Content-Based Filtering Using Neural Networks

This notebook relies on files created in the content_based_preproc.ipynb notebook. Be sure to run the code in there before completing this notebook.
Also, we'll be using the **python3** kernel from here on out so don't forget to change the kernel if it's still Python2.

This lab illustrates:

1. how to build feature columns for a model using tf.feature_column
2. how to create custom evaluation metrics and add them to Tensorboard
3. how to train a model and make predictions with the saved model

Tensorflow Hub should already be installed. You can check that it is by using "pip freeze".

In [1]:
```
%%bash
pip freeze | grep tensor
```

```
tensorboard==2.7.0
tensorboard-data-server==0.6.1
tensorboard-plugin-wit==1.8.0
tensorflow @ file:///opt/conda/conda-bld/dlenv-tf-1-15-cpu_1634862472464/work/te
nsorflow-1.15.5-cp37-cp37m-linux_x86_64.whl
tensorflow-cloud==0.1.13
tensorflow-data-validation==0.23.1
tensorflow-datasets==1.2.0
tensorflow-estimator==1.15.1
tensorflow-hub==0.6.0
tensorflow-io==0.8.1
tensorflow-metadata==0.23.0
tensorflow-model-analysis==0.23.0
tensorflow-probability==0.8.0
tensorflow-serving-api==1.15.0
tensorflow-transform==0.23.0
```

Let's make sure we install the necessary version of tensorflow-hub. After doing the pip install below, click **"Restart the kernel"** on the notebook so that the Python environment picks up the new packages.

