Clustering text data with Gaussian mixtures

4 questions

1 point

1.
Select all the topics that have a cluster in the model created above.

- [ ] Baseball
- [ ] Basketball
- [ ] Soccer/football
- [ ] Music
- [ ] Politics
- [ ] Law
- [ ] Finance

1 point

2.
Try fitting EM with the random initial parameters you created above. What is the final loglikelihood that the algorithm converges to? Choose the range that contains this value.

- ○ Less than 2.2e9
- ○ Between 2.2e9 and 2.3e9

- Between 2.3e9 and 2.4e9
- Between 2.4e9 and 2.5e9
- Greater than 2.5e9

1 point

3.
Is the final loglikelihood larger or smaller than the final loglikelihood we obtained above when initializing EM with the results from running k-means?

- Initializing EM with k-means led to a larger final loglikelihood
- Initializing EM with k-means led to a smaller final loglikelihood

1 point

4.
For the above model, `out_random_init`, use the `visualize_EM_clusters` method you created above. Are the clusters more or less interpretable than the ones found after initializing using k-means?

- More interpretable
- Less interpretable

Submit Quiz