ASSIGNMENT 0

1) Post A: "Cats are cute and funny"
Post B: "Dogs are funny animals"
Post C: "Cats and dogs rarely get along"

Unique Words: Cats, are, cute, and, funny, dogs, animals, rarely, get, along

In post A,	In post B,	In post C,
cats = 1	cats = 0	cats = 1
are = 1	are = 1	are = 0
cute = 1	cute	cute = 0
and = 1	~~funny~~ = 0	and = 1
funny = 1	and = 0	funny = 0
dogs = 0	funny = 1	dogs = 1
animals = 0	dogs = 1	animals = 0
rarely = 0	animals = 1	rarely = 1
get = 0	rarely = 0	get = 1
along = 0	get = 0	along = 1
	along = 0	

Dimension of bag-of-words for this corpus = 10

Matrix;

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

order = 3×10
rows = number of posts
columns = word is appearing in post or not i.e, unique words

---

2) Probability that a post of 10 words does not contain an occurence of "cat" = $(1-p)^{10}$

Probability it contains atleast one occurence of the word "cat"
$= 1 - (1-p)^{10}$

For post of length L being a "cat type" post = $1 - (1-p)^{L}$

Probability Vector ~~for Post A~~,

$$\left[\frac{2}{15}, \frac{2}{15}, \frac{1}{15}, \frac{2}{15}, \frac{2}{15}, \frac{2}{15}, \frac{1}{15}, \frac{1}{15}, \frac{1}{15}, \frac{1}{15}\right]$$

Probability of "cat-like" post for post A = $1 - \left(1 - \frac{2}{15}\right)^{5}$

Theoretical possibility = ½ since there are only two cases

~~or 2/3 if one of~~

Since Post A and C are 'cat' type posts,

$$P(\text{cat type}) = \frac{2}{3}$$

$$P(\text{cute}/\text{cat type}) = \frac{P(\text{cute and cat type})}{P(\text{cat type})} = \frac{1/15}{2/3} = 0.1$$

$$P(\text{cat type}|\text{cute}) = \frac{P(\text{cute}|\text{cat type}) \times P(\text{cat type})}{P(\text{cute})} = \frac{0.1 \times 2/3}{1/15} = 1$$

This tells us that if cute is given, it is always cat type post

3) $U'(L) = \frac{-L}{10} + 3 = 0 \Rightarrow L = 30$

$U'(L) = 0$ at $L = 30$

$U''(L) = -\frac{1}{10} < 0$ which according to second derivative test, shows that maximum upvotes is recieved when $L = 30$

We know, $P(L,p) = 1 - (1-p)^L$

$G_1(L,p) = P(L,p) \cdot U(L)$

$= \left[1 - (1-p)^L\right] \cdot \left[-\frac{L^2}{20} + 3L\right]$

as $L \to \infty$, $G_1(L,p) = [1-0] \cdot [-\infty] = -\infty$