```
import pandas as pd
import datetime
import numpy as np
import random
import re
import scipy
import collections
import matplotlib.pyplot as plt
# %matplotlib inline

from sklearn.model_selection import train_test_split
from sklearn.ensemble import GradientBoostingClassifier
```

Подход к решению задачи по прогнозированию покупок беременных. Тестовое задание к х
Бочаров А.М. скайп bam271074

Прогнозирование детских покупок можно сделать на основании покупок, которые женщин
Известно, что в период беременности женщина нуждается в витаминах и фолиевой кислот
возраста от 16 до 40 лет часто покупает овощи, фрукты, овес, постное мясо, молочные изд
информации можно построить модель, обучить ее выявлять данную закономерность и исг

```
pip list
```

⊡

```
Package                 Version
----------------------- ---------------
absl-py                 0.9.0
alabaster               0.7.12
albumentations          0.1.12
altair                  4.1.0
asgiref                 3.2.7
astor                   0.8.1
astropy                 4.0.1.post1
astunparse              1.6.3
atari-py                0.2.6
atomicwrites            1.4.0
attrs                   19.3.0
audioread               2.1.8
autograd                1.3
Babel                   2.8.0
backcall                0.1.0
beautifulsoup4          4.6.3
bleach                  3.1.5
blis                    0.4.1
bokeh                   1.4.0
boto                    2.49.0
boto3                   1.13.13
botocore                1.16.13
Bottleneck              1.3.2
branca                  0.4.1
bs4                     0.0.1
CacheControl            0.12.6
cachetools              3.1.1
catalogue               1.0.0
certifi                 2020.4.5.1
cffi                    1.14.0
chainer                 6.5.0
chardet                 3.0.4
click                   7.1.2
cloudpickle             1.3.0
cmake                   3.12.0
cmdstanpy               0.4.0
colorlover              0.3.0
community               1.0.0b1
contextlib2             0.5.5
convertdate             2.2.1
coverage                3.7.1
coveralls               0.5
crcmod                  1.7
cufflinks               0.17.3
cupy-cuda101            6.5.0
cvxopt                  1.2.5
cvxpy                   1.0.31
cycler                  0.10.0
cymem                   2.0.3
Cython                  0.29.18
daft                    0.0.4
dask                    2.12.0
dataclasses             0.7
datascience             0.10.6
decorator               4.4.2
defusedxml              0.6.0
descartes               1.1.0
dill                    0.3.1.1
distributed             1.25.3
```

```
Django                     3.0.6
dlib                       19.18.0
docopt                     0.6.2
docutils                   0.15.2
dopamine-rl                1.0.5
earthengine-api            0.1.221
easydict                   1.9
ecos                       2.0.7.post1
editdistance               0.5.3
en-core-web-sm             2.2.5
entrypoints                0.3
ephem                      3.7.7.1
et-xmlfile                 1.0.1
fa2                        0.3.5
fancyimpute                0.4.3
fastai                     1.0.61
fastdtw                    0.3.4
fastprogress               0.2.3
fastrlock                  0.4
fbprophet                  0.6
feather-format             0.4.1
featuretools               0.4.1
filelock                   3.0.12
firebase-admin             4.1.0
fix-yahoo-finance          0.0.22
Flask                      1.1.2
folium                     0.8.3
fsspec                     0.7.4
future                     0.16.0
gast                       0.3.3
GDAL                       2.2.2
gdown                      3.6.4
gensim                     3.6.0
geographiclib              1.50
geopy                      1.17.0
gin-config                 0.3.0
glob2                      0.7
google                     2.0.3
google-api-core            1.16.0
google-api-python-client   1.7.12
google-auth                1.7.2
google-auth-httplib2       0.0.3
google-auth-oauthlib       0.4.1
google-cloud-bigquery      1.21.0
google-cloud-core          1.0.3
google-cloud-datastore     1.8.0
google-cloud-firestore     1.7.0
google-cloud-language      1.2.0
google-cloud-storage       1.18.1
google-cloud-translate     1.5.0
google-colab               1.0.0
google-pasta               0.2.0
google-resumable-media     0.4.1
googleapis-common-protos   1.51.0
googledrivedownloader      0.4
graphviz                   0.10.1
grpcio                     1.29.0
gspread                    3.0.1
gspread-dataframe          3.0.6
gym                        0.17.2
h5py                       2.10.0
HeapDict                   1.0.1
```

```
holidays              0.9.12
html5lib              1.0.1
httpimport            0.5.18
httplib2              0.17.3
httplib2shim          0.0.3
humanize              0.5.1
hyperopt              0.1.2
ideep4py              2.0.0.post3
idna                  2.9
image                 1.5.31
imageio               2.4.1
imagesize             1.2.0
imbalanced-learn      0.4.3
imblearn              0.0
imgaug                0.2.9
importlib-metadata    1.6.0
imutils               0.5.3
inflect               2.1.0
intel-openmp          2020.0.133
intervaltree          2.1.0
ipykernel             4.10.1
ipython               5.5.0
ipython-genutils      0.2.0
ipython-sql           0.3.9
ipywidgets            7.5.1
itsdangerous          1.1.0
jax                   0.1.67
jaxlib                0.1.47
jdcal                 1.4.1
jedi                  0.17.0
jieba                 0.42.1
Jinja2                2.11.2
jmespath              0.10.0
joblib                0.15.1
jpeg4py               0.1.4
jsonschema            2.6.0
jupyter               1.0.0
jupyter-client        5.3.4