In [2]:
```
!pip3 install tensorflow-hub==0.7.0
!pip3 install --upgrade tensorflow==1.15.3
!pip3 install google-cloud-bigquery==1.10
```

```
Collecting tensorflow-hub==0.7.0
  Downloading tensorflow_hub-0.7.0-py2.py3-none-any.whl (89 kB)
     |████████████████████████████████| 89 kB 5.9 MB/s
Requirement already satisfied: six>=1.10.0 in /opt/conda/lib/python3.7/site-pack
ages (from tensorflow-hub==0.7.0) (1.16.0)
Requirement already satisfied: protobuf>=3.4.0 in /opt/conda/lib/python3.7/site-
packages (from tensorflow-hub==0.7.0) (3.18.1)
Requirement already satisfied: numpy>=1.12.0 in /opt/conda/lib/python3.7/site-pa
ckages (from tensorflow-hub==0.7.0) (1.19.5)
Installing collected packages: tensorflow-hub
  Attempting uninstall: tensorflow-hub
    Found existing installation: tensorflow-hub 0.6.0
```

```
    Uninstalling tensorflow-hub-0.6.0:
      Successfully uninstalled tensorflow-hub-0.6.0
Successfully installed tensorflow-hub-0.7.0
Collecting tensorflow==1.15.3
  Downloading tensorflow-1.15.3-cp37-cp37m-manylinux2010_x86_64.whl (110.5 MB)
     |████████████████████████████████| 110.5 MB 24 kB/s
Requirement already satisfied: absl-py>=0.7.0 in /opt/conda/lib/python3.7/site-p
ackages (from tensorflow==1.15.3) (0.8.1)
Requirement already satisfied: astor>=0.6.0 in /opt/conda/lib/python3.7/site-pac
kages (from tensorflow==1.15.3) (0.8.1)
Requirement already satisfied: gast==0.2.2 in /opt/conda/lib/python3.7/site-pack
ages (from tensorflow==1.15.3) (0.2.2)
Requirement already satisfied: keras-preprocessing>=1.0.5 in /opt/conda/lib/pyth
on3.7/site-packages (from tensorflow==1.15.3) (1.1.2)
Requirement already satisfied: protobuf>=3.6.1 in /opt/conda/lib/python3.7/site-
packages (from tensorflow==1.15.3) (3.18.1)
Requirement already satisfied: termcolor>=1.1.0 in /opt/conda/lib/python3.7/site
-packages (from tensorflow==1.15.3) (1.1.0)
Requirement already satisfied: wheel>=0.26 in /opt/conda/lib/python3.7/site-pack
ages (from tensorflow==1.15.3) (0.37.0)
Collecting tensorboard<1.16.0,>=1.15.0
  Downloading tensorboard-1.15.0-py3-none-any.whl (3.8 MB)
     |████████████████████████████████| 3.8 MB 47.1 MB/s
Requirement already satisfied: six>=1.10.0 in /opt/conda/lib/python3.7/site-pack
ages (from tensorflow==1.15.3) (1.16.0)
Requirement already satisfied: wrapt>=1.11.1 in /opt/conda/lib/python3.7/site-pa
ckages (from tensorflow==1.15.3) (1.13.2)
Requirement already satisfied: grpcio>=1.8.6 in /opt/conda/lib/python3.7/site-pa
ckages (from tensorflow==1.15.3) (1.41.0)
Requirement already satisfied: google-pasta>=0.1.6 in /opt/conda/lib/python3.7/s
ite-packages (from tensorflow==1.15.3) (0.2.0)
Requirement already satisfied: opt-einsum>=2.3.2 in /opt/conda/lib/python3.7/sit
e-packages (from tensorflow==1.15.3) (3.3.0)
Requirement already satisfied: numpy<2.0,>=1.16.0 in /opt/conda/lib/python3.7/si
te-packages (from tensorflow==1.15.3) (1.19.5)
Requirement already satisfied: keras-applications>=1.0.8 in /opt/conda/lib/pytho
n3.7/site-packages (from tensorflow==1.15.3) (1.0.8)
Requirement already satisfied: tensorflow-estimator==1.15.1 in /opt/conda/lib/py
thon3.7/site-packages (from tensorflow==1.15.3) (1.15.1)
Requirement already satisfied: h5py in /opt/conda/lib/python3.7/site-packages (f
rom keras-applications>=1.0.8->tensorflow==1.15.3) (2.10.0)
Requirement already satisfied: markdown>=2.6.8 in /opt/conda/lib/python3.7/site-
packages (from tensorboard<1.16.0,>=1.15.0->tensorflow==1.15.3) (3.3.4)
Requirement already satisfied: werkzeug>=0.11.15 in /opt/conda/lib/python3.7/sit
e-packages (from tensorboard<1.16.0,>=1.15.0->tensorflow==1.15.3) (2.0.2)
Requirement already satisfied: setuptools>=41.0.0 in /opt/conda/lib/python3.7/si
te-packages (from tensorboard<1.16.0,>=1.15.0->tensorflow==1.15.3) (58.2.0)
Requirement already satisfied: importlib-metadata in /opt/conda/lib/python3.7/si
te-packages (from markdown>=2.6.8->tensorboard<1.16.0,>=1.15.0->tensorflow==1.1
5.3) (4.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in /opt/conda/lib/python
3.7/site-packages (from importlib-metadata->markdown>=2.6.8->tensorboard<1.16.0,
>=1.15.0->tensorflow==1.15.3) (3.10.0.2)
Requirement already satisfied: zipp>=0.5 in /opt/conda/lib/python3.7/site-packag
es (from importlib-metadata->markdown>=2.6.8->tensorboard<1.16.0,>=1.15.0->tenso
rflow==1.15.3) (3.6.0)
Installing collected packages: tensorboard, tensorflow
  Attempting uninstall: tensorboard
    Found existing installation: tensorboard 2.7.0
    Uninstalling tensorboard-2.7.0:
      Successfully uninstalled tensorboard-2.7.0
```

```
  Attempting uninstall: tensorflow
    Found existing installation: tensorflow 1.15.5
    Uninstalling tensorflow-1.15.5:
      Successfully uninstalled tensorflow-1.15.5
ERROR: pip's dependency resolver does not currently take into account all the pa
ckages that are installed. This behaviour is the source of the following depende
ncy conflicts.
explainable-ai-sdk 1.3.2 requires xai-image-widget, which is not installed.
tfx 0.23.0 requires attrs<20,>=19.3.0, but you have attrs 21.2.0 which is incomp
atible.
tfx 0.23.0 requires click<8,>=7, but you have click 8.0.3 which is incompatible.
tfx 0.23.0 requires docker<5,>=4.1, but you have docker 5.0.3 which is incompati
ble.
tfx 0.23.0 requires google-api-python-client<2,>=1.7.8, but you have google-api-
python-client 2.27.0 which is incompatible.
tfx 0.23.0 requires google-resumable-media<0.7.0,>=0.6.0, but you have google-re
sumable-media 2.0.3 which is incompatible.
tfx 0.23.0 requires kubernetes<12,>=10.0.1, but you have kubernetes 18.20.0 whic
h is incompatible.
tfx 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyarrow 5.0.0 which is inc
ompatible.
tfx 0.23.0 requires pyyaml<6,>=3.12, but you have pyyaml 6.0 which is incompatib
le.
tfx-bsl 0.23.0 requires google-api-python-client<2,>=1.7.11, but you have google
-api-python-client 2.27.0 which is incompatible.
tfx-bsl 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyarrow 5.0.0 which is
incompatible.
tensorflow-model-analysis 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyar
row 5.0.0 which is incompatible.
tensorflow-data-validation 0.23.1 requires joblib<0.15,>=0.12, but you have jobl
ib 1.1.0 which is incompatible.
tensorflow-data-validation 0.23.1 requires pyarrow<0.18,>=0.17, but you have pya
rrow 5.0.0 which is incompatible.
tensorflow-cloud 0.1.13 requires tensorboard>=2.3.0, but you have tensorboard 1.
15.0 which is incompatible.
Successfully installed tensorboard-1.15.0 tensorflow-1.15.3
Collecting google-cloud-bigquery==1.10
  Downloading google_cloud_bigquery-1.10.0-py2.py3-none-any.whl (88 kB)
     |████████████████████████████████| 88 kB 4.8 MB/s
Requirement already satisfied: google-resumable-media>=0.3.1 in /opt/conda/lib/p
ython3.7/site-packages (from google-cloud-bigquery==1.10) (2.0.3)
Collecting google-cloud-core<0.30dev,>=0.29.0
  Downloading google_cloud_core-0.29.1-py2.py3-none-any.whl (25 kB)
Collecting google-api-core<2.0.0dev,>=1.6.0
  Downloading google_api_core-1.31.3-py2.py3-none-any.whl (93 kB)
     |████████████████████████████████| 93 kB 2.0 MB/s
Requirement already satisfied: setuptools>=40.3.0 in /opt/conda/lib/python3.7/si
te-packages (from google-api-core<2.0.0dev,>=1.6.0->google-cloud-bigquery==1.10)
(58.2.0)
Requirement already satisfied: googleapis-common-protos<2.0dev,>=1.6.0 in /opt/c
onda/lib/python3.7/site-packages (from google-api-core<2.0.0dev,>=1.6.0->google-
cloud-bigquery==1.10) (1.53.0)
Collecting google-auth<2.0dev,>=1.25.0
  Downloading google_auth-1.35.0-py2.py3-none-any.whl (152 kB)
     |████████████████████████████████| 152 kB 41.6 MB/s
Collecting protobuf<3.18.0,>=3.12.0
  Downloading protobuf-3.17.3-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.
whl (1.0 MB)
     |████████████████████████████████| 1.0 MB 48.2 MB/s
Requirement already satisfied: six>=1.13.0 in /opt/conda/lib/python3.7/site-pack
ages (from google-api-core<2.0.0dev,>=1.6.0->google-cloud-bigquery==1.10) (1.16.
```

```
0)
Requirement already satisfied: requests<3.0.0dev,>=2.18.0 in /opt/conda/lib/pyth
on3.7/site-packages (from google-api-core<2.0.0dev,>=1.6.0->google-cloud-bigquer
y==1.10) (2.25.1)
Requirement already satisfied: packaging>=14.3 in /opt/conda/lib/python3.7/site-
packages (from google-api-core<2.0.0dev,>=1.6.0->google-cloud-bigquery==1.10) (2
1.0)
Requirement already satisfied: pytz in /opt/conda/lib/python3.7/site-packages (f
rom google-api-core<2.0.0dev,>=1.6.0->google-cloud-bigquery==1.10) (2021.3)
Requirement already satisfied: google-crc32c<2.0dev,>=1.0 in /opt/conda/lib/pyth
on3.7/site-packages (from google-resumable-media>=0.3.1->google-cloud-bigquery==
1.10) (1.1.2)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /opt/conda/lib/python3.
7/site-packages (from google-auth<2.0dev,>=1.25.0->google-api-core<2.0.0dev,>=1.
6.0->google-cloud-bigquery==1.10) (0.2.7)
Requirement already satisfied: rsa<5,>=3.1.4 in /opt/conda/lib/python3.7/site-pa
ckages (from google-auth<2.0dev,>=1.25.0->google-api-core<2.0.0dev,>=1.6.0->goog
le-cloud-bigquery==1.10) (4.7.2)
Requirement already satisfied: cachetools<5.0,>=2.0.0 in /opt/conda/lib/python3.
7/site-packages (from google-auth<2.0dev,>=1.25.0->google-api-core<2.0.0dev,>=1.
6.0->google-cloud-bigquery==1.10) (4.2.4)
Requirement already satisfied: cffi>=1.0.0 in /opt/conda/lib/python3.7/site-pack
ages (from google-crc32c<2.0dev,>=1.0->google-resumable-media>=0.3.1->google-clo
ud-bigquery==1.10) (1.15.0)
Requirement already satisfied: pyparsing>=2.0.2 in /opt/conda/lib/python3.7/site
-packages (from packaging>=14.3->google-api-core<2.0.0dev,>=1.6.0->google-cloud-
bigquery==1.10) (2.4.7)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /opt/conda/lib/python3.
7/site-packages (from requests<3.0.0dev,>=2.18.0->google-api-core<2.0.0dev,>=1.
6.0->google-cloud-bigquery==1.10) (1.26.7)
Requirement already satisfied: chardet<5,>=3.0.2 in /opt/conda/lib/python3.7/sit
e-packages (from requests<3.0.0dev,>=2.18.0->google-api-core<2.0.0dev,>=1.6.0->g
oogle-cloud-bigquery==1.10) (4.0.0)
Requirement already satisfied: idna<3,>=2.5 in /opt/conda/lib/python3.7/site-pac
kages (from requests<3.0.0dev,>=2.18.0->google-api-core<2.0.0dev,>=1.6.0->google
-cloud-bigquery==1.10) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.7/si
te-packages (from requests<3.0.0dev,>=2.18.0->google-api-core<2.0.0dev,>=1.6.0->
google-cloud-bigquery==1.10) (2021.10.8)
Requirement already satisfied: pycparser in /opt/conda/lib/python3.7/site-packag
es (from cffi>=1.0.0->google-crc32c<2.0dev,>=1.0->google-resumable-media>=0.3.1-
>google-cloud-bigquery==1.10) (2.20)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /opt/conda/lib/python3.7/
site-packages (from pyasn1-modules>=0.2.1->google-auth<2.0dev,>=1.25.0->google-a
pi-core<2.0.0dev,>=1.6.0->google-cloud-bigquery==1.10) (0.4.8)
Installing collected packages: protobuf, google-auth, google-api-core, google-cl
oud-core, google-cloud-bigquery
  Attempting uninstall: protobuf
    Found existing installation: protobuf 3.18.1
    Uninstalling protobuf-3.18.1:
      Successfully uninstalled protobuf-3.18.1
  Attempting uninstall: google-auth
    Found existing installation: google-auth 2.3.0
    Uninstalling google-auth-2.3.0:
      Successfully uninstalled google-auth-2.3.0
  Attempting uninstall: google-api-core
    Found existing installation: google-api-core 2.1.0
    Uninstalling google-api-core-2.1.0:
      Successfully uninstalled google-api-core-2.1.0
  Attempting uninstall: google-cloud-core
    Found existing installation: google-cloud-core 2.1.0
```

```
    Uninstalling google-cloud-core-2.1.0:
      Successfully uninstalled google-cloud-core-2.1.0
  Attempting uninstall: google-cloud-bigquery
    Found existing installation: google-cloud-bigquery 2.28.1
    Uninstalling google-cloud-bigquery-2.28.1:
      Successfully uninstalled google-cloud-bigquery-2.28.1
ERROR: pip's dependency resolver does not currently take into account all the pa
ckages that are installed. This behaviour is the source of the following depende
ncy conflicts.
explainable-ai-sdk 1.3.2 requires xai-image-widget, which is not installed.
tfx 0.23.0 requires attrs<20,>=19.3.0, but you have attrs 21.2.0 which is incomp
atible.
tfx 0.23.0 requires click<8,>=7, but you have click 8.0.3 which is incompatible.
tfx 0.23.0 requires docker<5,>=4.1, but you have docker 5.0.3 which is incompati
ble.
tfx 0.23.0 requires google-api-python-client<2,>=1.7.8, but you have google-api-
python-client 2.27.0 which is incompatible.
tfx 0.23.0 requires google-resumable-media<0.7.0,>=0.6.0, but you have google-re
sumable-media 2.0.3 which is incompatible.
tfx 0.23.0 requires kubernetes<12,>=10.0.1, but you have kubernetes 18.20.0 whic
h is incompatible.
tfx 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyarrow 5.0.0 which is inc
ompatible.
tfx 0.23.0 requires pyyaml<6,>=3.12, but you have pyyaml 6.0 which is incompatib
le.
tfx-bsl 0.23.0 requires google-api-python-client<2,>=1.7.11, but you have google
-api-python-client 2.27.0 which is incompatible.
tfx-bsl 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyarrow 5.0.0 which is
incompatible.
tensorflow-model-analysis 0.23.0 requires pyarrow<0.18,>=0.17, but you have pyar
row 5.0.0 which is incompatible.
tensorflow-data-validation 0.23.1 requires joblib<0.15,>=0.12, but you have jobl
ib 1.1.0 which is incompatible.
tensorflow-data-validation 0.23.1 requires pyarrow<0.18,>=0.17, but you have pya
rrow 5.0.0 which is incompatible.
tensorflow-cloud 0.1.13 requires tensorboard>=2.3.0, but you have tensorboard 1.
15.0 which is incompatible.
google-cloud-translate 3.6.0 requires google-cloud-core<3.0.0dev,>=1.3.0, but yo
u have google-cloud-core 0.29.1 which is incompatible.
google-cloud-storage 1.42.3 requires google-cloud-core<3.0dev,>=1.6.0; python_ve
rsion >= "3.6", but you have google-cloud-core 0.29.1 which is incompatible.
google-cloud-spanner 3.11.1 requires google-cloud-core<3.0dev,>=1.4.1, but you h
ave google-cloud-core 0.29.1 which is incompatible.
google-cloud-logging 2.6.0 requires google-cloud-core<3.0.0dev,>=1.4.1, but you
 have google-cloud-core 0.29.1 which is incompatible.
google-cloud-firestore 2.3.4 requires google-cloud-core<3.0.0dev,>=1.4.1, but yo
u have google-cloud-core 0.29.1 which is incompatible.
google-cloud-datastore 2.3.0 requires google-cloud-core<3.0.0dev,>=1.4.0, but yo
u have google-cloud-core 0.29.1 which is incompatible.
google-cloud-bigtable 2.4.0 requires google-cloud-core<3.0.0dev,>=1.4.1, but you
have google-cloud-core 0.29.1 which is incompatible.
google-cloud-aiplatform 1.6.0 requires google-cloud-bigquery<3.0.0dev,>=1.15.0,
 but you have google-cloud-bigquery 1.10.0 which is incompatible.
apache-beam 2.33.0 requires httplib2<0.20.0,>=0.8, but you have httplib2 0.20.1
 which is incompatible.
apache-beam 2.33.0 requires pyarrow<5.0.0,>=0.15.1, but you have pyarrow 5.0.0 w
hich is incompatible.
Successfully installed google-api-core-1.31.3 google-auth-1.35.0 google-cloud-bi
gquery-1.10.0 google-cloud-core-1.7.2 protobuf-3.17.3
```

