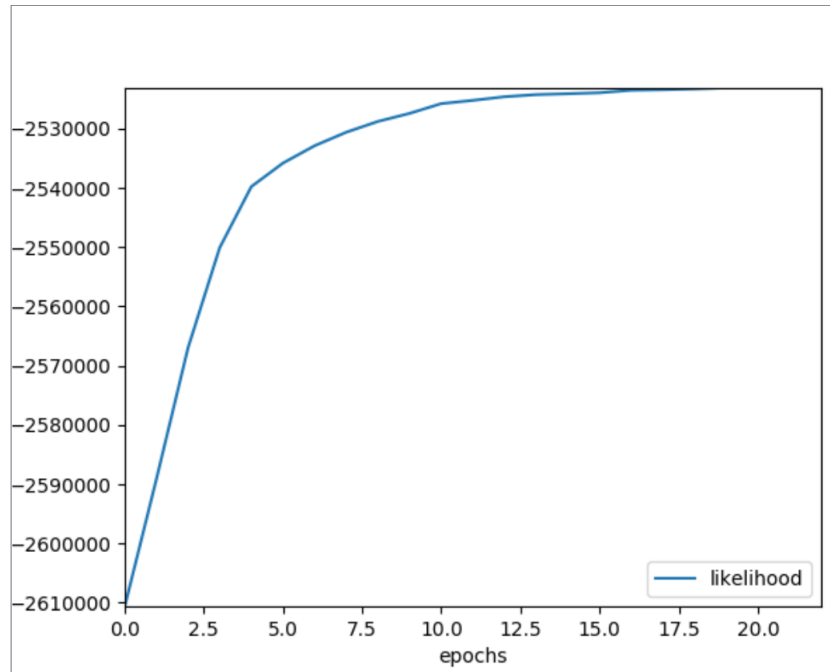


Coral Malachi – 314882853

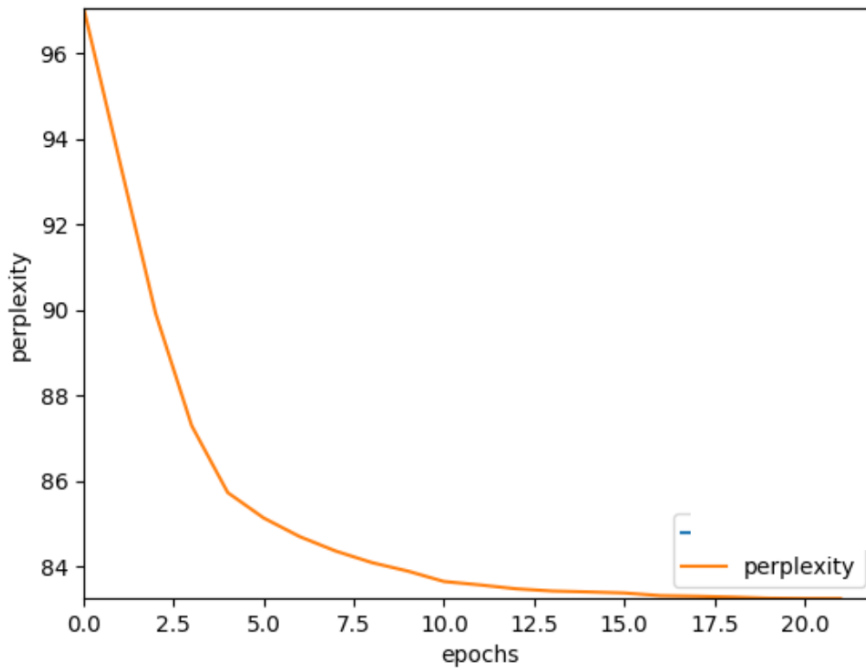
Avishai Zagoury - 209573633

REPORT:

1. EM threshold – 10
2. plot graph show: epochs vs. Log likelihood



plot graph show : epochs vs. Mean Perplexity



### 3. Confusion Matrix:

1	0	256	10	6	4	118	4	7	2	310
2	7	0	9	56	0	0	118	1	0	152
3	1	2	48	22	1	9	14	24	29	93
4	0	0	137	17	5	0	21	73	62	171
5	14	4	0	200	2	1	1	0	0	212
6	2	93	1	10	84	126	2	0	0	261
7	366	2	2	22	1	3	15	0	2	399
8	1	2	168	3	10	0	10	89	60	187
9	9	92	13	13	230	18	1	3	4	339

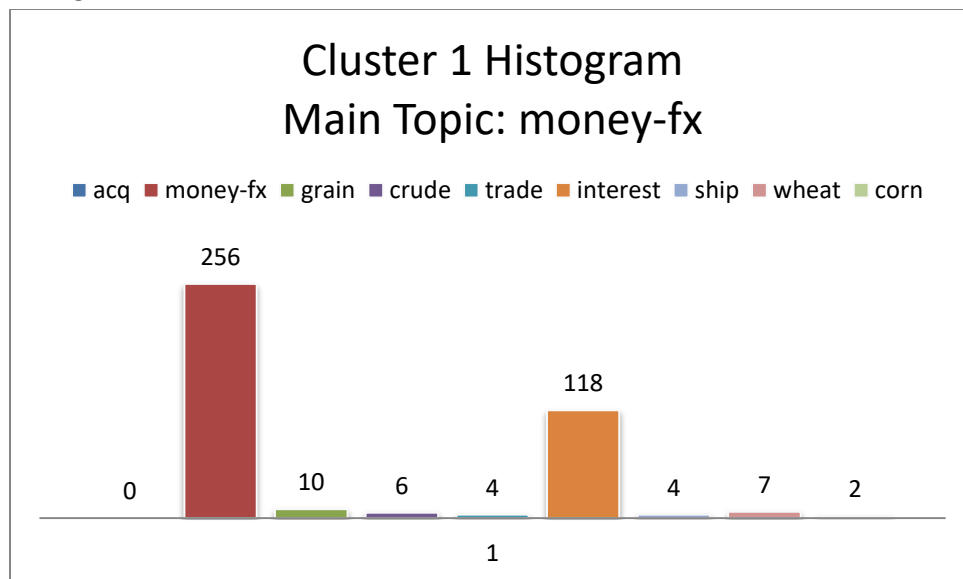
topic	acq	money-fx	grain	crude	trade	interest	ship	wheat	corn	sum
1	0	256	10	6	4	118	4	7	2	310
2	7	0	9	56	0	0	118	1	0	152
3	1	2	48	22	1	9	14	24	29	93
4	0	0	137	17	5	0	21	73	62	171
5	14	4	0	200	2	1	1	0	0	212
6	2	93	1	10	84	126	2	0	0	261
7	366	2	2	22	1	3	15	0	2	399
8	1	2	168	3	10	0	10	89	60	187
9	9	92	13	13	230	18	1	3	4	339

המטריצה לאחר מיונה לפי סדר יורד של ערך העמודה האחרונה:

topic	acq	money-fx	grain	crude	trade	interest	ship	wheat	corn	sum	..
7	366	2	2	22	1	3	15	0	2	399	
9	9	92	13	13	230	18	1	3	4	339	
1	0	256	10	6	4	118	4	7	2	310	
6	2	93	1	10	84	126	2	0	0	261	
5	14	4	0	200	2	1	1	0	0	212	
8	1	2	168	3	10	0	10	89	60	187	
4	0	0	137	17	5	0	21	73	62	171	
2	7	0	9	56	0	0	118	1	0	152	
3	1	2	48	22	1	9	14	24	29	93	

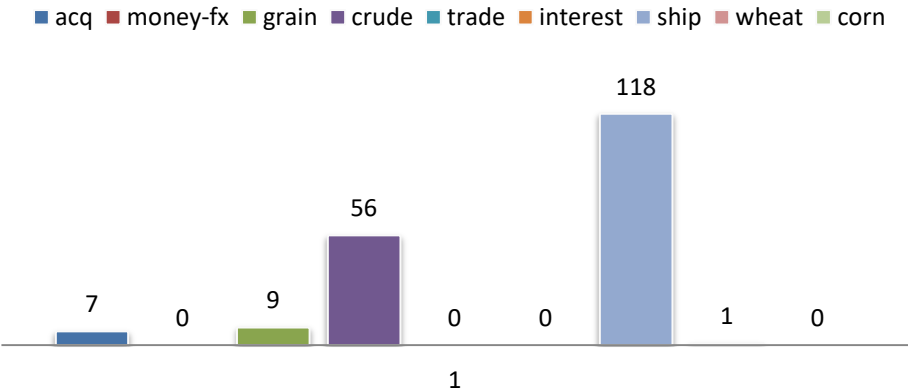
**בודק יקר: אני השארתי את מספור הקלאסטרים כפי שהוא, כלומר כאשר מיינתי את המטריצה מהגדול לקטן לפי העמודה האחרונה, מספור הקלאסטר זז יחד עם השורה. אם זו לא הייתה כוונתך הרי שהמספורים שלי משתנים, וכל קלאסטר יקבל טופיק ראשי אחר.**

