In [1]:
```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn import metrics
```

In [2]:
```
# Load Data

data = pd.read_csv('diabetes.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [4]:
```
# Features Selection

x = data.iloc[:,:-1].values
y = data.iloc[:,-1].values
```

In [5]:
```
#split Data
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3,random_state=1)
```

In [6]:
```
# Build model
dtc = DecisionTreeClassifier()
dtc = dtc.fit(x_train,y_train)
y_pred = dtc.predict(x_test)
```

In [7]:
```
# Evaluate Model

metrics.accuracy_score(y_test,y_pred)
```

Out[7]:
```
0.6796536796536796
```

In [ ]:

In [ ]:

In [ ]:

In [8]:
```
data.columns[:-1]
```

Out[8]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age'],
      dtype='object')
```

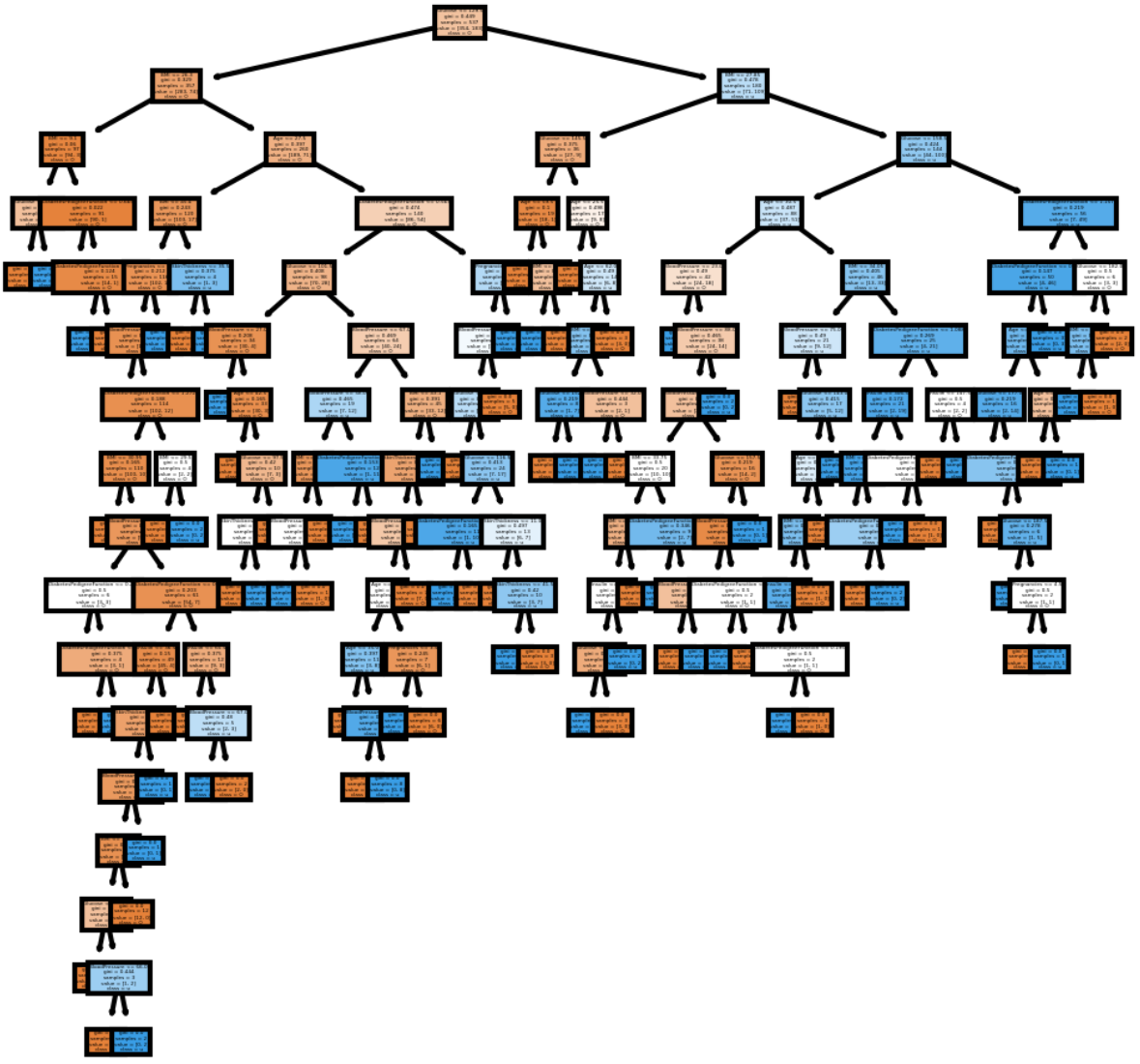
In [9]:
```
from sklearn import tree
import matplotlib.pyplot as plt

features = data.columns[:-1]

fig, axes = plt.subplots(nrows = 1,ncols = 1,figsize = (4,4), dpi=300)

tree.plot_tree(dtc,feature_names = features,class_names=data.columns[-1],
               filled = True);

fig.savefig('imagename.png')
```



In [12]:
```
from sklearn import tree
```

In [13]:
```
tree.plot_tree(dtc)
```

Out[13]:
```
[Text(134.60182584269663, 211.04470588235293, 'X[1] <= 129.5\ngini = 0.449\nsamples = 537\nvalue = [354, 183]'),
 Text(49.2560393258427, 198.25411764705882, 'X[5] <= 26.3\ngini = 0.329\nsamples = 357\nvalue = [283, 74]'),
 Text(15.047191011235956, 185.4635294117647, 'X[5] <= 9.1\ngini = 0.06\nsamples = 97\nvalue = [94, 3]'),
 Text(7.523595505617978, 172.6729411764706, 'X[1] <= 114.5\ngini = 0.444\nsamples = 6\nvalue = [4, 2]'),
 Text(3.761797752808989, 159.88235294117646, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(11.285393258426968, 159.88235294117646, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(22.570786516853936, 172.6729411764706, 'X[6] <= 0.669\ngini = 0.022\nsamples = 91\nvalue = [90, 1]'),
 Text(18.808988764044944, 159.88235294117646, 'gini = 0.0\nsamples = 76\nvalue = [76, 0]'),
 Text(26.332584269662924, 159.88235294117646, 'X[6] <= 0.705\ngini = 0.124\nsamples = 15\nvalue = [14, 1]'),
 Text(22.570786516853936, 147.09176470588235, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(30.094382022471912, 147.09176470588235, 'gini = 0.0\nsamples = 14\nvalue = [14, 0]'),
 Text(83.46488764044945, 185.4635294117647, 'X[7] <= 27.5\ngini = 0.397\nsamples = 260\nvalue = [189, 71]'),
```

In [21]:
```
import graphviz
```

In [22]:
```
dot_data = tree.export_graphviz(dtc, out_file=None)
```

In [23]:
```
graph = graphviz.Source(dot_data)
```

In [27]:
```
dot_data = tree.export_graphviz(dtc, out_file=None,
                                filled=True, rounded=True,special_characters=True)
```

In [28]:
```
graph = graphviz.Source(dot_data)
```

In [30]:
```
graph.view()
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call last)
C:\Anaconda3\envs\daailab2020\lib\site-packages\graphviz\backend.py in run(cmd, input, capture_out
put, check, encoding, quiet, **kwargs)
    163     try:
--> 164         proc = subprocess.Popen(cmd, startupinfo=get_startupinfo(), **kwargs)
    165     except OSError as e:

C:\Anaconda3\envs\daailab2020\lib\subprocess.py in __init__(self, args, bufsize, executable, stdi
n, stdout, stderr, preexec_fn, close_fds, shell, cwd, env, universal_newlines, startupinfo, creati
onflags, restore_signals, start_new_session, pass_fds, encoding, errors, text)
    799                                 errread, errwrite,
--> 800                                 restore_signals, start_new_session)
    801         except:

C:\Anaconda3\envs\daailab2020\lib\subprocess.py in _execute_child(self, args, executable, preexec_
fn, close_fds, pass_fds, cwd, env, startupinfo, creationflags, shell, p2cread, p2cwrite, c2pread,
 c2pwrite, errread, errwrite, unused_restore_signals, unused_start_new_session)
   1206                                         os.fspath(cwd) if cwd is not None else None,
-> 1207                                         startupinfo)
   1208             finally:

FileNotFoundError: [WinError 2] The system cannot find the file specified

During handling of the above exception, another exception occurred:

ExecutableNotFound                        Traceback (most recent call last)
<ipython-input-30-7860104da997> in <module>
----> 1 graph.view()

C:\Anaconda3\envs\daailab2020\lib\site-packages\graphviz\files.py in view(self, filename, director
y, cleanup, quiet, quiet_view)
    240         return self.render(filename=filename, directory=directory,
    241                            view=True, cleanup=cleanup,
--> 242                            quiet=quiet, quiet_view=quiet_view)
    243
    244     def _view(self, filepath, format, quiet):

C:\Anaconda3\envs\daailab2020\lib\site-packages\graphviz\files.py in render(self, filename, direct
ory, view, cleanup, format, renderer, formatter, quiet, quiet_view)
    207         rendered = backend.render(self._engine, format, filepath,
    208                                   renderer=renderer, formatter=formatter,
--> 209                                   quiet=quiet)
    210
    211         if cleanup:

C:\Anaconda3\envs\daailab2020\lib\site-packages\graphviz\backend.py in render(***failed resolving
 arguments***)
    219         cwd = None
    220
--> 221     run(cmd, capture_output=True, cwd=cwd, check=True, quiet=quiet)
    222     return rendered
    223

C:\Anaconda3\envs\daailab2020\lib\site-packages\graphviz\backend.py in run(cmd, input, capture_out
put, check, encoding, quiet, **kwargs)
    165     except OSError as e:
    166         if e.errno == errno.ENOENT:
--> 167             raise ExecutableNotFound(cmd)
    168         else:
    169             raise

ExecutableNotFound: failed to execute ['dot', '-Tpdf', '-O', 'Source.gv'], make sure the Graphviz
 executables are on your systems' PATH
```

In [ ]: