channels=3)
                parsed0 = tf.image.resize_images(parsed0, [224, 224])
                parsed0 = tf.cast(parsed0, dtype=tf.float32)
                parsed0 = parsed0 / 255
                parsed1 = tf.strings.to_number(label, out_type=tf.int64)
                return ({'inputs': parsed0}, parsed1)
        if num_epochs > 1:
                ds = ds.repeat(num_epochs)
                ds = ds.shuffle(32*batch_size)
        ds = ds.prefetch(batch_size)
        ds = ds.apply(tf.data.experimental.map_and_batch(parse, batch_size=batch_size,
                        num_parallel_batches=4))
        return ds
```

PipeModeDatasetを生成

ds.batchにはattribute_namesで指定したsource-refとcnn-sampleが引き渡される

データをパースしてCNNのインプット用に加工

model_fn

- model_fnにはCNNのモデルを記述する
  - 今回はTensorflow_hubのinception_v2を使用したい。
- SageMakerのトレーニングジョブ用コンテナには Tensoflow_hubが入っていない
  - ジョブコンテナ起動時にインストールする必要がある
  - OpenCVを使う場合も同様にインストールする必要がある

model_fnにtensorflow_hubを利用する

- pipmainを使い、動的にライブラリのインストールを行う。

```
from pip._internal import main as pipmain
pipmain(['install', 'tensorflow_hub’])
pipmain(['install', 'opencv-python’])

import tensorflow_hub as hub
import cv2

def model_fn(features, labels, mode, params):
        module =
        hub.Module("https://tfhub.dev/google/imagenet/inception_v2/feature_vector/1")
        ...
```

pipmainを使ってtensorflow_hubを動的にインストール

hub.Module()でinceptionモデルを読み込む

カスタムスクリプトができたら、model.fit()を呼び出して学習を開始

Ground Truthでラベリングジョブを作る際の注意点（１）

- 「教師データに含めない」ことができない
  - 明らかに不要な教師データについても、必ずいずれかのカテゴリに分類する必要がある
  - 例えば料理画像の中に、ごくわずかだけ関係ない画像（車など）が入っている場合など
  - これを回避するため、ラベリングジョブ作成時に予め「学習に使わない」ラベルを作っておき、完了した拡張マニフェストファイルから不要ラベルに属しているデータを削除する

拡張マニフェストファイルから不要なカテゴリを削除する

拡張マニフェストファイルのサンプル

class-nameに着目

{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/menu/ga6sm3xz0000_t_0n5p.jpg","cnn-sample":4,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"other","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/sweets/sfghj69d0000_t_0p2h.jpg","cnn-sample":8,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"sweets","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/meat/nvm3d10c0000_t_0nbl.jpg","cnn-sample":3,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"meat","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/kushi/8d5erz0e0000_t_0n8m.jpg","cnn-sample":2,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"kushi","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/drink/1evca6yx0000_t_0n9k.jpg","cnn-sample":1,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"drink","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/drink/h2vbgz190000_t_0n7t.jpg","cnn-sample":1,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"drink","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/restaurant/sfn4ay570000_t_001p.jpg","cnn-sample":6,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/menu/sptsrrbz0000_t_0n6f.jpg","cnn-sample":4,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"other","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/drink/ce67b4cu0000_t_0ntz.jpg","cnn-sample":1,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"drink","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/meat/srzn33240000_t_00jz.jpg","cnn-sample":3,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"meat","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/menu/sr50vhg60000_t_0n7x.jpg","cnn-sample":4,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"other","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/drink/1kfgubz50000_t_0ndr.jpg","cnn-sample":1,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"drink","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/sushi_sashimi/gvym8nt70000_t_0n7i.jpg","cnn-sample":7,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"sushi_sashimi","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/menu/68nttutk0000_t_0n5q.jpg","cnn-sample":4,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"other","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/sweets/dxb5sgs50000_t_0n9e.jpg","cnn-sample":8,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"sweets","human-annotated":"yes","creation-date":"2019-02-06T00:28:45.196810","type":"groundtruth/image-classification"}}
{"source-ref":"s3://sagemaker-ap-northeast-1-XXXXXXXXXXXX/data/train_data/kushi/8rzcwfa40000_t_008c.jpg","cnn-sample":2,"cnn-sample-metadata":{"confidence":0.68,"job-name":"labeling-job/cnn-sample","class-name":"kushi

"class-name":"other",

ラベリングジョブ時、教師データとして不要な画像をclass-name:”other”に分類

↓

拡張マニフェストファイルを直接編集し、”other”だけを削除

ぐるなび GURUNAVI

Ground Truthでラベリングジョブを作る際の注意点（２）

- ジョブを停止または完了するまで拡張マニフェストファイルは出力されない
  - ラベリングジョブ実行中にトレーニングを開始できない
- 停止すると再開できない
  - ラベリングジョブは管理画面から停止できる
  - 停止するとそのジョブは再開できないため、再開するのためにはchainジョブを別途立ち上げる必要がある

Ground Truthでラベリングジョブを作る際の注意点（３）

- 停止ラベリングジョブをテストに使用したい
  - ラベリングが完了しないまま停止すると、未ラベリングの教師データについては属性が付与されない
  - 上記の状態でトレーニングするとattribute_namesが参照できずインプットエラーとなる
  - これを避けるため、拡張マニフェストファイルから属性が付与されていないレコードを削除する必要がある

画像の類似度判定モデル

- 今回の目的は分類モデルから重みを取り出し、画像同士の類似度を算出すること
- 学習済みトレーニングデータからSageMakerのエンドポイント生成機能を利用し、推論エンドポイントをデプロイ
- 推論エンドポイントには、重みベクトルを返却する処理を追加しておく
- メニュー画像をリクエストとして推論エンドポイントに送信し、重みベクトルを返却、それを基に類似度を算出する

結果について

- 分類はかなり高い精度でできている
- ただ、画像同士の類似性についてはもう一工夫必要
  - 正しく分類でき、かつ画像として似ていても類似しているとは言い難いケースが意外と多い
    - コース料理で見た目は似てるがジャンルが異なる（洋食と和食）

画像引用元：https://www.gnavi.co.jp

結果について

- 一見してどのジャンルに属するか迷う料理が結構ある
  - こぼれ寿司は寿司なのか？
  - スイーツのような盛り付けの洋食
  - 料理人の方々の創意工夫は素晴らしいです。

画像引用元：https://www.gnavi.co.jp

今後について

- ラベルの再精査を進めつつ、レコメンドエンジンへの導入を図る
- 分類精度自体は高いので、画像のメタデータ付与など他ジョブへの利用についても検討

まとめ

- Ground Truthで行うアノテーションはとても便利
  - 教師データさえあればすぐにアノテーションできる
  - 専用ツールのインストールが不要
  - WEBアプリケーションなので、どこでも作業可能
  - スマホのブラウザで動かないのでここは改善してほしい
- アノテーション作業をしているとものすごく腹が減る
- GroundTruthにはテキスト分類ジョブもあるので、テキストを使ったメニューの分類も進めていきたい