4. Histogram for each class:



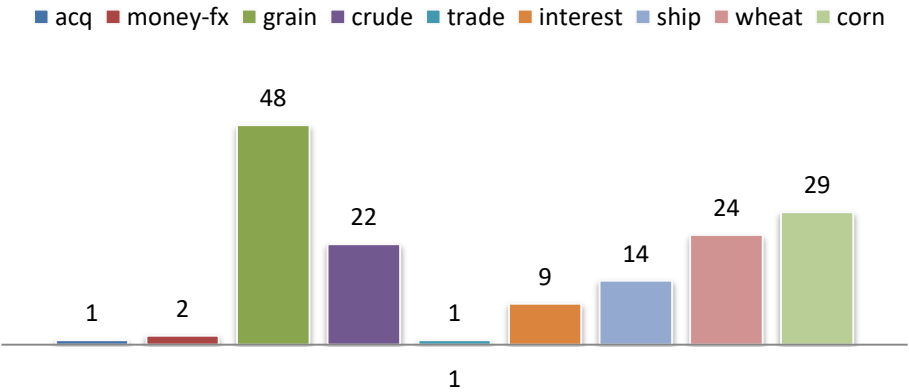
# Cluster 2 Histogram

## Main Topic: ship



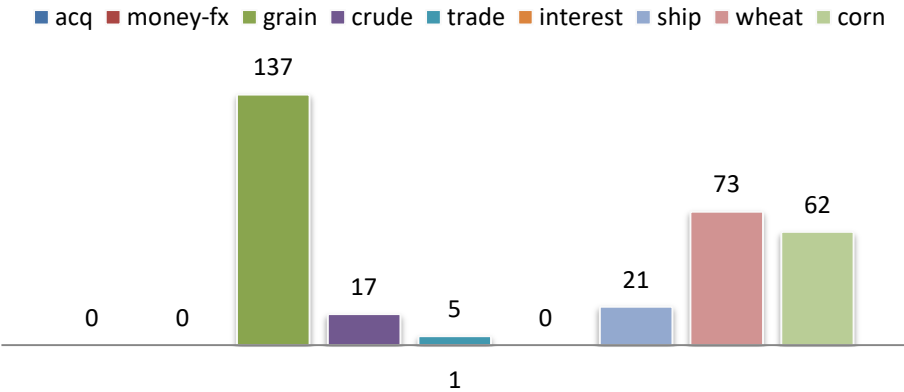
# Cluster 3 Histogram

## Main Topic: grain



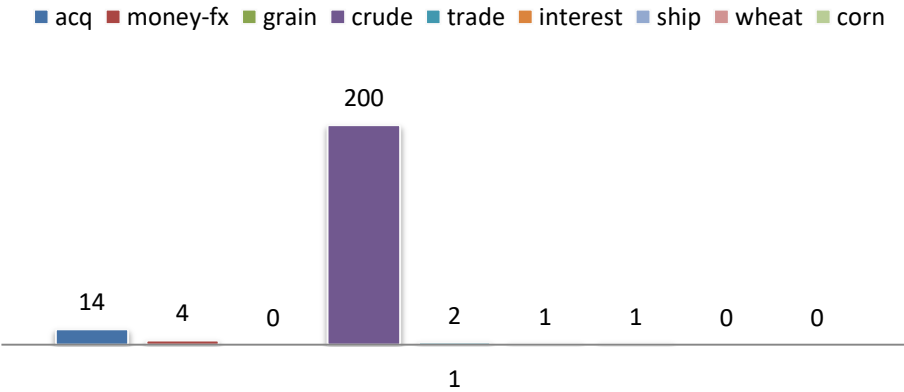
# Cluster 4 Histogram

## Main Topic: grain



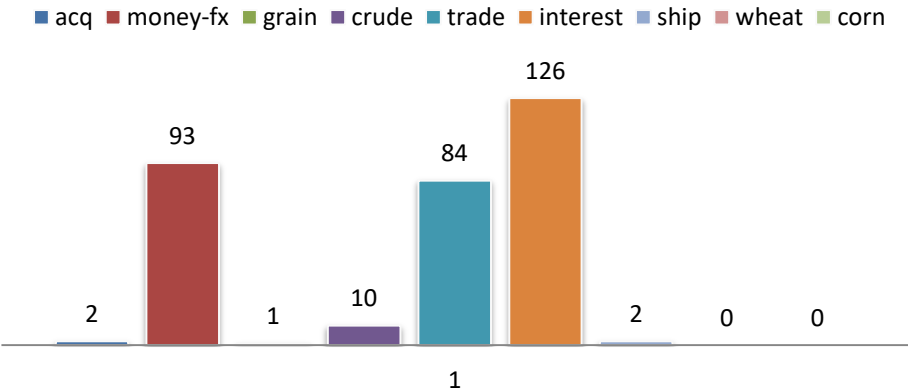
# Cluster 5 Histogram

## Main Topic: crude



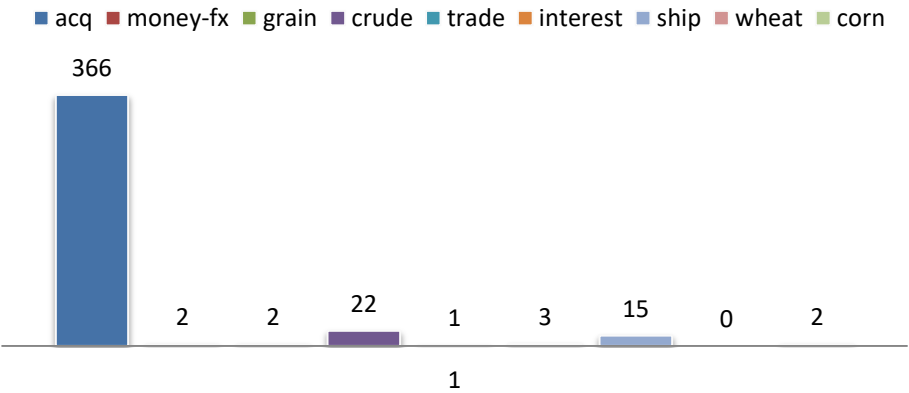
# Cluster 6 Histogram

## Main Topic: interest



# Cluster 7 Histogram

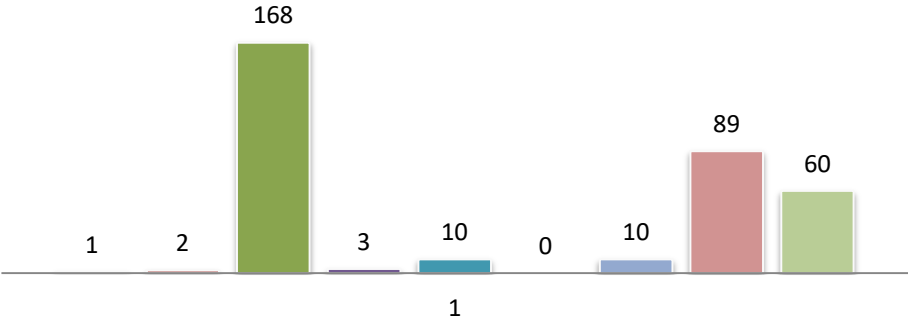
