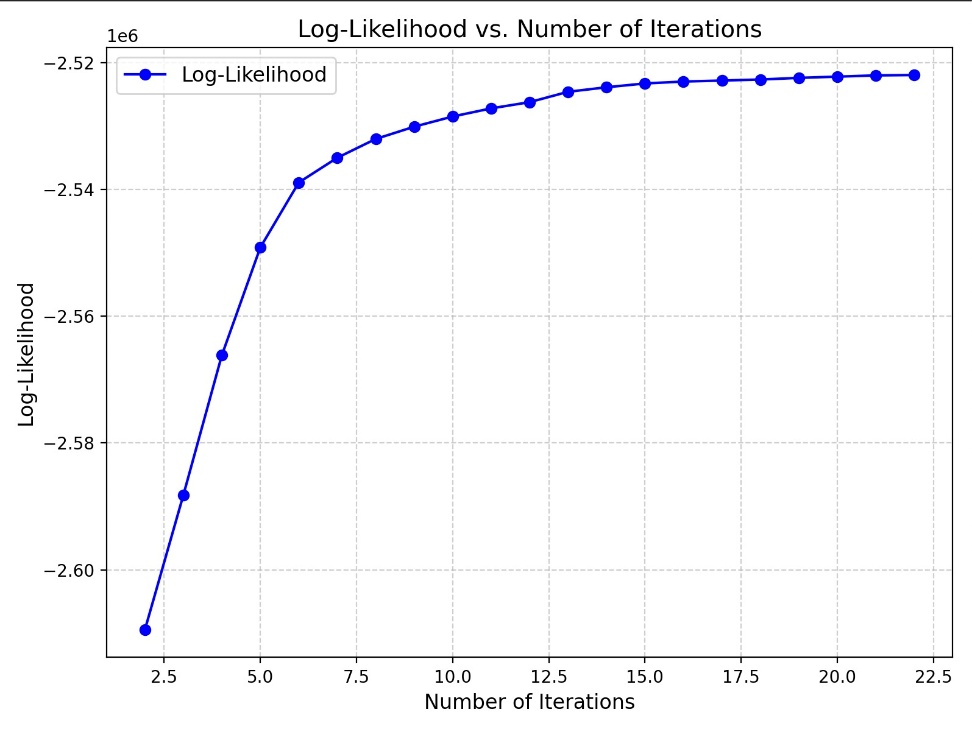
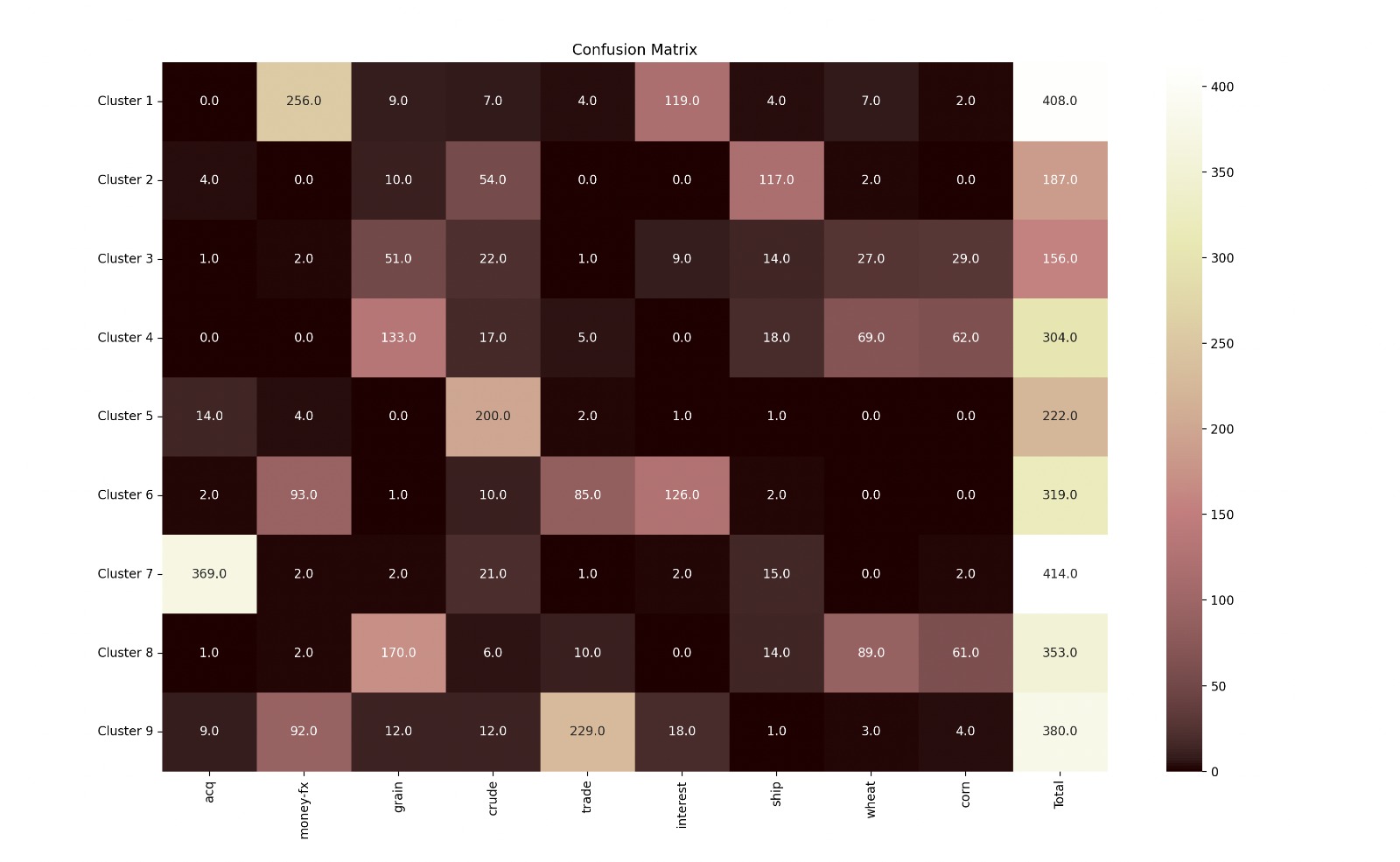
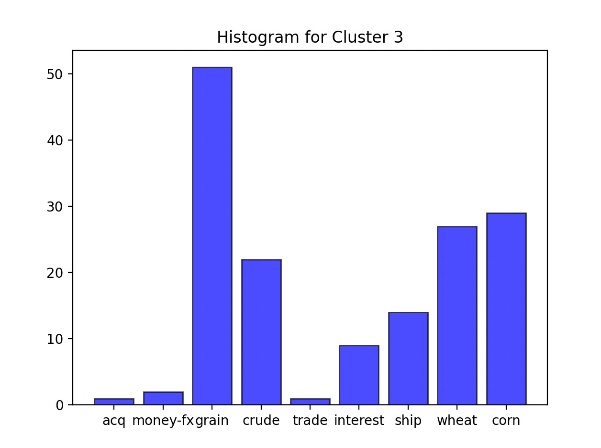
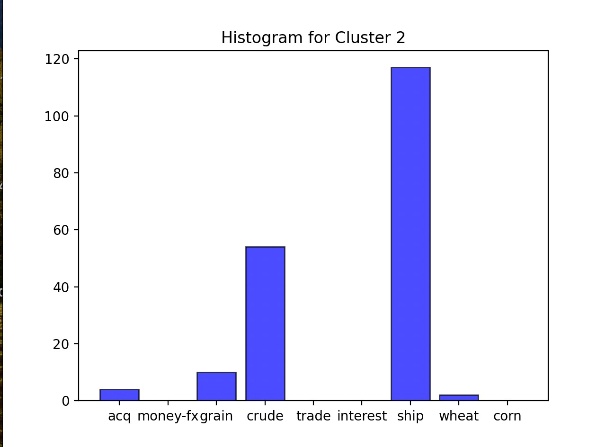
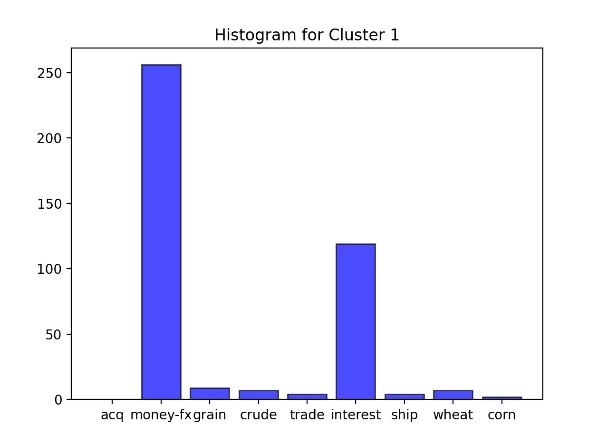
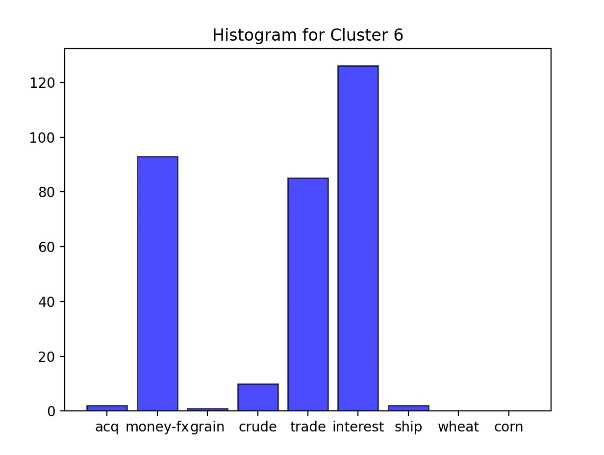
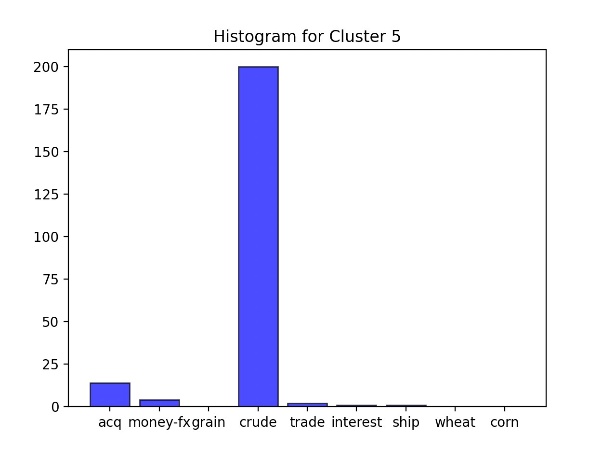
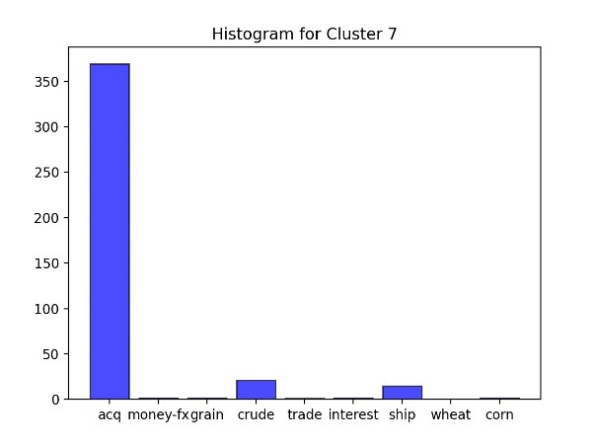
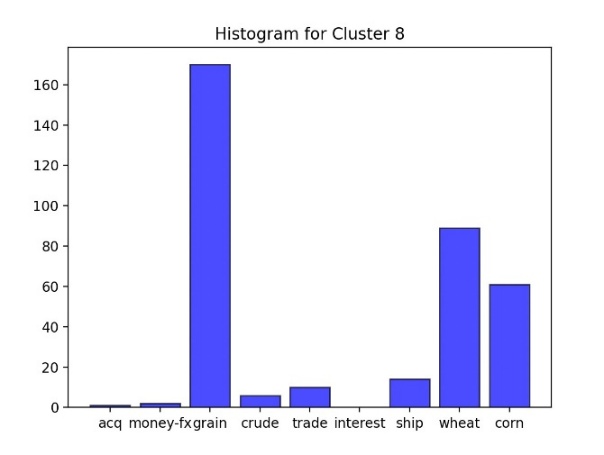
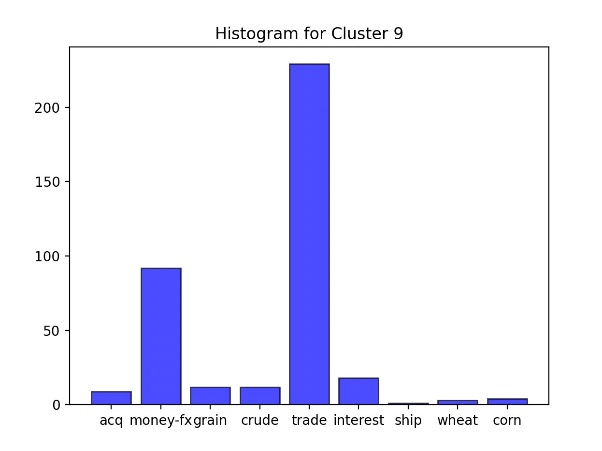