Note: Please ignore any incompatibility warnings and errors and re-run the cell

to view the installed tensorflow version.

In [1]:
```
import os
import tensorflow as tf
import numpy as np
import tensorflow_hub as hub
import shutil

PROJECT = 'qwiklabs-gcp-00-573e51f4471d' # REPLACE WITH YOUR PROJECT ID
BUCKET = 'qwiklabs-gcp-00-573e51f4471d' # REPLACE WITH YOUR BUCKET NAME
REGION = 'us-central1' # REPLACE WITH YOUR BUCKET REGION e.g. us-central1

# do not change these
os.environ['PROJECT'] = PROJECT
os.environ['BUCKET'] = BUCKET
os.environ['REGION'] = REGION
os.environ['TFVERSION'] = '1.15.3'
```

In [2]:
```
%%bash
gcloud config set project $PROJECT
gcloud config set compute/region $REGION
```

```
Updated property [core/project].
Updated property [compute/region].
```

Build the feature columns for the model.

To start, we'll load the list of categories, authors and article ids we created in the previous **Create Datasets** notebook.

In [3]:
```
categories_list = open("categories.txt").read().splitlines()
authors_list = open("authors.txt").read().splitlines()
content_ids_list = open("content_ids.txt").read().splitlines()
mean_months_since_epoch = 523
```

In the cell below we'll define the feature columns to use in our model. If necessary, remind yourself the various feature columns to use.

For the embedded_title_column feature column, use a Tensorflow Hub Module to create an embedding of the article title. Since the articles and titles are in German, you'll want to use a German language embedding module.

Explore the text embedding Tensorflow Hub modules available here. Filter by setting the language to 'German'. The 50 dimensional embedding should be sufficient for our purposes.

In [4]:
```
embedded_title_column = hub.text_embedding_column(
    key="title",
    module_spec="https://tfhub.dev/google/nnlm-de-dim50/1",
    trainable=False)

content_id_column = tf.feature_column.categorical_column_with_hash_bucket(
    key="content_id",
    hash_bucket_size= len(content_ids_list) + 1)
embedded_content_column = tf.feature_column.embedding_column(
    categorical_column=content_id_column,
    dimension=10)
```

```
author_column = tf.feature_column.categorical_column_with_hash_bucket(key="autho
    hash_bucket_size=len(authors_list) + 1)
embedded_author_column = tf.feature_column.embedding_column(
    categorical_column=author_column,
    dimension=3)

category_column_categorical = tf.feature_column.categorical_column_with_vocabula
    key="category",
    vocabulary_list=categories_list,
    num_oov_buckets=1)
category_column = tf.feature_column.indicator_column(category_column_categorical

months_since_epoch_boundaries = list(range(400,700,20))
months_since_epoch_column = tf.feature_column.numeric_column(
    key="months_since_epoch")
months_since_epoch_bucketized = tf.feature_column.bucketized_column(
    source_column = months_since_epoch_column,
    boundaries = months_since_epoch_boundaries)

crossed_months_since_category_column = tf.feature_column.indicator_column(tf.fea
  keys = [category_column_categorical, months_since_epoch_bucketized],
  hash_bucket_size = len(months_since_epoch_boundaries) * (len(categories_list)

feature_columns = [embedded_content_column,
                   embedded_author_column,
                   category_column,
                   embedded_title_column,
                   crossed_months_since_category_column]
```

Create the input function.

Next we'll create the input function for our model. This input function reads the data from the csv files we created in the previous labs.

In [5]:

```
record_defaults = [["Unknown"], ["Unknown"],["Unknown"],["Unknown"],["Unknown"],
column_keys = ["visitor_id", "content_id", "category", "title", "author", "month
label_key = "next_content_id"
def read_dataset(filename, mode, batch_size = 512):
  def _input_fn():
      def decode_csv(value_column):
          columns = tf.decode_csv(value_column,record_defaults=record_defaults)
          features = dict(zip(column_keys, columns))
          label = features.pop(label_key)
          return features, label

      # Create list of files that match pattern
      file_list = tf.io.gfile.glob(filename)

      # Create dataset from file list
      dataset = tf.data.TextLineDataset(file_list).map(decode_csv)

      if mode == tf.estimator.ModeKeys.TRAIN:
          num_epochs = None # indefinitely
          dataset = dataset.shuffle(buffer_size = 10 * batch_size)
      else:
          num_epochs = 1 # end-of-input after this
```

```
    dataset = dataset.repeat(num_epochs).batch(batch_size)
    return dataset.make_one_shot_iterator().get_next()
  return _input_fn
```

Create the model and train/evaluate

Next, we'll build our model which recommends an article for a visitor to the Kurier.at website. Look through the code below. We use the input_layer feature column to create the dense input layer to our network. This is just a single layer network where we can adjust the number of hidden units as a parameter.

Currently, we compute the accuracy between our predicted 'next article' and the actual 'next article' read next by the visitor. We'll also add an additional performance metric of top 10 accuracy to assess our model. To accomplish this, we compute the top 10 accuracy metric, add it to the metrics dictionary below and add it to the tf.summary so that this value is reported to Tensorboard as well.

In [6]:

```
def model_fn(features, labels, mode, params):
  net = tf.feature_column.input_layer(features, params['feature_columns'])
  for units in params['hidden_units']:
        net = tf.layers.dense(net, units=units, activation=tf.nn.relu)
   # Compute logits (1 per class).
  logits = tf.layers.dense(net, params['n_classes'], activation=None)

  predicted_classes = tf.argmax(logits, 1)
  from tensorflow.python.lib.io import file_io

  with file_io.FileIO('content_ids.txt', mode='r') as ifp:
    content = tf.constant([x.rstrip() for x in ifp])
  predicted_class_names = tf.gather(content, predicted_classes)
  if mode == tf.estimator.ModeKeys.PREDICT:
    predictions = {
        'class_ids': predicted_classes[:, tf.newaxis],
        'class_names' : predicted_class_names[:, tf.newaxis],
        'probabilities': tf.nn.softmax(logits),
        'logits': logits,
    }
    return tf.estimator.EstimatorSpec(mode, predictions=predictions)
  table = tf.contrib.lookup.index_table_from_file(vocabulary_file="content_ids.t
  labels = table.lookup(labels)
  # Compute loss.
  loss = tf.losses.sparse_softmax_cross_entropy(labels=labels, logits=logits)

  # Compute evaluation metrics.
  accuracy = tf.metrics.accuracy(labels=labels,
                                 predictions=predicted_classes,
                                 name='acc_op')
  top_10_accuracy = tf.metrics.mean(tf.nn.in_top_k(predictions=logits,
                                                   targets=labels,
                                                   k=10))

  metrics = {
    'accuracy': accuracy,
    'top_10_accuracy' : top_10_accuracy}

  tf.summary.scalar('accuracy', accuracy[1])
```

```
  tf.summary.scalar('top_10_accuracy', top_10_accuracy[1])

  if mode == tf.estimator.ModeKeys.EVAL:
      return tf.estimator.EstimatorSpec(
          mode, loss=loss, eval_metric_ops=metrics)

  # Create training op.
  assert mode == tf.estimator.ModeKeys.TRAIN

  optimizer = tf.train.AdagradOptimizer(learning_rate=0.1)
  train_op = optimizer.minimize(loss, global_step=tf.train.get_global_step())
  return tf.estimator.EstimatorSpec(mode, loss=loss, train_op=train_op)
```

