

## CMPE 258 – HW2 Abhishek Extra credit

### 1) Solve a Tabular regression problem using Auto AI/ML

#### a) Load and Inspect a tabular dataset

Google Cloud Platform My First Project Search products and resources

AI Platform (Unified) abais\_tabular\_regression

SOURCE ANALYZE

Dataset Info  
Created: Feb 22, 2021 10:33 PM  
Dataset format: CSV  
Dataset location: [gs://cloud-mi-.../bank-marketing.csv](#)

Summary  
Total columns: 17  
Total rows: 45,211

General statistics generated by Feb 25, 2021 10:10 AM [GENERATE STATISTICS](#)

Enter property name or value

Field Name	Missing % (count)	Distinct values
Age	-	77
Balance	-	7168
Campaign	-	48
Contact	-	3
Day	-	31
Default	-	2
Deposit	-	2
Duration	-	1573
Education	-	4

Training jobs and models  
abais\_tabular\_regression\_20212236352  
Model type: Tabular  
[TRAIN NEW MODEL](#)

#### b) Create a tabular regression model and evaluate it

Google Cloud Platform My First Project Search products and resources

AI Platform (Unified) abais\_tabular\_regression\_20212236352 [VIEW DATASET](#) [EXPORT](#)

EVALUATE DEPLOY & TEST BATCH PREDICTIONS MODEL PROPERTIES

Target column: Balance

MAE	MAPE	RMSE	RMSLE	R <sup>2</sup>
1,384.859	∞	2,554.933	-	0.06

Feature Importance

Feature	Importance
Age	High
Month	Medium-High
Day	Medium
Education	Medium-Low
Housing	Low-Medium
Loan	Low
Duration	Low

#### c) Examine R2 score for your chosen target variable (I chose Balance)

Google Cloud Platform My First Project Search products and resources

AI Platform (Unified) abais\_tabular\_regression\_20212236352 [VIEW DATASET](#) [EXPORT](#)

EVALUATE DEPLOY & TEST BATCH PREDICTIONS MODEL PROPERTIES

Target column: Balance

MAE	MAPE	RMSE	RMSLE	R <sup>2</sup>
1,384.859	∞	2,554.933	-	0.06

## d) Inspect model properties to see what is it optimized for (here RMSE)

[←](#) abais\_tabular\_regression\_20212236352 [VIEW](#)

[EVALUATE](#) [DEPLOY & TEST](#) [BATCH PREDICTIONS](#) [MODEL PROPERTIES](#)

**i** This model finished training on Feb 22, 2021, 11:43:21 PM.

Status	Succeeded
Model ID	6159534507545329664
Training pipeline ID	8169371394375680000
Created	Feb 22, 2021, 10:35:35 PM
Budget (original)	1 node hours
Training time	1 hr 7 min
Region	us-central1
Encryption type	Google-managed key

Dataset	<a href="#">abais_tabular_regression</a>
Dataset ID	6010704613609570304
Target column	Balance
Data split	Randomly assigned (80/10/10)
Model hyperparameters	<a href="#">Model</a> <a href="#">Trials</a>
Transformation options	<a href="#">View details</a>


Algorithm	AutoML
Objective	Tabular regression
Optimized for	RMSE

## e) Deploy model to endpoint

[←](#) abais\_tabular\_regression\_20212236352 [VIEW DATASET](#) [EXPORT](#)

[EVALUATE](#) [DEPLOY & TEST](#) [BATCH PREDICTIONS](#) [MODEL PROPERTIES](#)

### Use your edge-optimized model




Container

Export your model as a TF Saved Model to run on a Docker container.

### Deploy your model

Endpoints are machine learning models made available for online prediction requests. Endpoints are useful for timely predictions from many users (for example, in response to an application request). You can also request batch predictions if you don't need immediate results.

[DEPLOY TO ENDPOINT](#)

Endpoint	ID	Models	Region	Last updated
 <a href="#">abais_tabular_regression_ep</a>	6388180149564604416	1	us-central1	Feb 24, 2021, 10:21:20 AM

## f) Test model

### Test your model PREVIEW

Feature column name	Type	Required or optional	Value	Local feature importance
Age	Text	Required	<input type="text" value="38.000000"/>	0
Campaign	Text	Required	<input type="text" value="2.000000"/>	0
Contact	Text	Required	<input type="text" value="cellular"/>	0

Predict label

Prediction result

1082.2132568359375

95% prediction interval

[-395.3097839355469, 6374.0390625]

## 2) Image Object Detection

### a) Load an image dataset for object detection task

My First Project

Search products and resources

abais\_image\_object\_detection

abais\_image\_object\_dete...

IMPORT

BROWSE

ANALYZE

All 3,667

Labeled 0

Unlabeled 3,667











Filter labels

Images

ADD NEW LABEL

Filter items

Select all



No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

No bounding boxes

### b) Go to datasets-> analyze, validate the objective as Image object detection

abais\_image\_object\_detection

abais\_image\_object\_dete...

IMPORT

BROWSE

ANALYZE

After you create labels and assign images to those labels, label statistics will appear here.

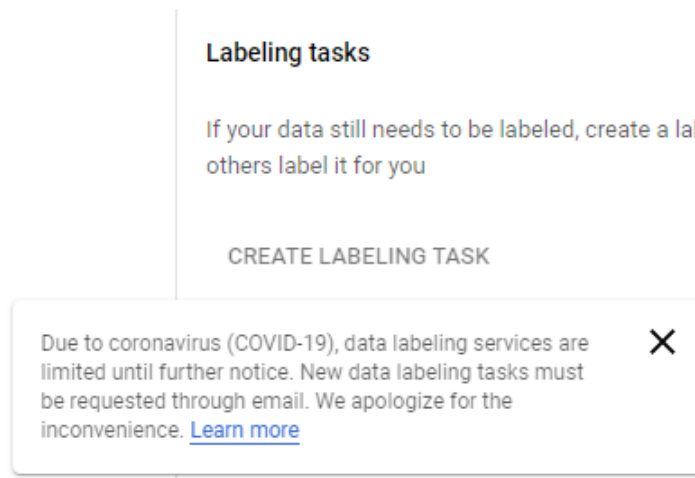
☐ Custom data split

You currently don't have any items assigned to Training, Validation, or Test sets. Don't worry, when you train a model, your items will be randomly divided into these sets, using recommended machine learning practices. [Learn more](#)

#### Properties

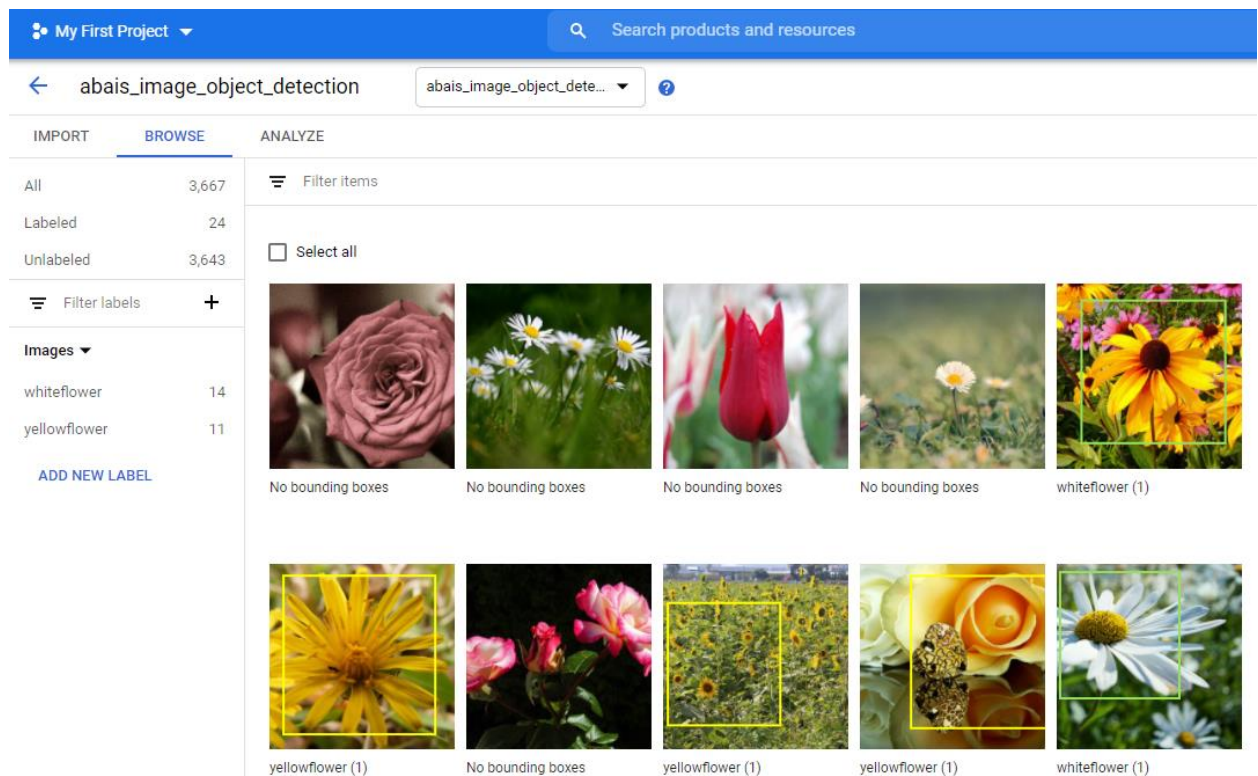
Annotation set	abais_image_object_detection_iod
Objective	Image object detection
Items	3667
Created	Feb 22, 2021, 10:11:58 PM
Last updated	Feb 22, 2021, 10:11:58 PM
Region	us-central1
Encryption type	Google-managed key

