DATE

S M T W T F S

Q1a) $Z = (X-Y)^2$

$E(Z) = E((X-Y)^2)$

$= E(X^2 - 2XY + Y^2)$

$= E(X^2) - 2E(YX) + E(Y^2)$

$= \frac{1}{3} - 2E(X)E(Y) + \frac{1}{3}$

$= \frac{2}{3} - \frac{2}{4}$

$= \frac{1}{6}$

$E(X) = E(Y) = \frac{1-0}{2} = \frac{1}{2}$ $\quad Var(X) = Var(Y) = \frac{(1-0)^2}{12} = \frac{1}{12}$

$E(X^2) = E(X^2) = Var(X) + E(X)^2$

$= \frac{1}{12} + \frac{1}{4}$

$= \frac{1}{3}$

integral formula

$E(X^3) = \int_0^1 x^3 f(x)\,dx = \frac{1}{4}$

$E(X^4) = \int_0^1 x^4 f(x)\,dx = \int_0^1 x^4\,dx = \frac{1}{5}$

$Var(Z) = E(Z^2) - E(Z)^2$

$= E((X-Y)^4) - (\frac{1}{36})$

$= E(X^4) - 4E(X^3Y) + 6E(X^2Y^2) - 4E(Y^3X) + E(Y^4) - \frac{1}{36}$

$= 0.2 - 4 \times 0.25 \times \frac{1}{2} + 6 \times (\frac{1}{3})^2 - 4 \times 0.25 \times \frac{1}{2} - \frac{1}{36}$

$= 0.0\overline{3}$

DATE

S M T W T F S

Q1b) $E(R) = E\left(\sum_{i=1}^{d} z_i\right)$

$= \sum_{i=1}^{d} E(z_i)$

$= \sum_{i=1}^{d} E(z)$

$= d\,E(z).$

$Var(R) = var\left(\sum_{i=1}^{d} z_i\right)$

$= \sum_{i,j=1}^{d} cov(z_i, z_j)$

$= \sum_{i=j}^{d} (ov(z_i, z_j) + \sum_{i \neq j}^{d} cov(z_i, z_j)$

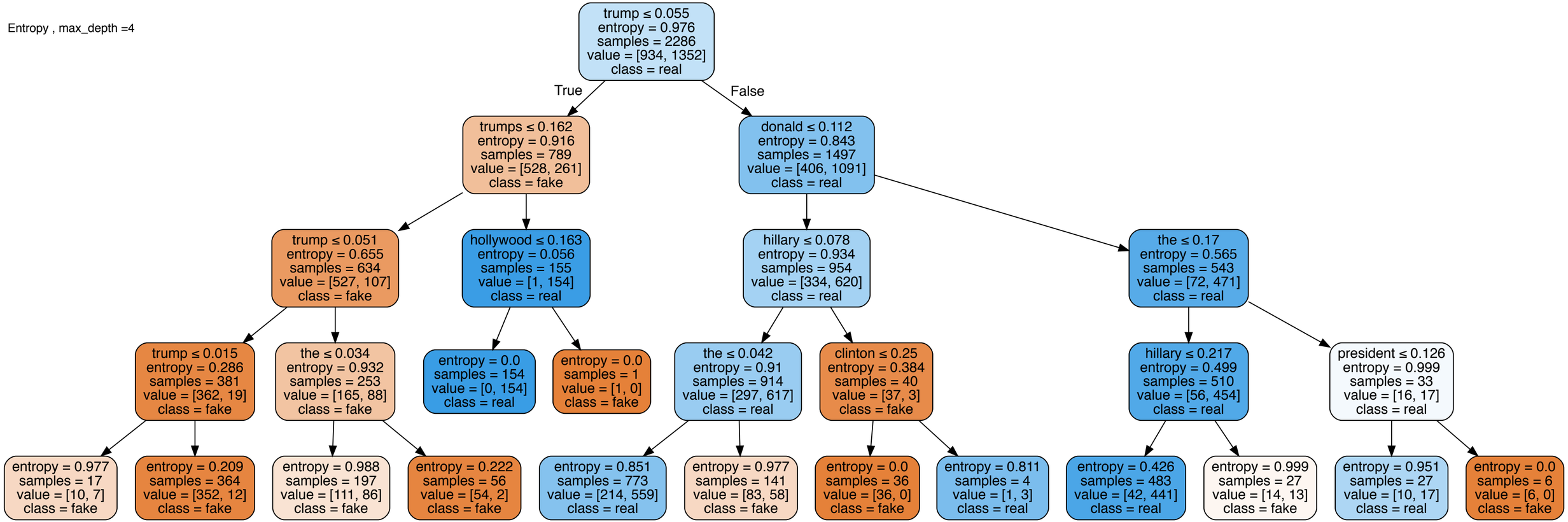
$= \sum_{i=1}^{d} var(z_i) + 0$

$= d\,var(z).$

z_i, z_j independent if $i \neq j$, $cov(z_i, z_j) = 0$

```
/anaconda3/bin/python /Users/yufei/PycharmProjects/csc411/hw1/hw1_code.py
criterion=gini, max depth = 4, score is 0.7868852459016393
criterion=gini, max depth = 6, score is 0.819672131147541
criterion=gini, max depth = 8, score is 0.819672131147541
criterion=gini, max depth = 10, score is 0.8135245901639344
criterion=gini, max depth = 12, score is 0.8114754098360656
criterion=gini, max depth = 14, score is 0.8114754098360656
criterion=entropy, max depth = 4, score is 0.7991803278688525
criterion=entropy, max depth = 6, score is 0.7889344262295082
criterion=entropy, max depth = 8, score is 0.7930327868852459
criterion=entropy, max depth = 10, score is 0.7868852459016393
criterion=entropy, max depth = 12, score is 0.7848360655737705
criterion=entropy, max depth = 14, score is 0.7622950819672131
keyword:  trump ;information gain:  0.0082

Process finished with exit code 0
```


Entropy , max_depth =4
trump ≤ 0.055
entropy = 0.976
samples = 2286
value = [934, 1352]
class = real
True
False
trumps ≤ 0.162
entropy = 0.916
samples = 789
value = [528, 261]
class = fake
donald ≤ 0.112
entropy = 0.843
samples = 1497
value = [406, 1091]
class = real
trump ≤ 0.051
entropy = 0.655
samples = 634
value = [527, 107]
class = fake
hollywood ≤ 0.163
entropy = 0.056
samples = 155
value = [1, 154]
class = real
hillary ≤ 0.078
entropy = 0.934
samples = 954
value = [334, 620]
class = real
the ≤ 0.17
entropy = 0.565
samples = 543
value = [72, 471]
class = real
trump ≤ 0.015
entropy = 0.286
samples = 381
value = [362, 19]
class = fake
the ≤ 0.034
entropy = 0.932
samples = 253
value = [165, 88]
class = fake
entropy = 0.0
samples = 154
value = [0, 154]
class = real
entropy = 0.0
samples = 1
value = [1, 0]
class = fake
the ≤ 0.042
entropy = 0.91
samples = 914
value = [297, 617]
class = real
clinton ≤ 0.25
entropy = 0.384
samples = 40
value = [37, 3]
class = fake
hillary ≤ 0.217
entropy = 0.499
samples = 510
value = [56, 454]
class = real
president ≤ 0.126
entropy = 0.999
samples = 33
value = [16, 17]
class = real
entropy = 0.977
samples = 17
value = [10, 7]
class = fake
entropy = 0.209
samples = 364
value = [352, 12]
class = fake
entropy = 0.988
samples = 197
value = [111, 86]
class = fake
entropy = 0.222
samples = 56
value = [54, 2]
class = fake
entropy = 0.851
samples = 773
value = [214, 559]
class = real
entropy = 0.977
samples = 141
value = [83, 58]
class = fake
entropy = 0.0
samples = 36
value = [36, 0]
class = fake
entropy = 0.811
samples = 4
value = [1, 3]
class = real
entropy = 0.426
samples = 483
value = [42, 441]
class = real
entropy = 0.999
samples = 27
value = [14, 13]
class = fake
entropy = 0.951
samples = 27
value = [10, 17]
class = real
entropy = 0.0
samples = 6
value = [6, 0]
class = fake