

README

BY PUTTING THE DATASET.TXT AND THE PLSA_50MB INTO 1 FOLDER THE CODE WOULD RUN SEAMLESSLY IF THE PLSA_50MB FILE IS RUN.

THE DATASET.TXT FILE IS THE SMALLER OF THE FILES WITH ABOUT A 200000 REVIEWS. THE ORIGINAL FILE ROUGHLY CONSISTS OF 4MILLION REVIEWS WHICH TAKES A LONG TIME TO RUN. EVEN THE FILE WITH 200000 REVIEWS TAKES ABOUT 20 MIN WHEN RUN.