

ai mse-2

Book Genre Classification using Machine Learning



April 22, 2025

PROBLEM STATEMENT:

Classify books into genres using metadata such as author popularity, book length, and number of keywords.

NAME :- AMAN UPADHYAY

ROLL NO :- 202401100300032

**Course:** Introduction to AI – Mid-Semester Evaluation (MSE-2)

INTRODUCTION

Book genre classification is a useful problem in information retrieval and recommendation systems. Given metadata about books—such as how well-known the author is, the length of the book, and the number of keywords—this project aims to train a machine learning model to predict a book’s genre (e.g., mystery, fantasy, etc.).

METHODOLOGY

**Dataset: book\_genres.csv**

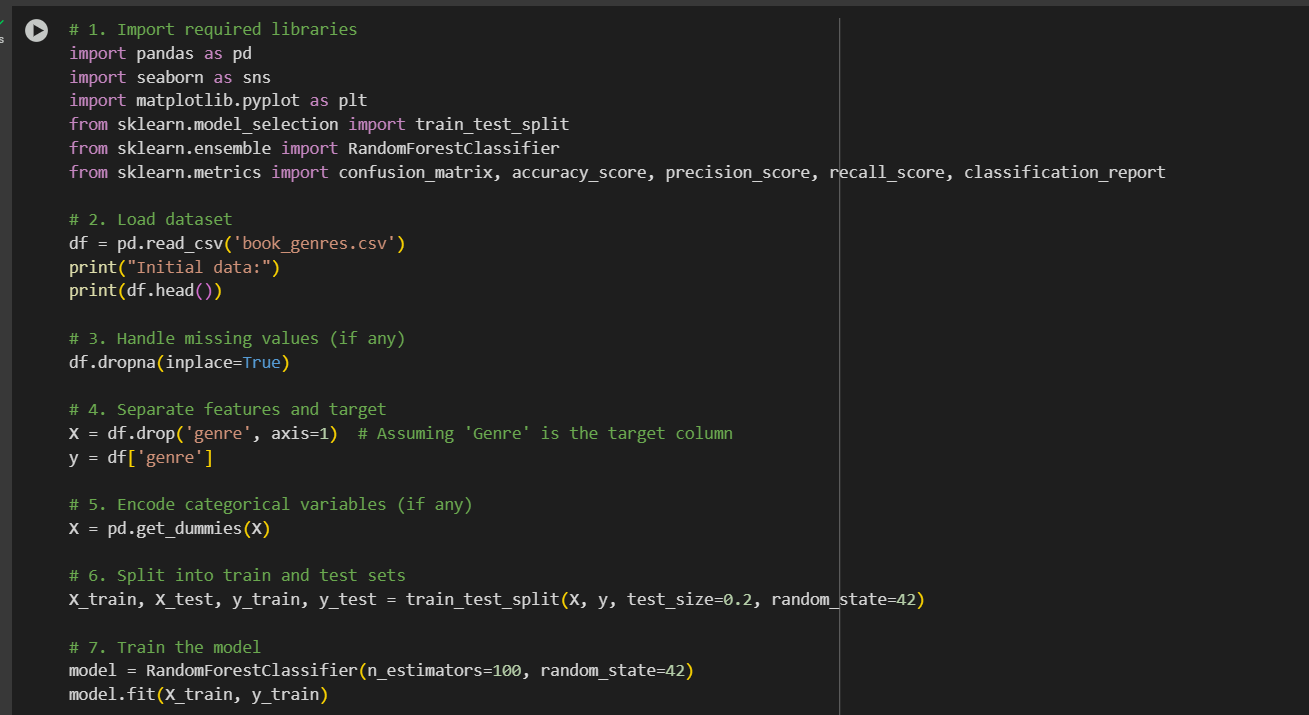
This dataset contains the following features:

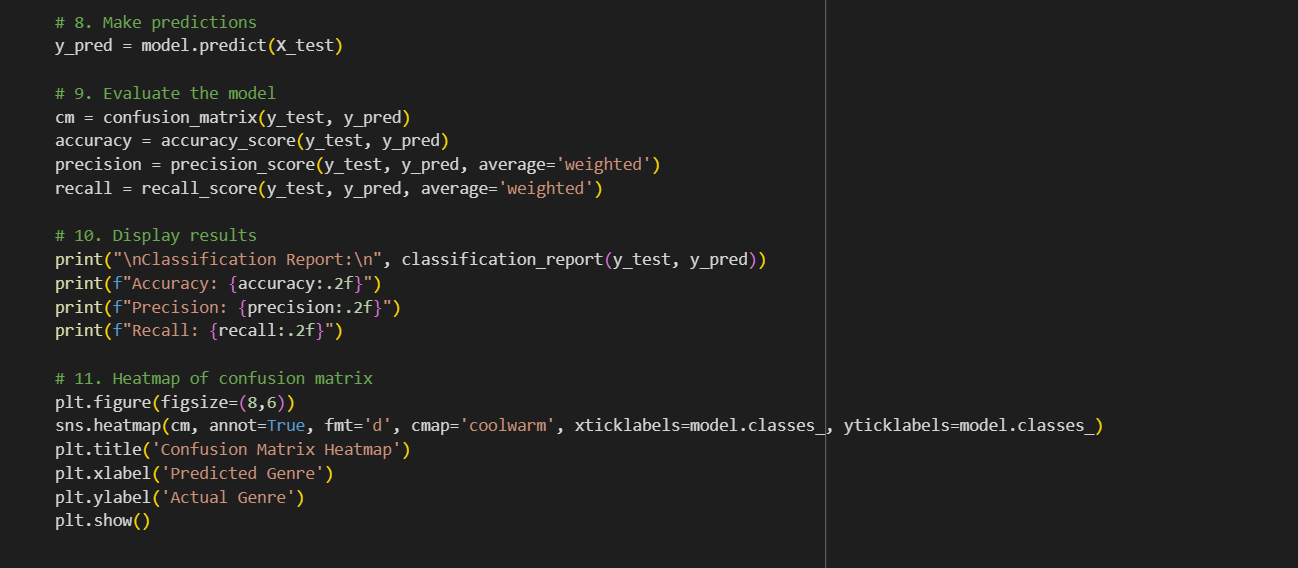
* Author\_popularity: Numerical score of author’s popularity
* Book\_length: Total number of pages
* Num\_keywords: Number of keywords associated with the book
* genre: The actual genre of the book (target variable)

**➤ Steps Taken:**

1. Loaded the dataset using Pandas.
2. Removed missing values.
3. Split data into training and test sets (80/20).
4. Trained a **Random Forest Classifier**.
5. Predicted genres on the test set.
6. Evaluated model using:
   * Confusion Matrix
   * Accuracy, Precision, and Recall
7. Visualized the confusion matrix using a heatmap with Seaborn.

CODE USED





**References**

1.pandas, matplotlib, seaborn, scikit-learn libraries

2.Mid-semester problem statement

3.Internal dataset: book\_genres.csv