**IMPLEMENT LSA & TOPIC MODEL:**

from gensim import corpora, models

from nltk.tokenize import word\_tokenize

import nltk

# Download NLTK data (only first time)

nltk.download("punkt")

# Sample documents

documents = [

"I love deep learning and natural language processing.",

"Machine learning is a part of artificial intelligence.",

"Deep learning models are widely used in computer vision.",

"Natural language processing helps machines understand human language.",

"Artificial intelligence is the future of technology."

]

# Step 1: Tokenization

texts = [word\_tokenize(doc.lower()) for doc in documents]

# Step 2: Create dictionary (word -> id mapping)

dictionary = corpora.Dictionary(texts)

# Step 3: Convert documents to Bag-of-Words (BoW) format

corpus = [dictionary.doc2bow(text) for text in texts]

# Step 4: Apply LSI model

lsi\_model = models.LsiModel(corpus, id2word=dictionary, num\_topics=2)

# Print the topics

print("=== LSI Topics ===")

for idx, topic in lsi\_model.print\_topics(num\_topics=2, num\_words=5):

print(f"Topic {idx}: {topic}")

# Step 5: Transform a new query into LSI space

query = "deep learning in AI"

query\_bow = dictionary.doc2bow(word\_tokenize(query.lower()))

query\_lsi = lsi\_model[query\_bow]

print("\n=== Query Representation in LSI Space ===")

print(list(query\_lsi))