

Judging a Book by its Cover

COL774 : Assignment 4

Sem I, 2022-23

Due Date: Monday November 28 (2022), 11:50 pm. No late submissions allowed. Total Points: 30 + 70

Notes:

- This assignment has **two parts**: non-competitive (30 marks) and competitive (70 marks)
- You should submit all your code (including any pre-processing scripts written by you)
- Do not submit the datasets
- **Include a write-up (pdf) file**, consolidated for both parts, which includes a brief description for each question explaining what you did. Include any observations or model details required by the question in this single write-up file
- You should use **Python** as your programming language and **PyTorch** as your deep learning framework
- Your code should have appropriate documentation for readability
- You will be graded based on what you have submitted as well as your ability to explain your code. Additionally, in competitive part, you will be graded based on your model performance relative to the class
- Refer to the Piazza for assignment submission instructions.
- This assignment should be done in groups of **in groups of 2**. If you are really keen on doing it **individually**, talk to us. You should carry out all the implementation by yourself (i.e., in your group).
- Since the assignment consists of a competitive part, no late submissions are allowed.
- For the competitive part, you will submit your class labels on a leaderboard, which will display your score and your ranking with respect to other submissions in the class.
- We plan to run **Moss** on the submissions. **We will also include solutions from the internet to maintain integrity**. Any cheating will result in a **zero on the assignment**, an additional penalty equivalent of the weight of the assignment and possibly much stricter penalties (including a **fail grade** and/or a **DISCO**).

1 Problem Statement

You are given a **dataset** consisting of book cover images and book titles. The task is to predict the book genre amongst 30 possible categories (example: Literature & Fiction, Medical Books, Romance, and Law). The task can be achieved using either the cover images, or title text, or both, in a "multimodal" fashion!

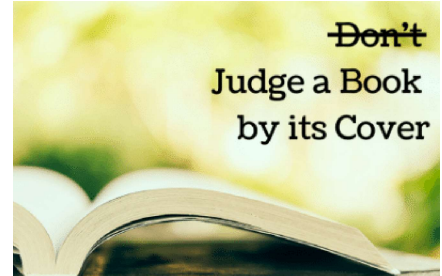


Image Source

For the non-competitive part, train your model using `train_x.csv` and `train_y.csv`, and test on `non_comp_test_x.csv`, and then report your accuracy by comparing the predicted labels with ground truth set of labels given in `non_comp_test_y.csv`. For the competitive part, you can use both the above splits for training the model, and submit predictions on `comp_test_x.csv` (target labels are not provided for this set).

2 Non-Competitive Part (30 Marks)

1. **Convolutional Neural Network (15 marks)** : For this part you have to implement a simple CNN which takes as input book cover images and returns the genre label. Implement a Convolutional Neural Network with the following structure:

- CONV1: Kernel Size $\rightarrow 5 \times 5$, Input Size $\rightarrow 3$, Output Size $\rightarrow 32$
- POOL1 : Kernel Size $\rightarrow 2 \times 2$
- CONV2 : Kernel Size $\rightarrow 5 \times 5$, Input Size $\rightarrow 32$, Output Size $\rightarrow 64$
- POOL2 : Kernel Size $\rightarrow 2 \times 2$
- CONV3 : Kernel Size $\rightarrow 5 \times 5$, Input Size $\rightarrow 64$, Output Size $\rightarrow 128$
- POOL3 : Kernel Size $\rightarrow 2 \times 2$
- FC1 : Fully Connected Layer with 128 outputs
- FC2 : Fully Connected Layer with 30 outputs

Use ReLU as the activation function for all layers apart from the Pooling layers.

Use the book cover images as input for this CNN and train the model. Finally, report the accuracy you get on the non-competitive test set.

2. **Recurrent Neural Network (15 marks)**: For this part, you have to implement a bidirectional RNN which takes as input book titles and returns the genre label. Tokenize the input text using any tokenizer of your choice and then obtain **GloVe Embeddings** for each word (specifically, glove_6b_300d). We recommend using torchtext to obtain these embeddings. The