Train and Evaluate

In [7]:

```
outdir = 'content_based_model_trained'
shutil.rmtree(outdir, ignore_errors = True) # start fresh each time
#tf.summary.FileWriterCache.clear() # ensure filewriter cache is clear for Tenso
estimator = tf.estimator.Estimator(
    model_fn=model_fn,
    model_dir = outdir,
    params={
     'feature_columns': feature_columns,
      'hidden_units': [200, 100, 50],
      'n_classes': len(content_ids_list)
    })

train_spec = tf.estimator.TrainSpec(
    input_fn = read_dataset("training_set.csv", tf.estimator.ModeKeys.TRAIN),
    max_steps = 2000)

eval_spec = tf.estimator.EvalSpec(
    input_fn = read_dataset("test_set.csv", tf.estimator.ModeKeys.EVAL),
    steps = None,
    start_delay_secs = 30,
    throttle_secs = 60)

tf.estimator.train_and_evaluate(estimator, train_spec, eval_spec)
```

```
INFO:tensorflow:Using default config.
INFO:tensorflow:Using default config.
INFO:tensorflow:Using config: {'_model_dir': 'content_based_model_trained', '_tf
_random_seed': None, '_save_summary_steps': 100, '_save_checkpoints_steps': Non
e, '_save_checkpoints_secs': 600, '_session_config': allow_soft_placement: true
graph_options {
  rewrite_options {
    meta_optimizer_iterations: ONE
  }
}
, '_keep_checkpoint_max': 5, '_keep_checkpoint_every_n_hours': 10000, '_log_step
_count_steps': 100, '_train_distribute': None, '_device_fn': None, '_protocol':
None, '_eval_distribute': None, '_experimental_distribute': None, '_experimental
_max_worker_delay_secs': None, '_session_creation_timeout_secs': 7200, '_servic
e': None, '_cluster_spec': <tensorflow.python.training.server_lib.ClusterSpec ob
ject at 0x7f837d9cbe90>, '_task_type': 'worker', '_task_id': 0, '_global_id_in_c
luster': 0, '_master': '', '_evaluation_master': '', '_is_chief': True, '_num_ps
_replicas': 0, '_num_worker_replicas': 1}
INFO:tensorflow:Using config: {'_model_dir': 'content_based_model_trained', '_tf
```

```
_random_seed': None, '_save_summary_steps': 100, '_save_checkpoints_steps': Non
e, '_save_checkpoints_secs': 600, '_session_config': allow_soft_placement: true
graph_options {
  rewrite_options {
    meta_optimizer_iterations: ONE
  }
}
, '_keep_checkpoint_max': 5, '_keep_checkpoint_every_n_hours': 10000, '_log_step
_count_steps': 100, '_train_distribute': None, '_device_fn': None, '_protocol':
None, '_eval_distribute': None, '_experimental_distribute': None, '_experimental
_max_worker_delay_secs': None, '_session_creation_timeout_secs': 7200, '_servic
e': None, '_cluster_spec': <tensorflow.python.training.server_lib.ClusterSpec ob
ject at 0x7f837d9cbe90>, '_task_type': 'worker', '_task_id': 0, '_global_id_in_c
luster': 0, '_master': '', '_evaluation_master': '', '_is_chief': True, '_num_ps
_replicas': 0, '_num_worker_replicas': 1}
INFO:tensorflow:Not using Distribute Coordinator.

INFO:tensorflow:Not using Distribute Coordinator.
INFO:tensorflow:Running training and evaluation locally (non-distributed).

INFO:tensorflow:Running training and evaluation locally (non-distributed).
INFO:tensorflow:Start train and evaluate loop. The evaluate will happen after ev
ery checkpoint. Checkpoint frequency is determined based on RunConfig arguments:
save_checkpoints_steps None or save_checkpoints_secs 600.

INFO:tensorflow:Start train and evaluate loop. The evaluate will happen after ev
ery checkpoint. Checkpoint frequency is determined based on RunConfig arguments:
save_checkpoints_steps None or save_checkpoints_secs 600.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/training/training_util.py:236: Variable.initialized_value (from tensorflo
w.python.ops.variables) is deprecated and will be removed in a future version.
Instructions for updating:
Use Variable.read_value. Variables in 2.X are initialized automatically both in
eager and graph (inside tf.defun) contexts.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/training/training_util.py:236: Variable.initialized_value (from tensorflo
w.python.ops.variables) is deprecated and will be removed in a future version.
Instructions for updating:
Use Variable.read_value. Variables in 2.X are initialized automatically both in
eager and graph (inside tf.defun) contexts.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/autograph/converters/directives.py:119: The name tf.decode_csv is deprecat
ed. Please use tf.io.decode_csv instead.


WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/autograph/converters/directives.py:119: The name tf.decode_csv is deprecat
ed. Please use tf.io.decode_csv instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/1111986337.py:25: DatasetV1.make_on
e_shot_iterator (from tensorflow.python.data.ops.dataset_ops) is deprecated and
will be removed in a future version.
Instructions for updating:
Use `for ... in dataset:` to iterate over a dataset. If using `tf.estimator`, re
turn the `Dataset` object directly from your input function. As a last resort, y
ou can use `tf.compat.v1.data.make_one_shot_iterator(dataset)`.

WARNING:tensorflow:From /tmp/ipykernel_25835/1111986337.py:25: DatasetV1.make_on
e_shot_iterator (from tensorflow.python.data.ops.dataset_ops) is deprecated and
will be removed in a future version.
Instructions for updating:
Use `for ... in dataset:` to iterate over a dataset. If using `tf.estimator`, re
turn the `Dataset` object directly from your input function. As a last resort, y
ou can use `tf.compat.v1.data.make_one_shot_iterator(dataset)`.
```

```
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Calling model_fn.
WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:2: The name tf.featur
e_column.input_layer is deprecated. Please use tf.compat.v1.feature_column.input
_layer instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:2: The name tf.featur
e_column.input_layer is deprecated. Please use tf.compat.v1.feature_column.input
_layer instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:206: EmbeddingColumn._get_dense_tensor (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:206: EmbeddingColumn._get_dense_tensor (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:3182: HashedCategoricalColumn._get_spa
rse_tensors (from tensorflow.python.feature_column.feature_column_v2) is depreca
ted and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:3182: HashedCategoricalColumn._get_spa
rse_tensors (from tensorflow.python.feature_column.feature_column_v2) is depreca
ted and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: HashedCategoricalColumn._transform_
feature (from tensorflow.python.feature_column.feature_column_v2) is deprecated
and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: HashedCategoricalColumn._transform_
feature (from tensorflow.python.feature_column.feature_column_v2) is deprecated
and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:3122: HashedCategoricalColumn._num_buc
kets (from tensorflow.python.feature_column.feature_column_v2) is deprecated and
will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
```

```
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:3122: HashedCategoricalColumn._num_buc
kets (from tensorflow.python.feature_column.feature_column_v2) is deprecated and
will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/ops/embedding_ops.py:802: where (from tensorflow.python.ops.array_ops) is
deprecated and will be removed in a future version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/ops/embedding_ops.py:802: where (from tensorflow.python.ops.array_ops) is
deprecated and will be removed in a future version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:207: EmbeddingColumn._variable_shape (fro
