**Social Media and Text Analytics**

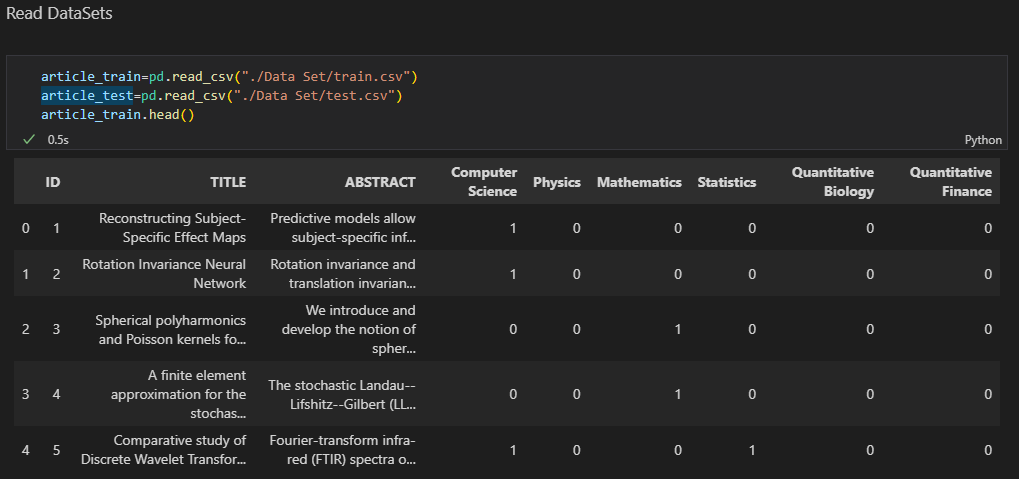
**Topic Modeling**

**Problem Statement:** Researchers have access to large online archives of scientific articles. As a consequence, finding relevant articles has become more difficult. Tagging or topic modelling provides a way to give token of identification to research articles which facilitates recommendation and search process.

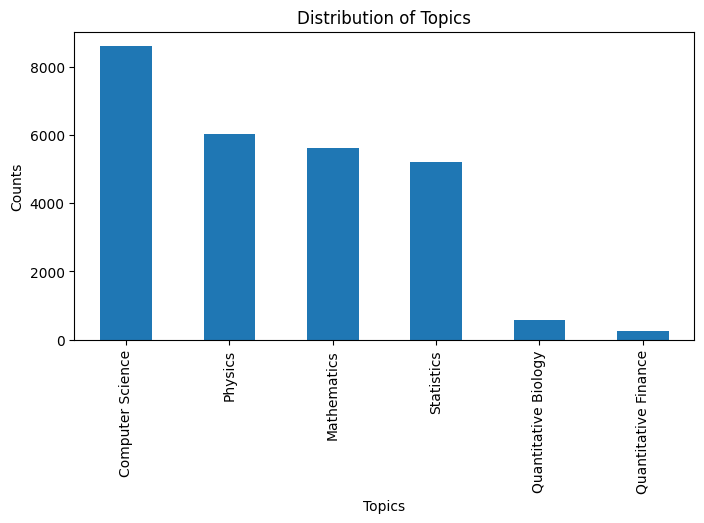
Given the abstract and title for a set of research articles, predict the topics for each article included in the test set.

**Solution:**

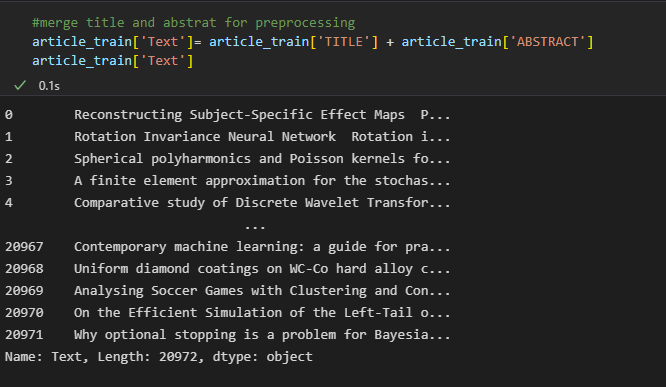
1. Install necessary libraries then load dataset

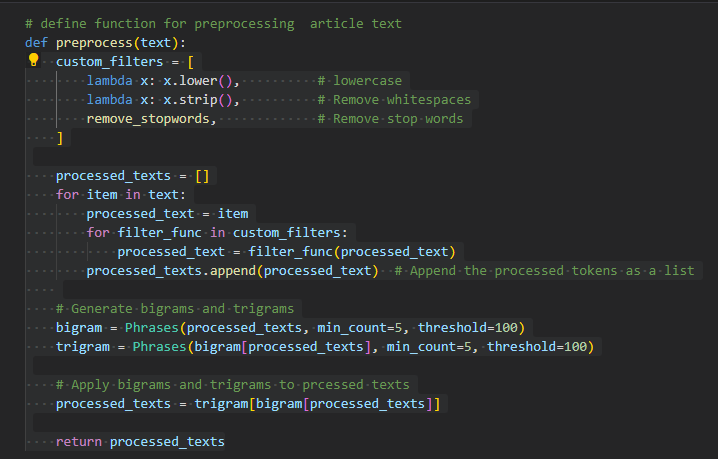


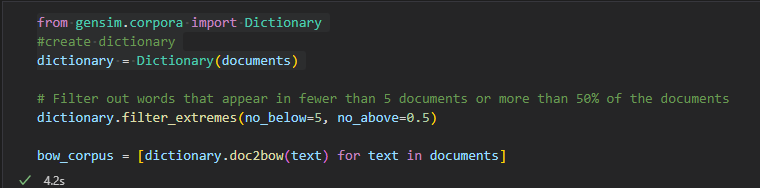
2)Visualize topic distribution in train dataset