### c) Create a labelling task

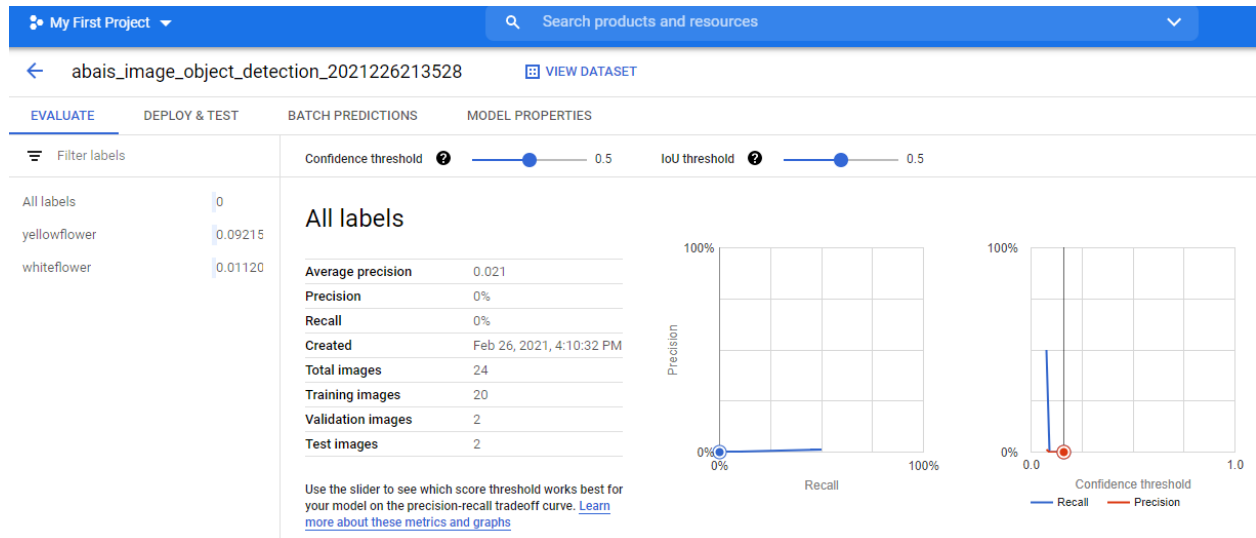


Google has limited data labelling due to covid-19 which suggests humans maybe involved in this task. This is also an important learning that – Labelling doesn't come cheap.

### d) AutoML requires atleast 2 labels of 10 images each. Manually create labels and assign images to labels. In this example, I created 2 labels namely whiteflower and yellowflower with > 10 images each/ label



- e) Create an image object detection model and train it with 80/10/10 split. Once training completes evaluate the model over various metrics such as confidence, precision, and recall



### 3) Text sentiment Analysis

#### a) Upload a text dataset for sentiment analysis

The screenshot shows the 'BROWSE' tab for a project named 'My First Project'. The dataset is 'abais\_text\_sentiment\_analysis'. The interface includes a 'Filter items' section on the left with a list of items: 'All' (11,947), 'Labeled' (0), and 'Unlabeled' (11,947). The main area displays a table of text items with their sentiment scores. The table has columns for 'Text' and 'Sentiment score'. The sentiment scores are 0 or 1. The table is filtered by 'Sentiment score'.

