In [1]:
```
from azureml import Workspace

ws = Workspace(
    workspace_id='7eba2ffdc55c4bf8afc0c8bf59a6bae4',
    authorization_token='114ab543068a447ead6e42797939b154',
    endpoint='https://studioapi.azureml.net')
ds = ws.datasets['Churn.csv']
frame = ds.to_dataframe()
```

In [2]:
```
frame
```

Out[2]:

| | COLLEGE | INCOME | OVERAGE | LEFTOVER | HOUSE | HANDSET_PRICE |
|---|---|---|---|---|---|---|
| **0** | 0 | 45094 | 0 | 0 | 668409 | 182 |
| **1** | 0 | 156268 | 0 | 9 | 223030 | 703 |
| **2** | 0 | 29679 | 53 | 58 | 289246 | 222 |
| **3** | 0 | 34200 | 108 | 66 | 180900 | 233 |
| **4** | 1 | 102370 | 0 | 17 | 259047 | 525 |
| **5** | 1 | 88490 | 0 | 6 | 984146 | 283 |
| **6** | 0 | 45370 | 182 | 19 | 187214 | 863 |
| **7** | 1 | 115339 | 0 | 23 | 730166 | 429 |
| **8** | 1 | 48886 | 0 | 23 | 183889 | 239 |
| **9** | 1 | 83827 | 0 | 7 | 579943 | 255 |
| **10** | 0 | 75859 | 0 | 5 | 266790 | 368 |
| **11** | 0 | 119745 | 41 | 44 | 269792 | 499 |
| **12** | 1 | 98258 | 190 | 22 | 297399 | 303 |
| **13** | 0 | 36819 | 223 | 0 | 366634 | 147 |
| **14** | 0 | 45758 | 67 | 61 | 293683 | 213 |
| **15** | 0 | 65343 | 0 | 55 | 871972 | 785 |
| **16** | 0 | 151267 | 221 | 14 | 897373 | 726 |
| **17** | 1 | 124336 | 0 | 0 | 459136 | 587 |
| **18** | 0 | 119512 | 51 | 31 | 248566 | 229 |
| **19** | 1 | 75947 | 0 | 0 | 721881 | 393 |
| **20** | 1 | 137729 | 46 | 11 | 207827 | 594 |
| **21** | 1 | 69169 | 235 | 0 | 686451 | 384 |
| **22** | 0 | 157812 | 188 | 0 | 920460 | 515 |
| **23** | 1 | 34181 | 0 | 60 | 996116 | 136 |
| **24** | 0 | 70000 | 0 | 83 | 620731 | 321 |
| **25** | 1 | 84215 | 40 | 20 | 651590 | 301 |
| **26** | 1 | 52650 | 0 | 23 | 673502 | 269 |
| **27** | 1 | 32450 | 86 | 52 | 587726 | 535 |
| **28** | 0 | 31086 | 0 | 0 | 442509 | 228 |
| **29** | 1 | 66329 | 166 | 0 | 549879 | 340 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ... | ... | ... | ... | ... | ... | ... |
| **13970** | 0 | 65452 | 172 | 0 | 761792 | 284 |
| **13971** | 0 | 105034 | 0 | 0 | 694417 | 492 |

In [5]:
```
import pandas as pd
y = targets = labels = frame["CHURN"].values

columns = ["COLLEGE", "INCOME", "OVER_15MINS_CALLS_PER_MONTH", "HOUSE",
"REPORTED_3ISFACTION"]
features = frame[list(columns)].values
features
```

Out[5]:
```
array([[     0,  45094,      1, 668409,      4],
       [     0, 156268,      1, 223030,      3],
       [     0,  29679,      5, 289246,      0],
       ...,
       [     0, 129228,      5, 293351,      4],
       [     1,  47398,      0, 434358,      0],
       [     1,  73795,     16, 759481,      0]])
```

In [6]:
```
from sklearn import tree
clf = tree.DecisionTreeClassifier(criterion="entropy", max_depth=3)
clf = clf.fit(X, y)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call l
ast)
<ipython-input-6-9091272ecb4b> in <module>()
      1 from sklearn import tree
      2 clf = tree.DecisionTreeClassifier(criterion="entropy", max_dept
h=3)
----> 3 clf = clf.fit(X, y)

NameError: name 'X' is not defined
```

