CSC411 - Fall18
Assignment 1

Mahsa Naserifar - 1002330872

1A:
In order to calculate the mean value of the function we need to perform a double integral on the 2d space on the probability density function, which in this case is a uniformly distributed 1 along the space.

```
>>> z = lambda y, x: (x-y)**2
>>> E = dblquad(z, 0, 1, lambda x: 0, lambda x: 1)[0]
>>> E
```
0.16666666666666666

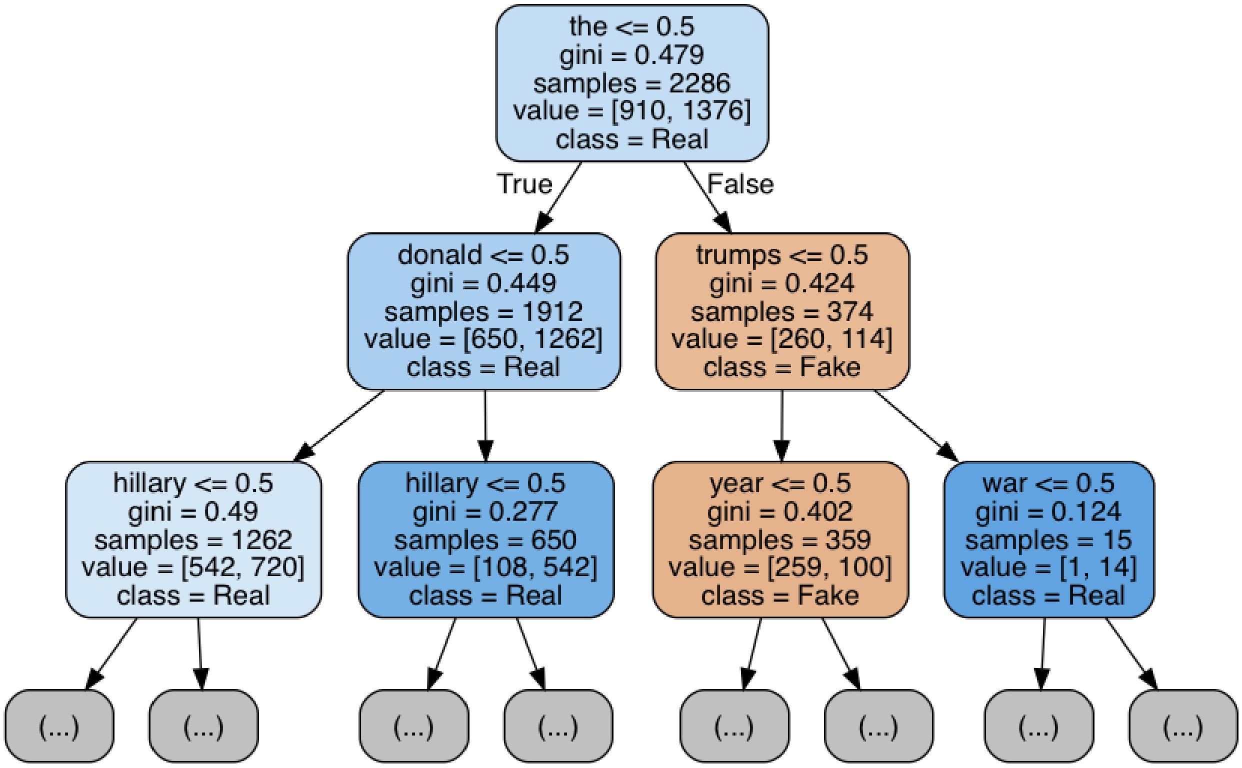
Variance of a squared function is equal to: var[X^2] = E[X^2] - E[X]^2. Therefore we need to perform a double integral to get E[X^2] with uniform probability density function.

```
>>> z = lambda y, x: (x-y)**4
>>> var = dblquad(z, 0, 1 , lambda x: 0, lambda x: 1)[0] - E**2
>>> var
```
0.03888888888888889

1B:
Since the two operations of expected value and variance are linear and distributable, we can simplify the summation of Z1, Z2, Z3 … to d * Z, hence we have:
E[R] = d * E[Z]
var[R] = d * var[Z]

2B:
Criterion: gini & Max Depth: 1 | Accuracy: 0.677551020408
Criterion: entropy & Max Depth: 1 | Accuracy: 0.59387755102
Criterion: gini & Max Depth: 10 | Accuracy: 0.738775510204
Criterion: entropy & Max Depth: 10 | Accuracy: 0.724489795918
Criterion: gini & Max Depth: 50 | Accuracy: 0.761224489796
Criterion: entropy & Max Depth: 50 | Accuracy: 0.769387755102
Criterion: gini & Max Depth: 100 | Accuracy: 0.795918367347
Criterion: entropy & Max Depth: 100 | Accuracy: 0.775510204082
Criterion: gini & Max Depth: 200 | Accuracy: 0.779591836735
Criterion: entropy & Max Depth: 200 | Accuracy: 0.773469387755

2C:



2D:

trump 0.0235811950844
win 0.389685648004
war 0.353359335021
america 0.0946357977378
tribute 0.0666808930443