File Modeling:

Nhìn chung tách các cột categorical có nhiều thành phần (author, genre, category) thành các cột mới (dạng one hot coding)

Dùng Random Forest