In [7]:
```
from sklearn.preprocessing import Imputer
imp = Imputer(missing_values='NaN', strategy='mean', axis=0)
X = imp.fit_transform(features)
X
```

Out[7]:
```
array([[  0.00000000e+00,   4.50940000e+04,   1.00000000e+00,
          6.68409000e+05,   4.00000000e+00],
       [  0.00000000e+00,   1.56268000e+05,   1.00000000e+00,
          2.23030000e+05,   3.00000000e+00],
       [  0.00000000e+00,   2.96790000e+04,   5.00000000e+00,
          2.89246000e+05,   0.00000000e+00],
       ...,
       [  0.00000000e+00,   1.29228000e+05,   5.00000000e+00,
          2.93351000e+05,   4.00000000e+00],
       [  1.00000000e+00,   4.73980000e+04,   0.00000000e+00,
          4.34358000e+05,   0.00000000e+00],
       [  1.00000000e+00,   7.37950000e+04,   1.60000000e+01,
          7.59481000e+05,   0.00000000e+00]])
```

In [8]:
```
features
```

Out[8]:
```
array([[     0,  45094,      1, 668409,      4],
       [     0, 156268,      1, 223030,      3],
       [     0,  29679,      5, 289246,      0],
       ...,
       [     0, 129228,      5, 293351,      4],
       [     1,  47398,      0, 434358,      0],
       [     1,  73795,     16, 759481,      0]])
```

In [9]:
```
X= features
```

In [10]:
```
from sklearn import tree
clf = tree.DecisionTreeClassifier(criterion="entropy", max_depth=3)
clf = clf.fit(X, y)
```

In [23]:
```
from sklearn.externals.six import StringIO
with open("Churn.dot", 'w') as f:
    f = tree.export_graphviz(clf, out_file=f, feature_names=columns)
```

In [19]:
```
dot -Tpng churn.dot -o churn.png
```

```
  File "<ipython-input-19-87ce5789710c>", line 1
    dot -Tpng churn.dot -o churn.png
                ^
SyntaxError: invalid syntax
```

In [24]:
```
from sklearn.datasets import load_iris
from sklearn import tree
iris = load_iris()
clf = tree.DecisionTreeClassifier()
clf = clf.fit(iris.data, iris.target)
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call l
ast)
<ipython-input-24-9aa4e8437e99> in <module>()
      1 from subprocess import check_call
----> 2 check_call(['dot','-Tpng','Churn.dot','-o','churn.png'])

/home/nbuser/env3/lib/python3.4/subprocess.py in check_call(*popenargs,
**kwargs)
    554     check_call(["ls", "-l"])
    555     """
--> 556     retcode = call(*popenargs, **kwargs)
    557     if retcode:
    558         cmd = kwargs.get("args")

/home/nbuser/env3/lib/python3.4/subprocess.py in call(timeout, *popenar
gs, **kwargs)
    535     retcode = call(["ls", "-l"])
    536     """
--> 537     with Popen(*popenargs, **kwargs) as p:
    538         try:
    539             return p.wait(timeout=timeout)

/home/nbuser/env3/lib/python3.4/subprocess.py in __init__(self, args, b
ufsize, executable, stdin, stdout, stderr, preexec_fn, close_fds, shel
l, cwd, env, universal_newlines, startupinfo, creationflags, restore_si
gnals, start_new_session, pass_fds)
    857                                 c2pread, c2pwrite,
    858                                 errread, errwrite,
--> 859                                 restore_signals, start_new_sess
ion)
    860         except:
    861             # Cleanup if the child failed starting.

/home/nbuser/env3/lib/python3.4/subprocess.py in _execute_child(self, a
rgs, executable, preexec_fn, close_fds, pass_fds, cwd, env, startupinf
o, creationflags, shell, p2cread, p2cwrite, c2pread, c2pwrite, errread,
errwrite, restore_signals, start_new_session)
   1461                             else:
   1462                                 err_msg += ': ' + repr(orig_exe
cutable)
-> 1463                     raise child_exception_type(errno_num, err_m
sg)
   1464                 raise child_exception_type(err_msg)
   1465

FileNotFoundError: [Errno 2] No such file or directory: 'dot'
```

In [ ]: