CS498AML HW7 Report

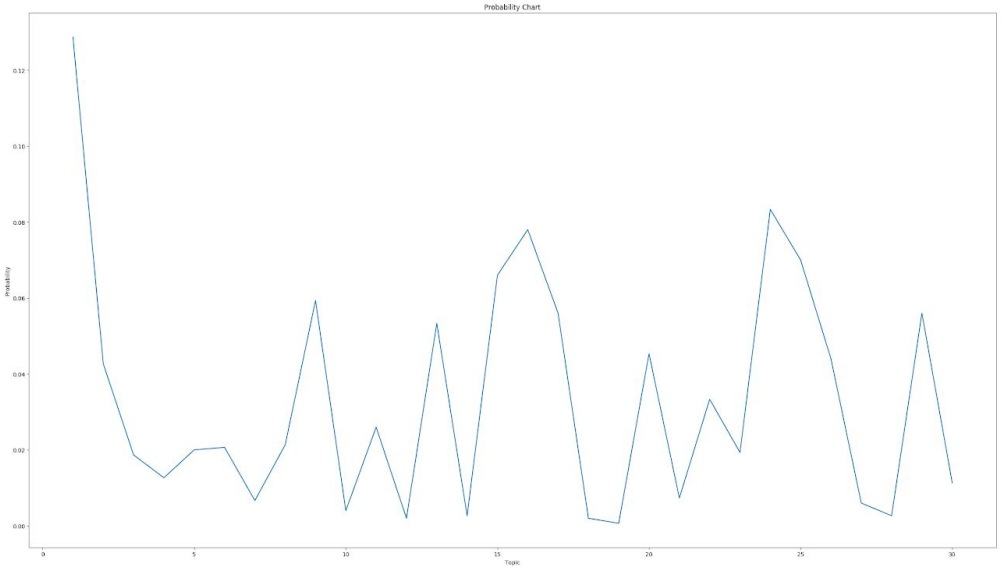
Qingkang Zhang(qzhang72), Ramya Narayanaswamy(rpn2)

Code : multi_EM.py

- Cluster this to 30 topics, using a simple mixture of multinomial topic model, as lectured in class.

For this part, we utilized random initialization with 1500 documents and were able to generate initial 30 multinomial distributions(clusters). To find near local optimization, we implemented EM algorithm as mentioned in lecture.

- Produce a graph showing, for each topic, the probability with which the topic is selected.



- Produce a table showing, for each topic, the 10 words with the highest probability for that topic.

	0	1	2	3	4	5	6	7	8	9
0	system	model	network	neural	function	input	signal	output	circuit	information
1	unit	network	input	learning	weight	hidden	layer	output	pattern	function
2	learning	action	model	task	control	reinforcement	robot	function	system	states

3	algorithm	vector	function	learning	loss	class	set	weight	bound	problem
4	network	unit	input	hidden	output	learning	function	training	pattern	weight
5	weight	network	error	training	set	input	noise	generalization	function	learning
6	network	task	neural	learning	training	architecture	control	solution	input	problem
7	input	network	output	neural	noise	function	training	set	data	information
8	network	training	set	data	neural	error	input	output	unit	learning
9	classifier	training	network	rbf	set	error	neural	problem	center	gaussian
10	word	network	recognition	training	system	model	speech	hmm	neural	set
11	cell	head	direction	rat	model	angular	system	velocity	mcnaughton	neural
12	model	data	network	set	neural	parameter	learning	algorithm	training	function
13	character	field	system	window	network	input	net	set	word	training
14	data	model	algorithm	set	parameter	point	learning	distribution	method	function
15	network	neural	system	input	function	learning	weight	output	model	unit
16	learning	algorithm	function	problem	policy	action	system	optimal	model	result
17	hint	learning	examples	function	error	market	performance	method	information	network
18	monte	carlo	player	decision	policy	base	move	rollout	network	trial
19	object	image	network	images	model	recognition	view	system	set	feature

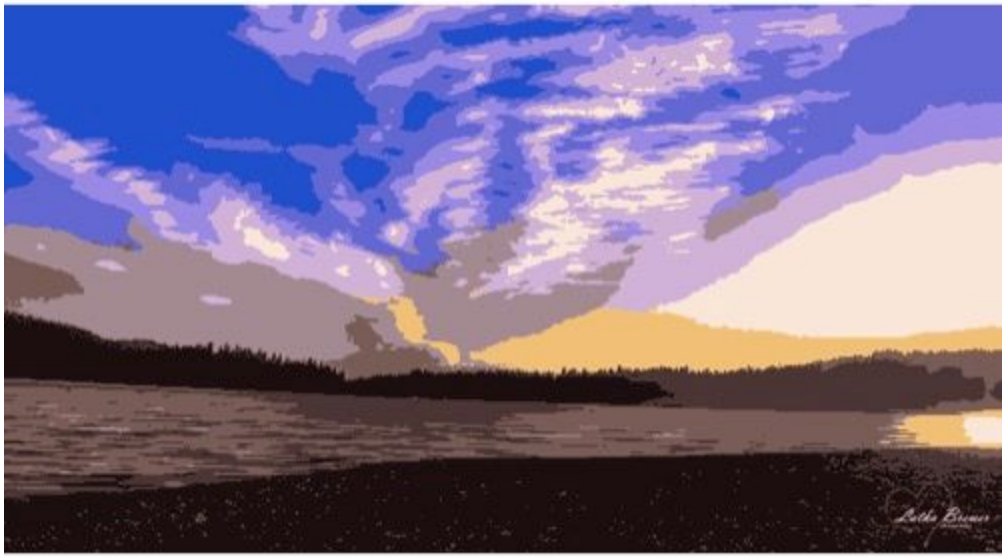
20	function	threshold	network	weight	neural	input	circuit	size	number	result
21	function	set	training	vector	algorithm	error	kernel	data	problem	classifier
22	speech	network	system	model	input	signal	recognition	neural	output	information
23	function	network	algorithm	learning	neural	model	input	problem	set	data
24	cell	model	input	neuron	visual	field	cortex	orientation	response	network
25	neuron	network	input	model	neural	synaptic	system	function	learning	firing
26	david	michael	john	richard	peter	author	index	thomas	eric	paul
27	eeg	component	response	trial	artifact	ica	data	single	visual	erp
28	learning	network	error	weight	training	input	function	algorithm	neural	set
29	model	learning	control	movement	motor	forward	field	arm	dynamic	trajectory

Part2:

Code : Part2Final.R

SmallSunset

# of clusters = 10



# of clusters = 20

# of clusters = 50

RobertMixed

# of clusters = 10

# of clusters = 20

# of clusters = 50

Smallstrelitzia

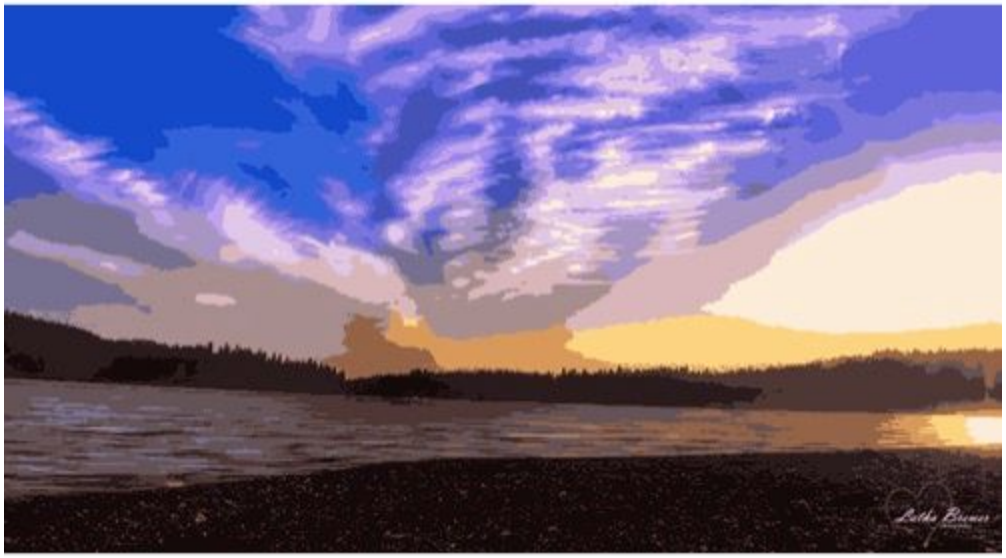
# of clusters = 10

# of clusters = 20

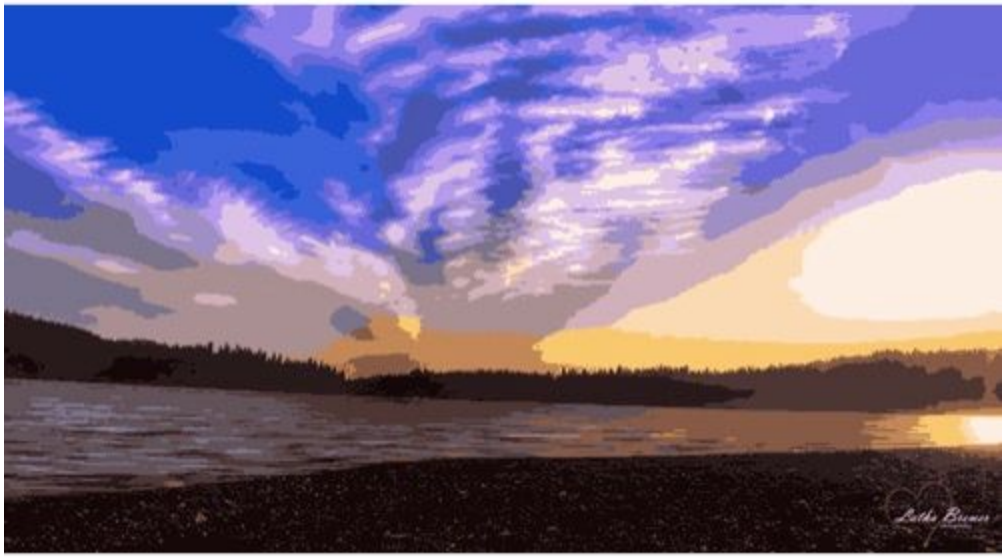
# of clusters = 50

Different Seeds for small sunset

Seed1:



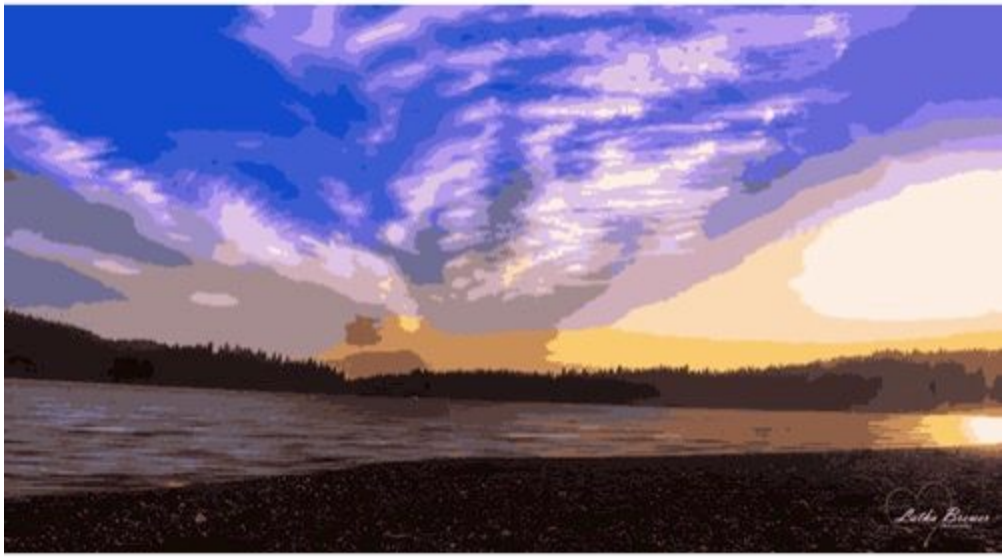
Seed2:



Seed3:

Seed4:

Seed5:



Conclusion: There are very minuscule variations when images are obtained with different seed. One needs to look keenly to see the differences.