m tensorflow.python.feature_column.feature_column_v2) is deprecated and will be
removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:207: EmbeddingColumn._variable_shape (fro
m tensorflow.python.feature_column.feature_column_v2) is deprecated and will be
removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:206: IndicatorColumn._get_dense_tensor (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:206: IndicatorColumn._get_dense_tensor (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: IndicatorColumn._transform_feature
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: IndicatorColumn._transform_feature
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
```

```
ython/feature_column/feature_column_v2.py:4300: CrossedColumn._get_sparse_tensor
s (from tensorflow.python.feature_column.feature_column_v2) is deprecated and wi
ll be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4300: CrossedColumn._get_sparse_tensor
s (from tensorflow.python.feature_column.feature_column_v2) is deprecated and wi
ll be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: CrossedColumn._transform_feature (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: CrossedColumn._transform_feature (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4118: VocabularyListCategoricalColumn.
_get_sparse_tensors (from tensorflow.python.feature_column.feature_column_v2) is
deprecated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4118: VocabularyListCategoricalColumn.
_get_sparse_tensors (from tensorflow.python.feature_column.feature_column_v2) is
deprecated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: VocabularyListCategoricalColumn._tr
ansform_feature (from tensorflow.python.feature_column.feature_column_v2) is dep
recated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: VocabularyListCategoricalColumn._tr
ansform_feature (from tensorflow.python.feature_column.feature_column_v2) is dep
recated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4118: BucketizedColumn._get_sparse_ten
sors (from tensorflow.python.feature_column.feature_column_v2) is deprecated and
will be removed in a future version.
Instructions for updating:
```

```
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4118: BucketizedColumn._get_sparse_ten
sors (from tensorflow.python.feature_column.feature_column_v2) is deprecated and
will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: BucketizedColumn._transform_feature
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: BucketizedColumn._transform_feature
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: NumericColumn._transform_feature (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column.py:2158: NumericColumn._transform_feature (f
rom tensorflow.python.feature_column.feature_column_v2) is deprecated and will b
e removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4271: IndicatorColumn._variable_shape
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4271: IndicatorColumn._variable_shape
(from tensorflow.python.feature_column.feature_column_v2) is deprecated and will
be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4326: CrossedColumn._num_buckets (from
tensorflow.python.feature_column.feature_column_v2) is deprecated and will be re
moved in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
```

```
ython/feature_column/feature_column_v2.py:4326: CrossedColumn._num_buckets (from
tensorflow.python.feature_column.feature_column_v2) is deprecated and will be re
moved in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4326: VocabularyListCategoricalColumn.
_num_buckets (from tensorflow.python.feature_column.feature_column_v2) is deprec
ated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/feature_column/feature_column_v2.py:4326: VocabularyListCategoricalColumn.
_num_buckets (from tensorflow.python.feature_column.feature_column_v2) is deprec
ated and will be removed in a future version.
Instructions for updating:
The old _FeatureColumn APIs are being deprecated. Please use the new FeatureColu
mn APIs instead.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore

2021-11-04 19:33:06.411123: W tensorflow/core/graph/graph_constructor.cc:1491] I
mporting a graph with a lower producer version 26 into an existing graph with pr
oducer version 134. Shape inference will have run different parts of the graph w
ith different producer versions.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore

INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:4: dense (from tensor
flow.python.layers.core) is deprecated and will be removed in a future version.
Instructions for updating:
Use keras.layers.Dense instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:4: dense (from tensor
flow.python.layers.core) is deprecated and will be removed in a future version.
Instructions for updating:
Use keras.layers.Dense instead.
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/layers/core.py:187: Layer.apply (from tensorflow.python.keras.engine.base_
layer) is deprecated and will be removed in a future version.
Instructions for updating:
Please use `layer.__call__` method instead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/layers/core.py:187: Layer.apply (from tensorflow.python.keras.engine.base_
layer) is deprecated and will be removed in a future version.
Instructions for updating:
Please use `layer.__call__` method instead.
WARNING:tensorflow:
The TensorFlow contrib module will not be included in TensorFlow 2.0.
For more information, please see:
  * https://github.com/tensorflow/community/blob/master/rfcs/20180907-contrib-su
nset.md
  * https://github.com/tensorflow/addons
  * https://github.com/tensorflow/io (for I/O related ops)
If you depend on functionality not listed there, please file an issue.
```

```
WARNING:tensorflow:
The TensorFlow contrib module will not be included in TensorFlow 2.0.
For more information, please see:
  * https://github.com/tensorflow/community/blob/master/rfcs/20180907-contrib-su
nset.md
  * https://github.com/tensorflow/addons
  * https://github.com/tensorflow/io (for I/O related ops)
If you depend on functionality not listed there, please file an issue.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:25: The name tf.losse
s.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.spa
rse_softmax_cross_entropy instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:25: The name tf.losse
s.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.spa
rse_softmax_cross_entropy instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:28: The name tf.metri
cs.accuracy is deprecated. Please use tf.compat.v1.metrics.accuracy instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:28: The name tf.metri
cs.accuracy is deprecated. Please use tf.compat.v1.metrics.accuracy instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:31: The name tf.metri
cs.mean is deprecated. Please use tf.compat.v1.metrics.mean instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:31: The name tf.metri
cs.mean is deprecated. Please use tf.compat.v1.metrics.mean instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:39: The name tf.summa
ry.scalar is deprecated. Please use tf.compat.v1.summary.scalar instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:39: The name tf.summa
ry.scalar is deprecated. Please use tf.compat.v1.summary.scalar instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:49: The name tf.trai
n.AdagradOptimizer is deprecated. Please use tf.compat.v1.train.AdagradOptimizer
instead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:49: The name tf.trai