Text	Sentiment score
My eldest son who is 27 just got word he has a new job after finishing his bache...	0
I visited my best friend at her school on St. Patrick's day.	0
My mom cooked some delicious rice for me with curd.	0
Today I make Eye contact with my crush. She Also look into my Eyes For a Seco...	0
I was dropping off my son for a sleepover. He was really excited to go. I droppe...	0
Dinner tonight was really good.	0
I WENT TO MEENAKSHI AMMAN TEMPLE WITH MY FAMILY MEMBERS.	0
I got the test results back from my father's echo and neck arteries taken at the ...	0
I was selected as the winner for a random lottery drawing from an mturk hit. It ...	0
My brother told me he got into med school!	0

- b) Labelling tasks are currently disabled, so manually label the text. This is done by assigning sentiment scores to the text under browse. Once filtered one can filter by label (sentiment score) and inspect.

My First Project

Search products and resources

abais\_text\_sentiment\_analysis
abais\_text\_sentiment\_ana...

IMPORTBROWSEANALYZE

All11,947
Labeled143
Unlabeled11,804

Filter labels
060
183

Filter items

☐

Text

Sentiment score

☐

▼

My eldest son who is 27 just got word he has a new job after finishing his bache...

1

☐

▼

I visited my best friend at her school on St. Patrick's day.

1

☐

▼

My mom cooked some delicious rice for me with curd.

1

☐

▼

Today I make Eye contact with my crush. She Also look into my Eyes For a Seco...

1

☐

▼

I was dropping off my son for a sleepover. He was really excited to go. I droppe...

1

☐

▼

Dinner tonight was really good.

1

☐

▼

I WENT TO MEENAKSHI AMMAN TEMPLE WITH MY FAMILY MEMBERS.

1

☐

▼

I was selected as the winner for a random lottery drawing from an mturk hit. It ...

1

☐

▼

My brother told me he got into med school!

1

☐

▼

I was very happy sitting down to a home-cooked meal with a group of good frie...

1

c) Analyze the labeled text to see enough text are labelled. >10 texts per label are required to train the model.

My First Project

Search products and

abais\_text\_sentiment\_analysis
abais\_text\_sentiment\_ana...

IMPORTBROWSEANALYZE

Consider assigning more labels

For best results, it's recommended that at least 100 text items are assigned to each label. Consider assigning more text items to labels with less than the recommended number. [Learn more](#)

Here's what you can do:

- Assign scores: Assign more scores on the Browse tab or create a labeling task to do it automatically
- Add text items: Upload more text items on the Import tab

Labels
↑

Text items

0

60

1

83

d) Google cloud fails to split the labelled data into 50/30/20 split of train, test, validate. Several other splits like 80/10/10 and 70/20/10 were also tried. The following error is seen for the 50/30/20 random split.

My First Project Search products and resources

abais\_text\_sentiment\_analysis\_2021226222856 VIEW TRAINING DATASET

Training pipeline failed with error message: Text Sentiment training requires training, validation and test datasets to be non-empty. The number of data items selected for training is 152, for validation is 0, for test is 0.

Status	Failed
Training pipeline ID	7970790798306312192
Created	Feb 26, 2021, 2:29:25 PM
Start time	Feb 26, 2021, 2:29:26 PM
Elapsed time	13 sec
Region	us-central1
Encryption type	Google-managed key

Dataset	abais_text_sentiment_analysis
Annotation set	<a href="#">abais_text_sentiment_analysis_tsa</a>
Data split	Randomly assigned (50/30/20)

Algorithm	AutoML
Objective	Text sentiment analysis

#### 4) Video object tracking

a) Load a video dataset and manually label it ensuring at least 2 labels with  $\geq 15$  videos per label. I created abais\_gymnastics\_label, abais\_play\_golf\_label

My First Project Search products and resources

abais\_video\_object\_tracking abais\_video\_object\_tracki...