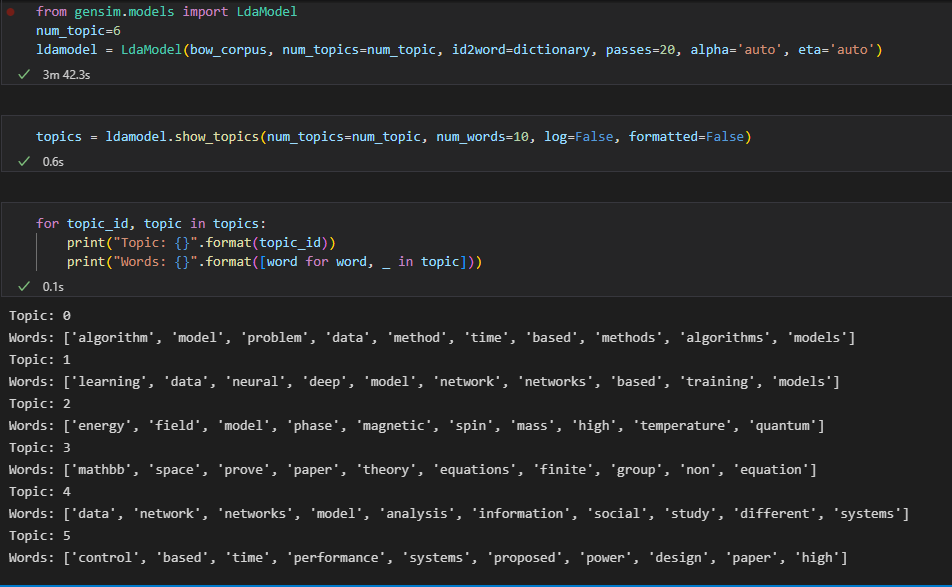
3)Merge Title and abstract

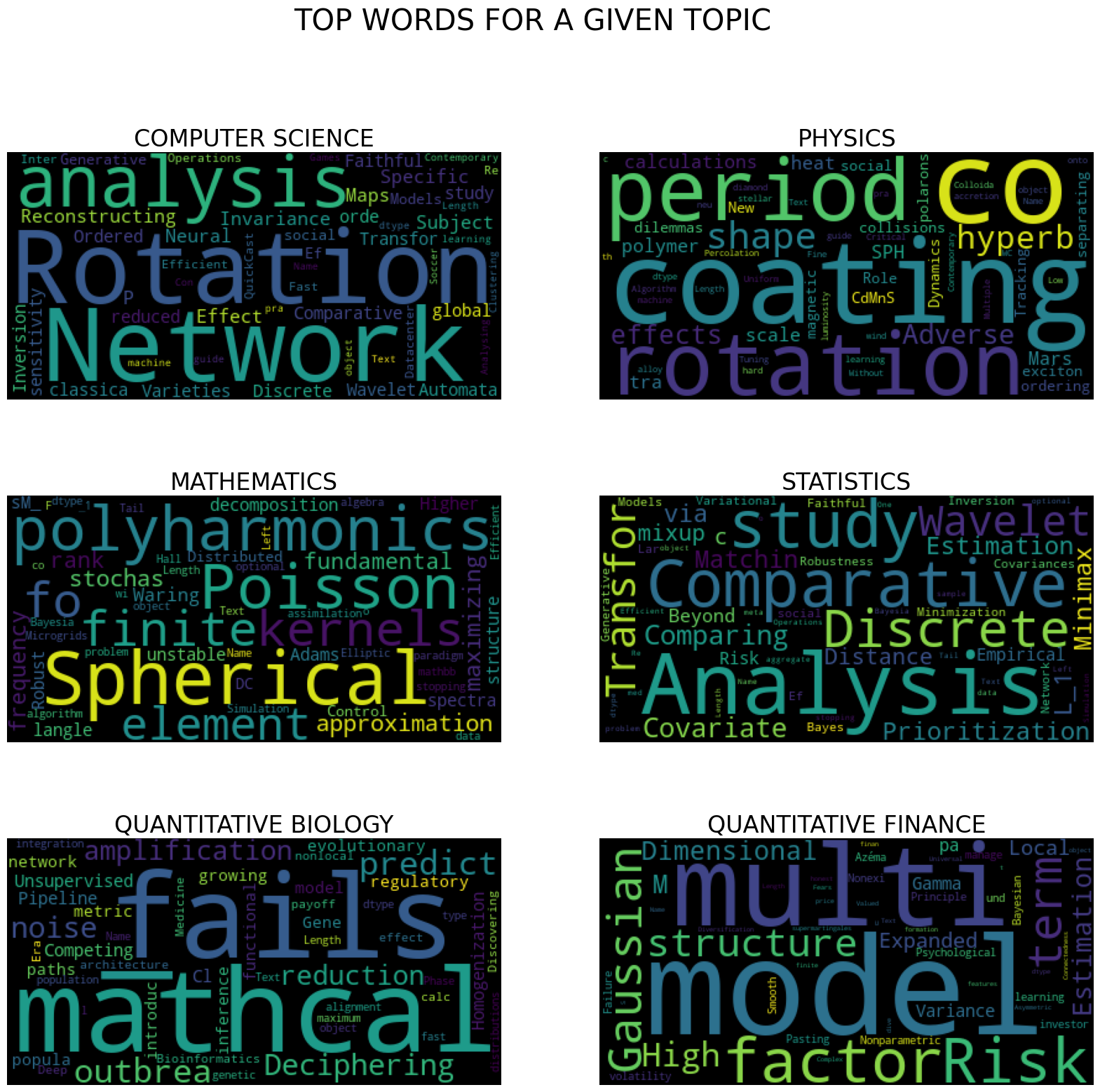


 4)Pre-processing

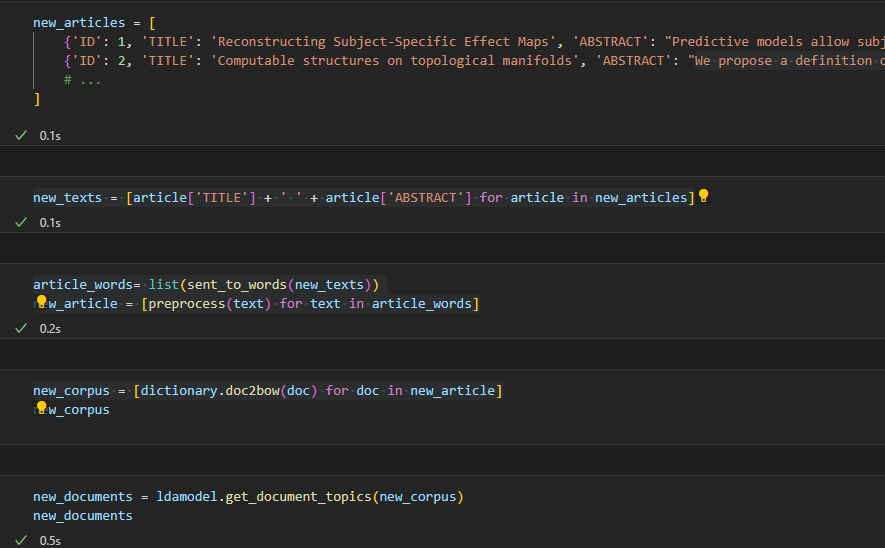
 5) Create Dictionary

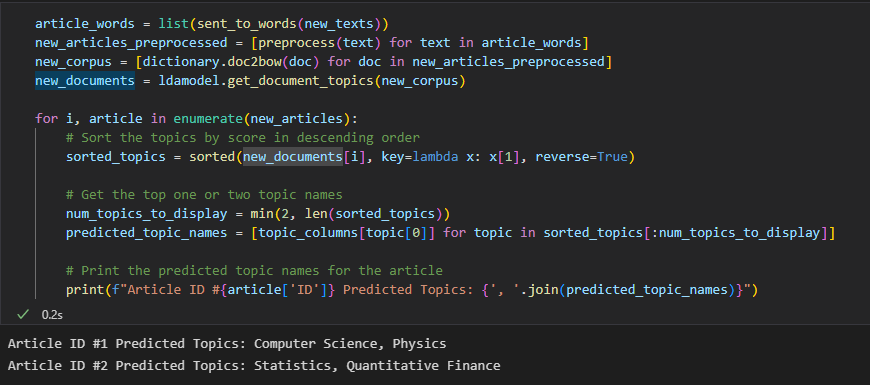
6) LDA modeling



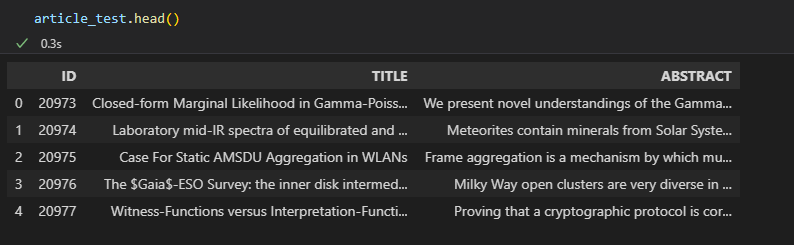
 7)Word Cloud for each topic

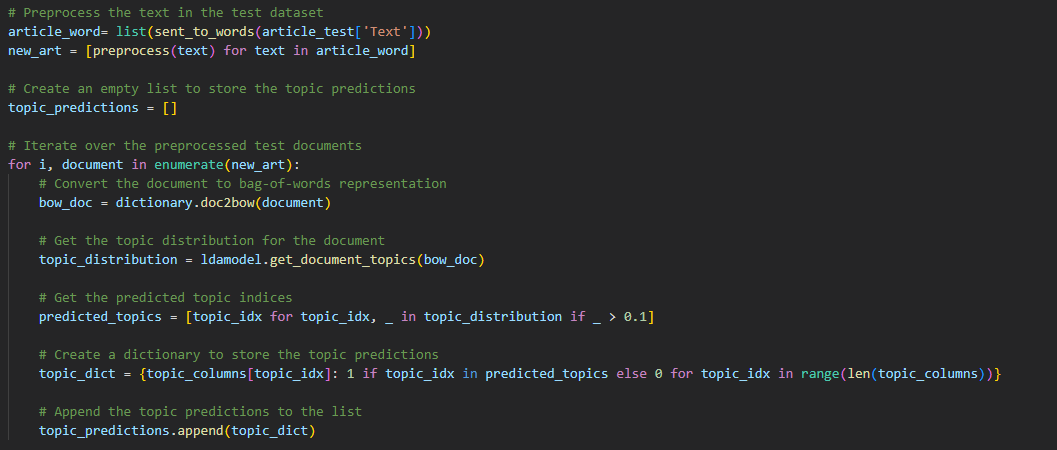
8)Manual Prediction

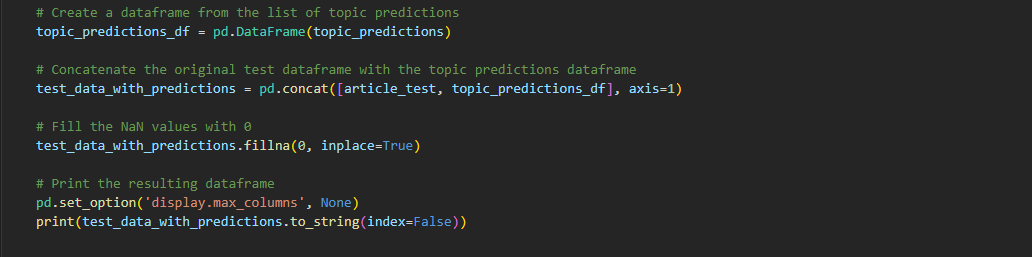




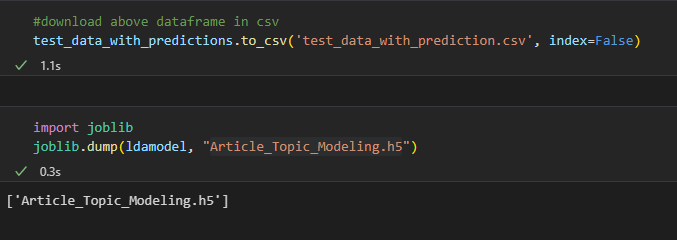
9) Test.csv file prediction



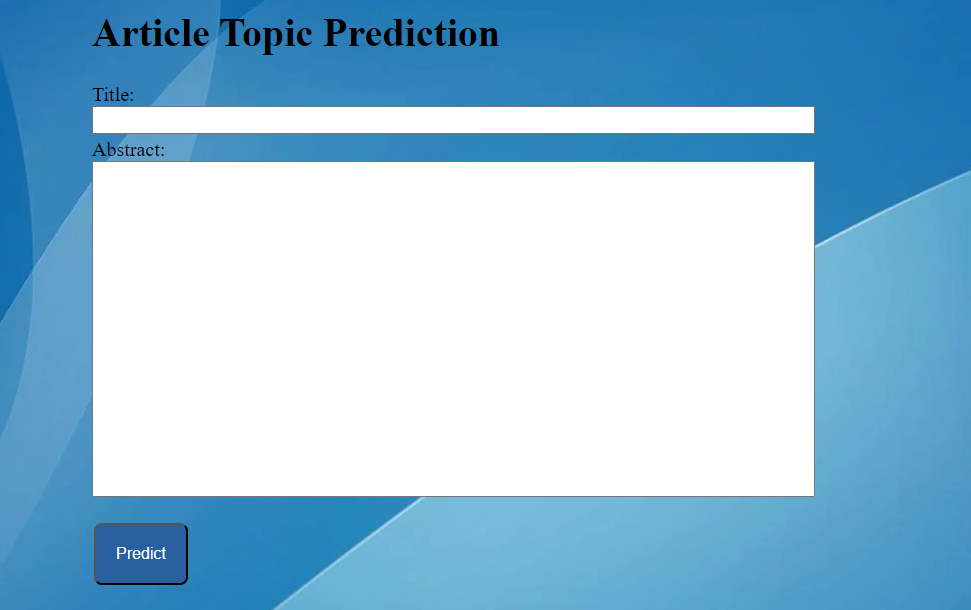




10)Text Prediction file downloaded & dump model



11) Create Flask UI for above model



12) Deploy above project on Heroku.