## Main Topic: acq



# Cluster 8 Histogram

## Main Topic: grain

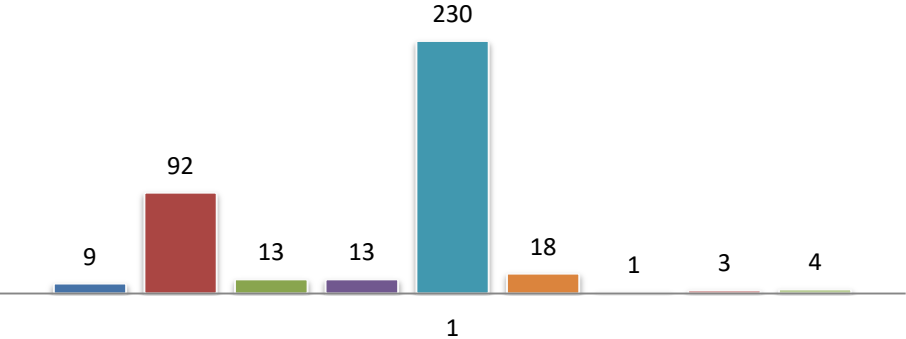
■ acq ■ money-fx ■ grain ■ crude ■ trade ■ interest ■ ship ■ wheat ■ corn



# Cluster 9 Histogram

## Main Topic: trade

■ acq ■ money-fx ■ grain ■ crude ■ trade ■ interest ■ ship ■ wheat ■ corn



5. The accuracy of our model is : 0.7763
6. The parameters:
  - Lambda= 1.1
  - K=10
  - vocabulary size after filtering - 6800