IMPORT BROWSE ANALYZE

All 500  
Labeled 31  
Unlabeled 469  
Training 4  
Test 4

Filter labels +

Videos

- abais\_gymnastics\_label 16
- abais\_play\_golf\_label 15

ADD NEW LABEL

Filter items

Select all

abais\_play\_golf\_label

abais\_gymnastics\_label

abais\_gymnastics\_label

abais\_gymnastics\_label

abais\_gymnastics\_label

abais\_gymnastics\_label

abais\_gymnastics\_label

## b) Create a video object tracking model and initiate training with 80/20 split and monitor progress of training pipeline under the training tab

Filter training pipelines...						
Name	ID	Job type	Model type	Status	Created	Elapsed time
✅ abais_image_object_detection_2021226213528	7988805196815794176	Training pipeline	🖼️ Image object detection	Succeeded	Feb 26, 2021, 1:36:35 PM	2 hr 33 min
❌ abais_video_object_tracking_2021226221818	6007221360772775936	Training pipeline	🎥 Video object tracking	Failed	Feb 26, 2021, 2:18:28 PM	58 min 58 sec
❌ abais_text_sentiment_analysis_2021226222856	7970790798306312192	Training pipeline	📄 Text sentiment analysis	Failed	Feb 26, 2021, 2:29:25 PM	13 sec
✅ abais_tabular_regression_20212236352	8169371394375680000	Training pipeline	📊 Tabular regression	Succeeded	Feb 22, 2021, 10:35:35 PM	1 hr 7 min

## c) Training pipeline failed with internal error, google cloud does not provide further details, inspection log underneath

My First Project 🔍

← abais\_video\_object\_tracking\_2021226221818 📄

❗ Training pipeline failed with error message: INTERNAL

Status	Failed
Training pipeline ID	6007221360772775936
Created	Feb 26, 2021, 2:18:28 PM
Start time	Feb 26, 2021, 2:18:28 PM
Elapsed time	58 min 58 sec
Region	us-central1
Encryption type	Google-managed key
Dataset	abais_video_object_tracking
Annotation set	<a href="#">abais_video_object_tracking_vot</a>
Data split	Randomly assigned (80/0/20)
Algorithm	AutoML
Objective	Video object tracking

## Training pipeline

My First Project 🔍 Search products and resources

Training **PREVIEW** + CREATE

TRAINING PIPELINE CUSTOM JOB HYPERPARAMETER TUNING

Training pipelines are the primary model training workflow in AI Platform (Unified). You can use training pipelines to create an AutoML-trained model or a custom-trained model. For custom-trained models, training pipelines orchestrate custom training jobs and hyperparameter tuning with additional steps like adding a dataset or uploading the model to AI Platform for prediction serving. [Learn More](#)

Region: us-central1 (Iowa) ⓘ

Filter training pipelines...						
Name	ID	Job type	Model type	Status	Created	Elapsed time
✅ abais_image_object_detection_2021226213528	7988805196815794176	Training pipeline	🖼️ Image object detection	Succeeded	Feb 26, 2021, 1:36:35 PM	2 hr 33 min
❌ abais_video_object_tracking_2021226221818	6007221360772775936	Training pipeline	🎥 Video object tracking	Failed	Feb 26, 2021, 2:18:28 PM	58 min 58 sec
❌ abais_text_sentiment_analysis_2021226222856	7970790798306312192	Training pipeline	📄 Text sentiment analysis	Failed	Feb 26, 2021, 2:29:25 PM	13 sec
✅ abais_tabular_regression_20212236352	8169371394375680000	Training pipeline	📊 Tabular regression	Succeeded	Feb 22, 2021, 10:35:35 PM	1 hr 7 min
✅ migrate-automl-abaisflowers_20210222012303	522681439565643776	Training pipeline	⚙️ Custom	Succeeded	Feb 22, 2021, 9:42:39 AM	
✅ abais_hello_custom_2	6214809156096884736	Training pipeline	⚙️ Custom	Succeeded	Feb 22, 2021, 9:08:03 AM	14 min 32 sec
✅ Structured_AutoML_Tutorial_202122243140	6261534002230853632	Training pipeline	📊 Tabular classification	Succeeded	Feb 21, 2021, 8:32:56 PM	1 hr 6 min
✅ abais_hello_video_dataset_202122233610	2522983361559920640	Training pipeline	🎥 Video classification	Succeeded	Feb 21, 2021, 7:36:33 PM	1 hr 59 min
✅ abais_hello_text_dataset_20212221474	6491780533180170240	Training pipeline	📄 Text classification (Single-label)	Succeeded	Feb 21, 2021, 5:48:01 PM	3 hr 44 min
✅ hello_custom	1526561944004198400	Training pipeline	⚙️ Custom	Succeeded	Feb 21, 2021, 8:13:52 PM	15 min 7 sec

Rows per page: 10 1 - 10 of many