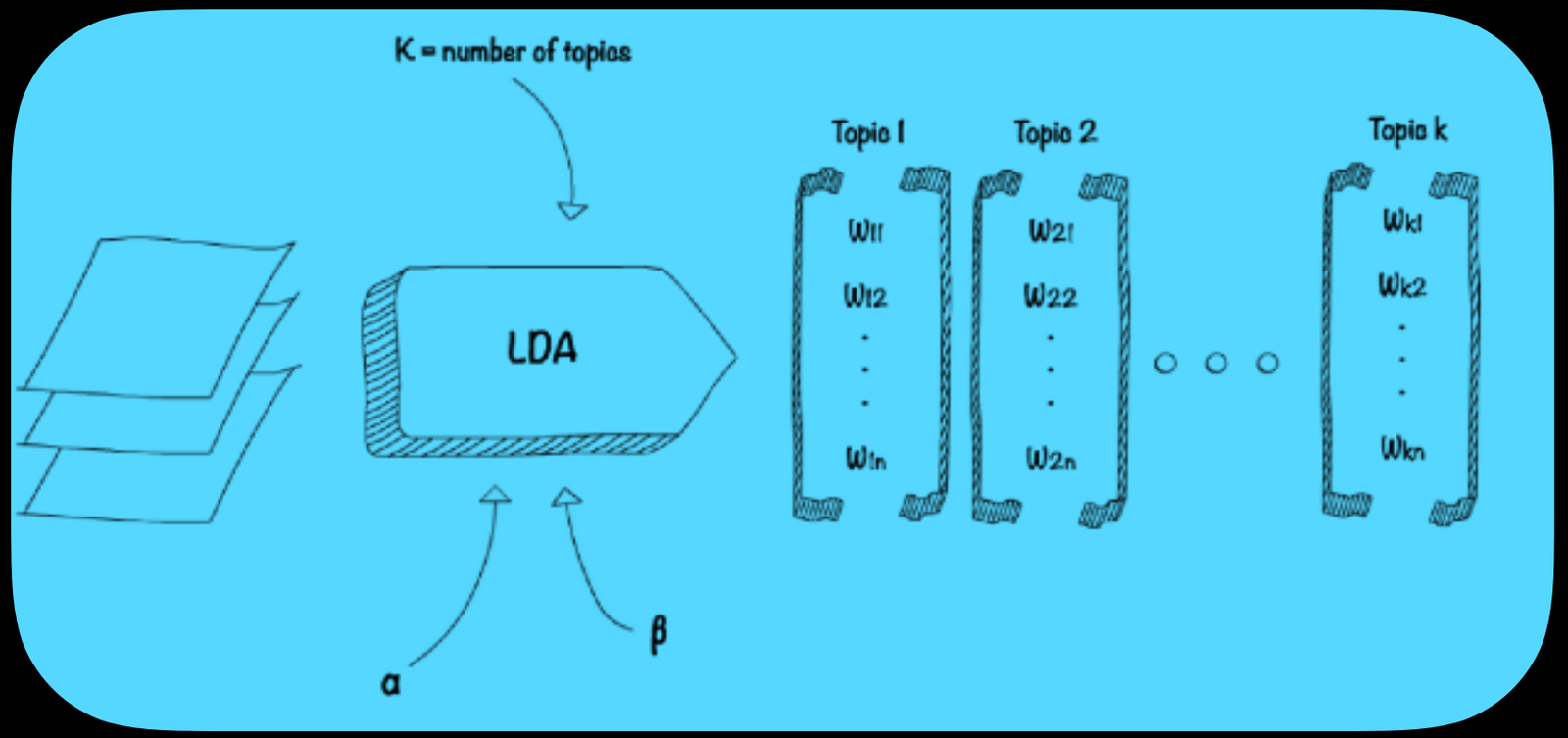
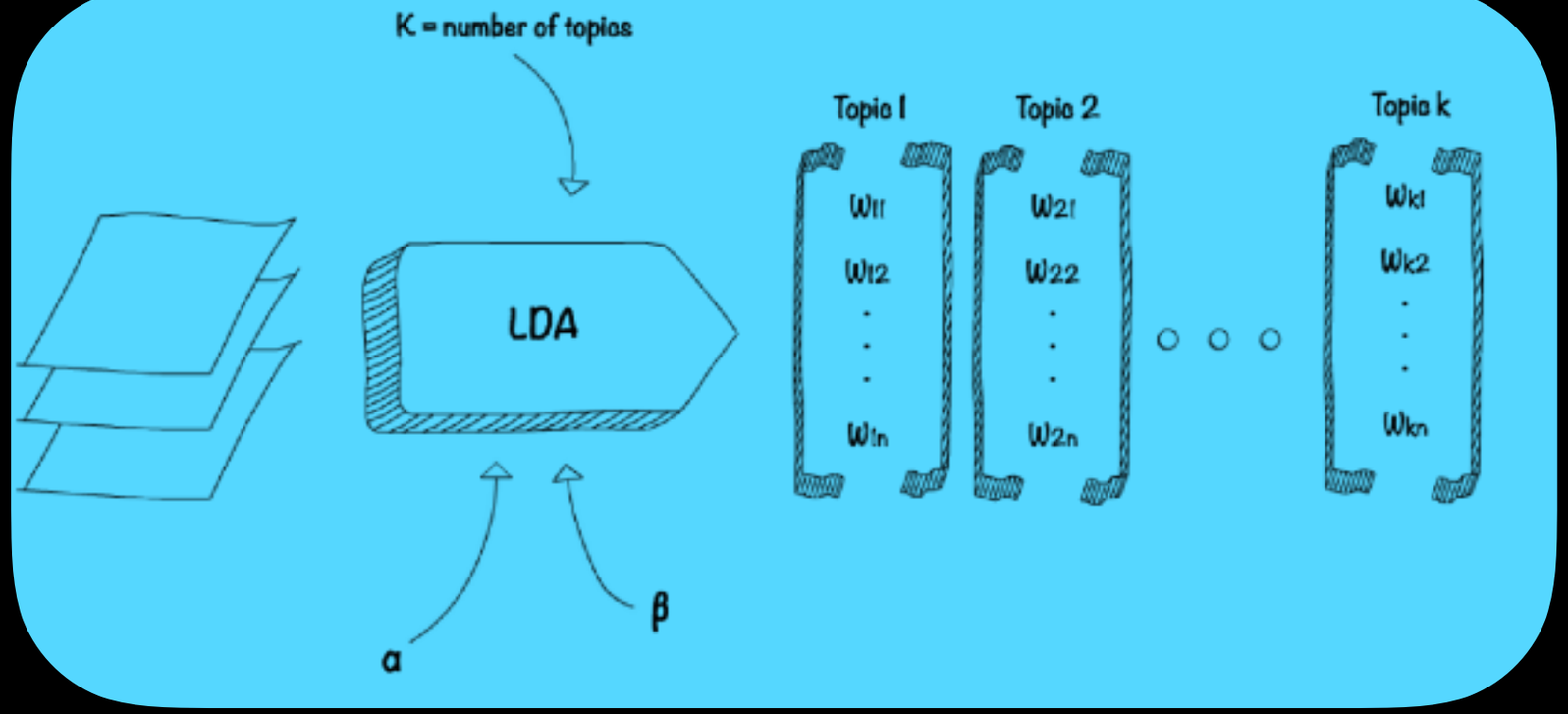


- Obviously, this random distribution is probably very wrong, as it was completely assigned at random.

- So, for each **word w** in **document m**, we assign **word w** to **a new topic** based on two things:
  - What topics are already in **document m**?
  - How many times **word w** has been assigned a **particular topic** across all of the documents: This term-to-topic distribution is called **beta**.

- We then update our **alpha** and **beta** values, but it's still probably very wrong. So, what do we do?



- We **repeat** this process a large number of times (like around 1,000), going through **every term** in **every document** in the corpus, and in each cycle **updating alpha** and **beta**.

- This is the **training process**, which works just like it did for other supervised machine learning techniques.

- However, once we get our trained model, this is just still for some value of **K**, which we don't know is the best value.

- While we don't cover this here, you'll need to repeat this entire process for various values of **K** until we get topics that are distinct from one another.

we get topics that are distinct from one another.

- While we don't cover this here, you'll need to repeat this entire process for various values of K until the best value.
- However, once we get our trained model, this is just still for some value of K, which we don't know is technique.
- This is the **training process**, which works just like it did for other supervised machine learning
- We repeat this process a large number of times (like around 1,000), during which every document in the corpus, and in each cycle **updating** alpha and beta.