jupyter-console       5.2.0
jupyter-core          4.6.3
kaggle                1.5.6
kapre                 0.1.3.1
Keras                 2.3.1
Keras-Applications    1.0.8
Keras-Preprocessing   1.1.2
keras-vis             0.4.1
kiwisolver            1.2.0
knnimpute             0.1.0
librosa               0.6.3
lightgbm              2.2.3
llvmlite              0.31.0
lmdb                  0.98
lucid                 0.3.8
LunarCalendar         0.0.9
lxml                  4.2.6
Markdown              3.2.2
MarkupSafe            1.1.1
matplotlib            3.2.1
matplotlib-venn       0.11.5
missingno             0.4.2
mistune               0.8.4
```

```
mizani                0.6.0
mkl                   2019.0
mlxtend               0.14.0
more-itertools        8.3.0
moviepy               0.2.3.5
mpmath                1.1.0
msgpack               1.0.0
multiprocess          0.70.9
multitasking          0.0.9
murmurhash            1.0.2
music21               5.5.0
natsort               5.5.0
nbconvert             5.6.1
nbformat              5.0.6
networkx              2.4
nibabel               3.0.2
nltk                  3.2.5
notebook              5.2.2
np-utils              0.5.12.1
numba                 0.48.0
numexpr               2.7.1
numpy                 1.18.4
nvidia-ml-py3         7.352.0
oauth2client          4.1.3
oauthlib              3.1.0
okgrade               0.4.3
opencv-contrib-python 4.1.2.30
opencv-python         4.1.2.30
openpyxl              2.5.9
opt-einsum            3.2.1
osqp                  0.6.1
packaging             20.4
palettable            3.3.0
pandas                1.0.3
pandas-datareader     0.8.1
pandas-gbq            0.11.0
pandas-profiling      1.4.1
pandocfilters         1.4.2
parso                 0.7.0
pathlib               1.0.1
patsy                 0.5.1
pexpect               4.8.0
pickleshare           0.7.5
Pillow                7.0.0
pip                   19.3.1
pip-tools             4.5.1
plac                  1.1.3
plotly                4.4.1
plotnine              0.6.0
pluggy                0.7.1
portpicker            1.3.1
prefetch-generator    1.0.1
preshed               3.0.2
prettytable           0.7.2
progressbar2          3.38.0
prometheus-client     0.8.0
promise               2.3
prompt-toolkit        1.0.18
protobuf              3.10.0
psutil                5.4.8
psycopg2              2.7.6.1
ptvsd                 5.0.0a12
```

```
ptyprocess            0.6.0
py                    1.8.1
pyarrow               0.14.1
pyasn1                0.4.8
pyasn1-modules        0.2.8
pycocotools           2.0.0
pycparser             2.20
pydata-google-auth    1.1.0
pydot                 1.3.0
pydot-ng              2.0.0
pydotplus             2.0.2
PyDrive               1.3.1
pyemd                 0.5.1
pyglet                1.5.0
Pygments              2.1.3
pygobject             3.26.1
pymc3                 3.7
PyMeeus               0.3.7
pymongo               3.10.1
pymystem3             0.2.0
PyOpenGL              3.1.5
pyparsing             2.4.7
pyrsistent            0.16.0
pysndfile             1.3.8
PySocks               1.7.1
pystan                2.19.1.1
pytest                3.6.4
python-apt            1.6.5+ubuntu0.2
python-chess          0.23.11
python-dateutil       2.8.1
python-louvain        0.14
python-slugify        4.0.0
python-utils          2.4.0
pytz                  2018.9
PyWavelets            1.1.1
PyYAML                3.13
pyzmq                 19.0.1
qtconsole             4.7.4
QtPy                  1.9.0
regex                 2019.12.20
requests              2.23.0
requests-oauthlib     1.3.0
resampy               0.2.2
retrying              1.3.3
rpy2                  3.2.7
rsa                   4.0
s3fs                  0.4.2
s3transfer            0.3.3
scikit-image          0.16.2
scikit-learn          0.22.2.post1
scipy                 1.4.1
screen-resolution-extra  0.0.0
scs                   2.1.2
seaborn               0.10.1
Send2Trash            1.5.0
setuptools            46.4.0
setuptools-git        1.2
Shapely               1.7.0
simplegeneric         0.8.1
six                   1.12.0
sklearn               0.0
```

```
sklearn-pandas            1.8.0
smart-open                2.0.0
snowballstemmer           2.0.0
sortedcontainers          2.1.0
spacy                     2.2.4
Sphinx                    1.8.5
sphinxcontrib-websupport  1.2.2
SQLAlchemy                1.3.17
sqlparse                  0.3.1
srsly                     1.0.2
statsmodels               0.10.2
sympy                     1.1.1
tables                    3.4.4
tabulate                  0.8.7
tbb                       2020.0.133
tblib                     1.6.0
tensorboard               2.2.1
tensorboard-plugin-wit    1.6.0.post3
tensorboardcolab          0.0.22
tensorflow                2.2.0
tensorflow-addons         0.8.3
tensorflow-datasets       2.1.0
tensorflow-estimator      2.2.0
tensorflow-gcs-config     2.1.8
tensorflow-hub            0.8.0
tensorflow-metadata       0.22.0
tensorflow-privacy        0.2.2
tensorflow-probability    0.10.0
termcolor                 1.1.0
terminado                 0.8.3
testpath                  0.4.4
text-unidecode            1.3
textblob                  0.15.3
textgenrnn                1.4.1
Theano                    1.0.4
thinc                     7.4.0
```

Версии:

- numpy : 1.18.4
- pandas : 1.0.3
- sklearn : 0.21.3

```
traitlets                 4.3.3
```

▾ загрузим сгенерированный датасет и обучим классифи

```
tzlocal                   1.5.1

#let s fix random state
value = 42
random.seed(value)
np.random.seed(value)

wcwidth                   0.1.9

#let s upload file with dataset pokupty20000.csv
from google.colab import files
uploaded = files.upload()
for fn in uploaded.keys():
```