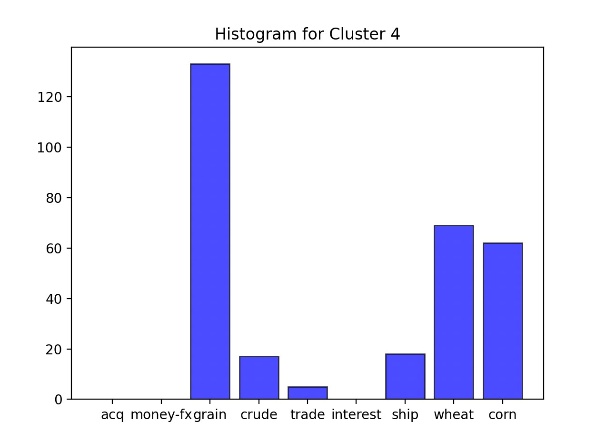
**Applied probability models for CS Exercise 3**

**Students: Lior Tzahar, Adva Cohen**

Report:

1. The threshold to stop the EM algorithm is when the log likelihood stops increasing.
2. A graph showing the number of iterations

   Description automatically generatedLog likelihood plot:  
   We can see that during the EM algorithm run the log likelihood increase as expected.  
    Perplexity plot:  
   We can see that during the EM algorithm run the log Perplexity decreased as expected.
3. Confusion Matrix:
4. The histograms for each cluster:



1. The accuracy we have got after running EM algorithm for 23 iterations is 0.777306967984934 ~ **78%**   
   this result is above 60% as requested.
2. The constants that were best for our model:

K = 10, λ = 1.05, The vocabulary size (V) = 6800