n.AdagradOptimizer is deprecated. Please use tf.compat.v1.train.AdagradOptimizer
instead.

WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:50: The name tf.trai
n.get_global_step is deprecated. Please use tf.compat.v1.train.get_global_step i
nstead.


WARNING:tensorflow:From /tmp/ipykernel_25835/2314360427.py:50: The name tf.trai
n.get_global_step is deprecated. Please use tf.compat.v1.train.get_global_step i
nstead.

WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/training/adagrad.py:76: calling Constant.__init__ (from tensorflow.python.
ops.init_ops) with dtype is deprecated and will be removed in a future version.
Instructions for updating:
Call initializer instance with the dtype argument instead of passing it to the c
onstructor
WARNING:tensorflow:From /opt/conda/lib/python3.7/site-packages/tensorflow_core/p
ython/training/adagrad.py:76: calling Constant.__init__ (from tensorflow.python.
```

```
ops.init_ops) with dtype is deprecated and will be removed in a future version.
Instructions for updating:
Call initializer instance with the dtype argument instead of passing it to the c
onstructor
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Create CheckpointSaverHook.
INFO:tensorflow:Create CheckpointSaverHook.
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Graph was finalized.
2021-11-04 19:33:07.118562: I tensorflow/core/platform/cpu_feature_guard.cc:142]
Your CPU supports instructions that this TensorFlow binary was not compiled to u
se: AVX2 FMA
2021-11-04 19:33:07.124231: I tensorflow/core/platform/profile_utils/cpu_utils.c
c:94] CPU Frequency: 2200155000 Hz
2021-11-04 19:33:07.124498: I tensorflow/compiler/xla/service/service.cc:168] XL
A service 0x56110dc4acc0 initialized for platform Host (this does not guarantee
that XLA will be used). Devices:
2021-11-04 19:33:07.124526: I tensorflow/compiler/xla/service/service.cc:176]
StreamExecutor device (0): Host, Default Version
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Saving checkpoints for 0 into content_based_model_trained/model.
ckpt.
INFO:tensorflow:Saving checkpoints for 0 into content_based_model_trained/model.
ckpt.
INFO:tensorflow:loss = 9.656642, step = 1
INFO:tensorflow:loss = 9.656642, step = 1
INFO:tensorflow:global_step/sec: 8.46066
INFO:tensorflow:global_step/sec: 8.46066
INFO:tensorflow:loss = 5.8901443, step = 101 (11.822 sec)
INFO:tensorflow:loss = 5.8901443, step = 101 (11.822 sec)
INFO:tensorflow:global_step/sec: 8.44057
INFO:tensorflow:global_step/sec: 8.44057
INFO:tensorflow:loss = 4.7733927, step = 201 (11.848 sec)
INFO:tensorflow:loss = 4.7733927, step = 201 (11.848 sec)
INFO:tensorflow:global_step/sec: 8.51153
INFO:tensorflow:global_step/sec: 8.51153
INFO:tensorflow:loss = 4.9673395, step = 301 (11.750 sec)
INFO:tensorflow:loss = 4.9673395, step = 301 (11.750 sec)
INFO:tensorflow:global_step/sec: 8.73552
INFO:tensorflow:global_step/sec: 8.73552
INFO:tensorflow:loss = 4.505791, step = 401 (11.446 sec)
INFO:tensorflow:loss = 4.505791, step = 401 (11.446 sec)
INFO:tensorflow:global_step/sec: 8.5173
INFO:tensorflow:global_step/sec: 8.5173
INFO:tensorflow:loss = 5.727135, step = 501 (11.741 sec)
INFO:tensorflow:loss = 5.727135, step = 501 (11.741 sec)
INFO:tensorflow:global_step/sec: 8.39534
INFO:tensorflow:global_step/sec: 8.39534
INFO:tensorflow:loss = 5.5824623, step = 601 (11.912 sec)
INFO:tensorflow:loss = 5.5824623, step = 601 (11.912 sec)
INFO:tensorflow:global_step/sec: 8.51405
```

```
INFO:tensorflow:global_step/sec: 8.51405
INFO:tensorflow:loss = 4.8128004, step = 701 (11.745 sec)
INFO:tensorflow:loss = 4.8128004, step = 701 (11.745 sec)
INFO:tensorflow:global_step/sec: 8.51511
INFO:tensorflow:global_step/sec: 8.51511
INFO:tensorflow:loss = 4.6534204, step = 801 (11.744 sec)
INFO:tensorflow:loss = 4.6534204, step = 801 (11.744 sec)
INFO:tensorflow:global_step/sec: 8.34677
INFO:tensorflow:global_step/sec: 8.34677
INFO:tensorflow:loss = 4.316224, step = 901 (11.980 sec)
INFO:tensorflow:loss = 4.316224, step = 901 (11.980 sec)
INFO:tensorflow:global_step/sec: 8.16986
INFO:tensorflow:global_step/sec: 8.16986
INFO:tensorflow:loss = 5.617612, step = 1001 (12.240 sec)
INFO:tensorflow:loss = 5.617612, step = 1001 (12.240 sec)
INFO:tensorflow:global_step/sec: 8.41208
INFO:tensorflow:global_step/sec: 8.41208
INFO:tensorflow:loss = 4.5326986, step = 1101 (11.888 sec)
INFO:tensorflow:loss = 4.5326986, step = 1101 (11.888 sec)
INFO:tensorflow:global_step/sec: 8.54456
INFO:tensorflow:global_step/sec: 8.54456
INFO:tensorflow:loss = 4.989264, step = 1201 (11.703 sec)
INFO:tensorflow:loss = 4.989264, step = 1201 (11.703 sec)
INFO:tensorflow:global_step/sec: 8.55272
INFO:tensorflow:global_step/sec: 8.55272
INFO:tensorflow:loss = 4.640286, step = 1301 (11.692 sec)
INFO:tensorflow:loss = 4.640286, step = 1301 (11.692 sec)
INFO:tensorflow:global_step/sec: 8.32734
INFO:tensorflow:global_step/sec: 8.32734
INFO:tensorflow:loss = 5.4034114, step = 1401 (12.009 sec)
INFO:tensorflow:loss = 5.4034114, step = 1401 (12.009 sec)
INFO:tensorflow:global_step/sec: 8.38427
INFO:tensorflow:global_step/sec: 8.38427
INFO:tensorflow:loss = 5.460411, step = 1501 (11.928 sec)
INFO:tensorflow:loss = 5.460411, step = 1501 (11.928 sec)
INFO:tensorflow:global_step/sec: 8.48495
INFO:tensorflow:global_step/sec: 8.48495
INFO:tensorflow:loss = 4.6751447, step = 1601 (11.785 sec)
INFO:tensorflow:loss = 4.6751447, step = 1601 (11.785 sec)
INFO:tensorflow:global_step/sec: 8.4357
INFO:tensorflow:global_step/sec: 8.4357
INFO:tensorflow:loss = 4.489712, step = 1701 (11.857 sec)
INFO:tensorflow:loss = 4.489712, step = 1701 (11.857 sec)
INFO:tensorflow:global_step/sec: 8.5491
INFO:tensorflow:global_step/sec: 8.5491
INFO:tensorflow:loss = 4.2188635, step = 1801 (11.694 sec)
INFO:tensorflow:loss = 4.2188635, step = 1801 (11.694 sec)
INFO:tensorflow:global_step/sec: 8.57636
INFO:tensorflow:global_step/sec: 8.57636
INFO:tensorflow:loss = 5.5490894, step = 1901 (11.661 sec)
INFO:tensorflow:loss = 5.5490894, step = 1901 (11.661 sec)
INFO:tensorflow:Saving checkpoints for 2000 into content_based_model_trained/mod
el.ckpt.
```

```
INFO:tensorflow:Saving checkpoints for 2000 into content_based_model_trained/mod
el.ckpt.
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
2021-11-04 19:37:07.551733: W tensorflow/core/graph/graph_constructor.cc:1491] I
mporting a graph with a lower producer version 26 into an existing graph with pr
oducer version 134. Shape inference will have run different parts of the graph w
ith different producer versions.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Starting evaluation at 2021-11-04T19:37:07Z
INFO:tensorflow:Starting evaluation at 2021-11-04T19:37:07Z
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Restoring parameters from content_based_model_trained/model.ckpt
-2000
INFO:tensorflow:Restoring parameters from content_based_model_trained/model.ckpt
-2000
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Finished evaluation at 2021-11-04-19:37:13
INFO:tensorflow:Finished evaluation at 2021-11-04-19:37:13
INFO:tensorflow:Saving dict for global step 2000: accuracy = 0.03976718, global_
step = 2000, loss = 5.1095486, top_10_accuracy = 0.25395525
INFO:tensorflow:Saving dict for global step 2000: accuracy = 0.03976718, global_
step = 2000, loss = 5.1095486, top_10_accuracy = 0.25395525
INFO:tensorflow:Saving 'checkpoint_path' summary for global step 2000: content_b
ased_model_trained/model.ckpt-2000
INFO:tensorflow:Saving 'checkpoint_path' summary for global step 2000: content_b
ased_model_trained/model.ckpt-2000
INFO:tensorflow:Loss for final step: 4.588237.
INFO:tensorflow:Loss for final step: 4.588237.
```

