import nltk

from nltk.corpus import gutenberg

from gensim import corpora, models

import pyLDAvis.gensim\_models as gensimvis

import pyLDAvis

nltk.download('gutenberg')

doc = gutenberg.open('carroll-alice.txt').read()

# Tokenize the text

tokens = [word.lower() for word in nltk.word\_tokenize(doc)]

# Remove stopwords

stop\_words = set(nltk.corpus.stopwords.words('english'))

filtered\_tokens = [token for token in tokens if token not in stop\_words]

# Create bag-of-words representation

dictionary = corpora.Dictionary([filtered\_tokens])

corpus = [dictionary.doc2bow(filtered\_tokens)]

# Train the LDA model

lda\_model = models.LdaModel(corpus, id2word=dictionary, num\_topics=5)

# Convert the LDA model to a format compatible with pyLDAvis

vis\_data = gensimvis.prepare(lda\_model, corpus, dictionary)

# Display the visualization

pyLDAvis.display(vis\_data)