Out[7]:

```
({'accuracy': 0.03976718,
  'loss': 5.1095486,
  'top_10_accuracy': 0.25395525,
  'global_step': 2000},
 [])
```

This takes a while to complete but in the end, I get about **30% top 10 accuracy**.

Make predictions with the trained model.

With the model now trained, we can make predictions by calling the predict method on the estimator. Let's look at how our model predicts on the first five examples of the training set.
To start, we'll create a new file 'first_5.csv' which contains the first five elements of our training

set. We'll also save the target values to a file 'first_5_content_ids' so we can compare our results.

In [8]:
```
%%bash
head -5 training_set.csv > first_5.csv
head first_5.csv
awk -F "\"*,\"*" '{print $2}' first_5.csv > first_5_content_ids
```

```
1038643850985118087,299828023,News,Glyphosat geht in die Verlängerung,Andreas An
zenberger,574,299425707
1083265653486482344,299828023,News,Glyphosat geht in die Verlängerung,Andreas An
zenberger,574,299957318
1090462532616257705,299696307,News,Patient mit Taxi heimgeschickt,Johannes Weich
hart,574,299793337
1090462532616257705,299793337,News,""Sittenwächter"" an See in Niederösterreic
h: Hauptverdächtiger in Haft",,574,299836255
1122220186524713655,299915364,News,Glyphosat-Verlängerung: SPD wirft CDU Vertrau
ensbruch vor,,574,299949290
```

Recall, to make predictions on the trained model we pass a list of examples through the input function. Complete the code below to make predictions on the examples contained in the "first_5.csv" file we created above.

In [9]:
```
output = list(estimator.predict(input_fn=read_dataset("first_5.csv", tf.estimato
```

```
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
2021-11-04 19:37:14.161038: W tensorflow/core/graph/graph_constructor.cc:1491] I
mporting a graph with a lower producer version 26 into an existing graph with pr
oducer version 134. Shape inference will have run different parts of the graph w
ith different producer versions.
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Saver not created because there are no variables in the graph to
restore
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Restoring parameters from content_based_model_trained/model.ckpt
-2000
INFO:tensorflow:Restoring parameters from content_based_model_trained/model.ckpt
-2000
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Done running local_init_op.
```

In [10]:
```
import numpy as np
recommended_content_ids = [np.asscalar(d["class_names"]).decode('UTF-8') for d i
content_ids = open("first_5_content_ids").read().splitlines()
```

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:2: DeprecationWarni
ng: np.asscalar(a) is deprecated since NumPy v1.16, use a.item() instead
```

Finally, we map the content id back to the article title. Let's compare our model's recommendation for the first example. This can be done in BigQuery. Look through the query below and make sure it is clear what is being returned.

In [11]:

```
from google.cloud import bigquery
recommended_title_sql="""
#standardSQL
SELECT
(SELECT MAX(IF(index=6, value, NULL)) FROM UNNEST(hits.customDimensions)) AS tit
FROM `cloud-training-demos.GA360_test.ga_sessions_sample`,
  UNNEST(hits) AS hits
WHERE
  # only include hits on pages
  hits.type = "PAGE"
  AND (SELECT MAX(IF(index=10, value, NULL)) FROM UNNEST(hits.customDimensions))
LIMIT 1""".format(recommended_content_ids[0])

current_title_sql="""
#standardSQL
SELECT
(SELECT MAX(IF(index=6, value, NULL)) FROM UNNEST(hits.customDimensions)) AS tit
FROM `cloud-training-demos.GA360_test.ga_sessions_sample`,
  UNNEST(hits) AS hits
WHERE
  # only include hits on pages
  hits.type = "PAGE"
  AND (SELECT MAX(IF(index=10, value, NULL)) FROM UNNEST(hits.customDimensions))
LIMIT 1""".format(content_ids[0])
recommended_title = bigquery.Client().query(recommended_title_sql).to_dataframe(
current_title = bigquery.Client().query(current_title_sql).to_dataframe()['title
print("Current title: {} ".format(current_title))
print("Recommended title: {}".format(recommended_title))
```

```
Current title: b'Glyphosat geht in die Verl\xc3\xa4ngerung'
Recommended title: b'Pammesberger 2017: Die Karikaturen aus dem KURIER'
```

In [ ]: