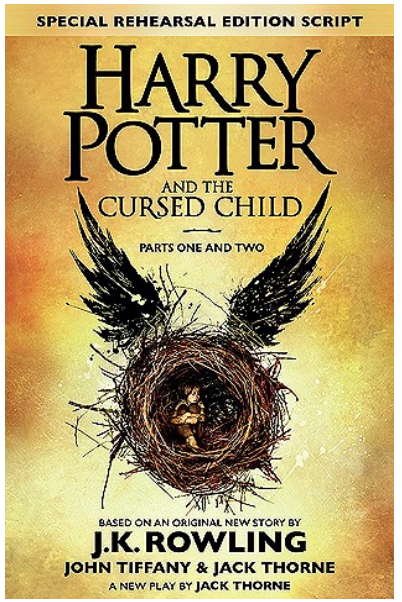
Judging a Book by Its Cover

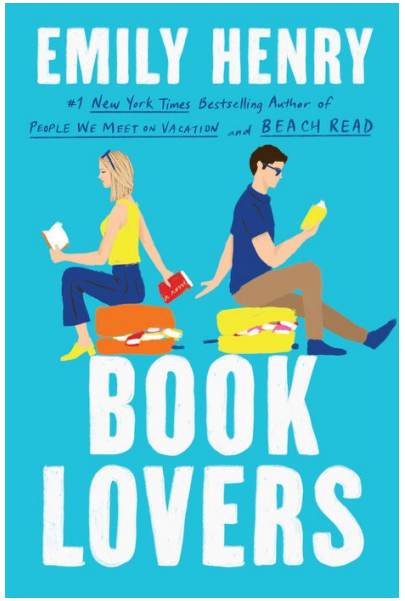
Book Genre Classification and Image Captioning

UNI: pc3019
By: Preethi Chandirasekeran

Introduction

- ➔ It is often said that one should not judge a book by its cover. However, most readers analyze the cover while forming an opinion on a book
- ➔ This project aims to leverages book cover images for genre classification and image captioning
- ➔ These models can be used to improve book recommender systems and their accessibility towards the visually impaired
- ➔ The genre classification models can be used by authors who find the genre of their book ambiguous and want to determine the genre of their book to make sure that they reach their target audience





Key Objectives

- Creating a dataset of book cover images and the genres of the corresponding books
- Building, comparing, and evaluating the performance of various deep learning models for genre classification using book cover images. Aim to achieve 40-50% accuracy on the genre classification
- Using multimodal data to perform image captioning of the book covers. Since book covers are often highly stylized, this project aims to implement an image captioning model which can identify the basic content of the book covers

Data Collection


```
def fetch_image_url_and_text(query,  wd, sleep_between_interactions=1):
    wd.get(query)
    try:
        checkquote = wd.find_element(By.ID,"quoteAuthor")
        myflag=True
    except:
        myflag=False
    while myflag==True:
        wd.get(query)
        try:
            checkquote = wd.find_element(By.ID,"quoteAuthor")
            myflag=True
        except:
            myflag=False
    my_title = None
    my_url = None
    actual_image = wd.find_element(By.CLASS_NAME,"ResponsiveImage")
    my_url = actual_image.get_attribute('src')
    actual_title = wd.find_element(By.CLASS_NAME,"Text__title1")
    my_title = actual_title.text
    return my_title, my_url
```

Snippet of Code used for web scraping Goodreads.com

Dataset Description

Genre	Number of Images
Fantasy	1,238
Horror	1,235
Memoir	1,247
Romance	1,227
Science Fiction	1,240
Self Help	1,248
Total	**7,435**

Genre Classification Results

Model	Accuracy	ROC AUC Score
Resnet-18	43.73%	0.7274
Resnet-50	19.45%	0.4903
VGG-16	45.5%	0.7778

From the above results we infer that the VGG-16 model had the best performance on the given dataset. It was followed closely by the Resnet-18 model. The ROC AUC scores align with the accuracy values. Obtaining an accuracy of around 45% is very good for the given dataset considering that there are 6 classes.

As previously mentioned in the project proposal, one of the key objectives was to perform genre classification with an accuracy between 40-50%. This objective has been successfully achieved.

Image Captioning Model

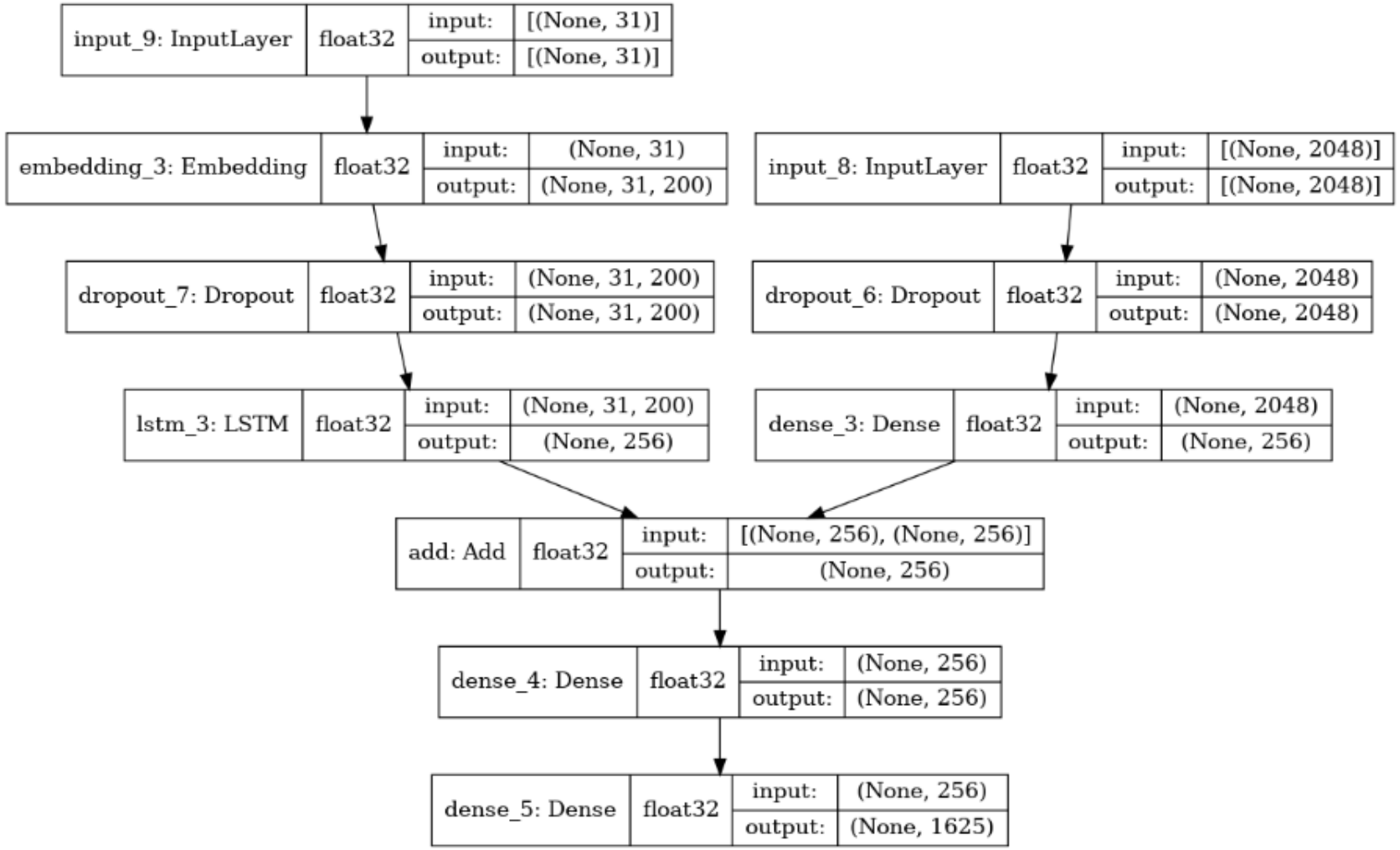
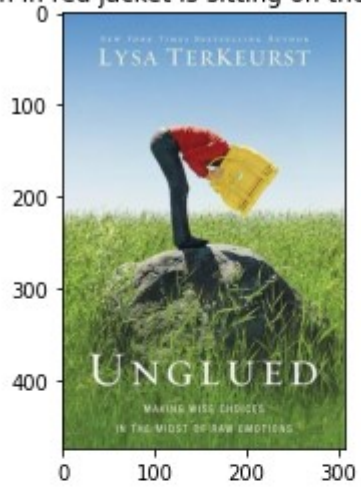
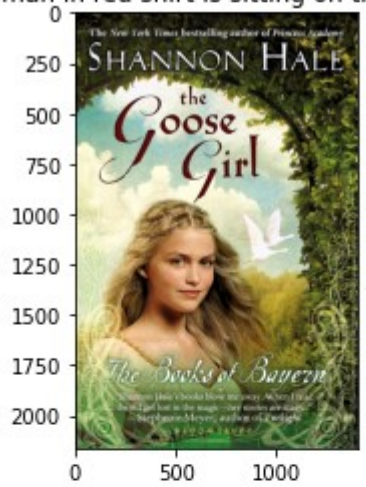


Image Captioning Results

man in red jacket is sitting on the street



woman in red shirt is sitting on the grass

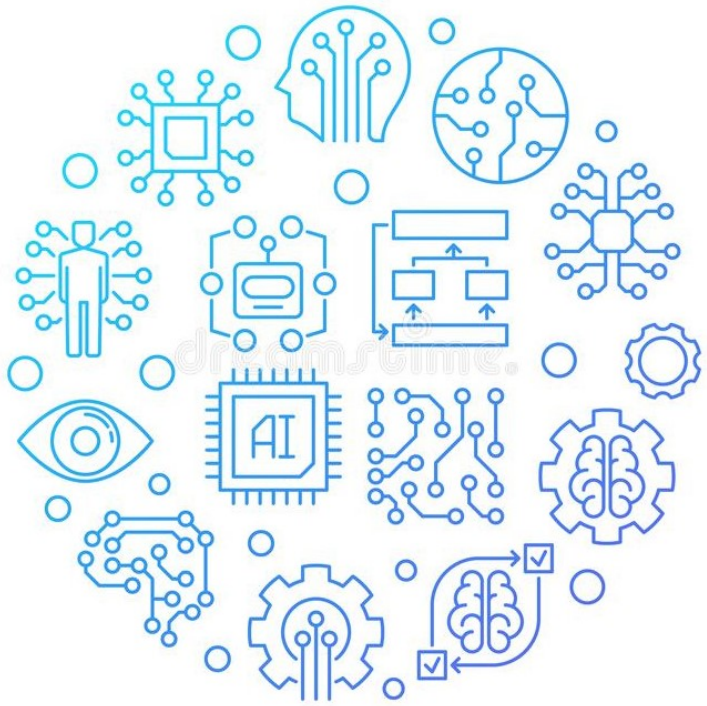


man in white shirt is sitting on the floor

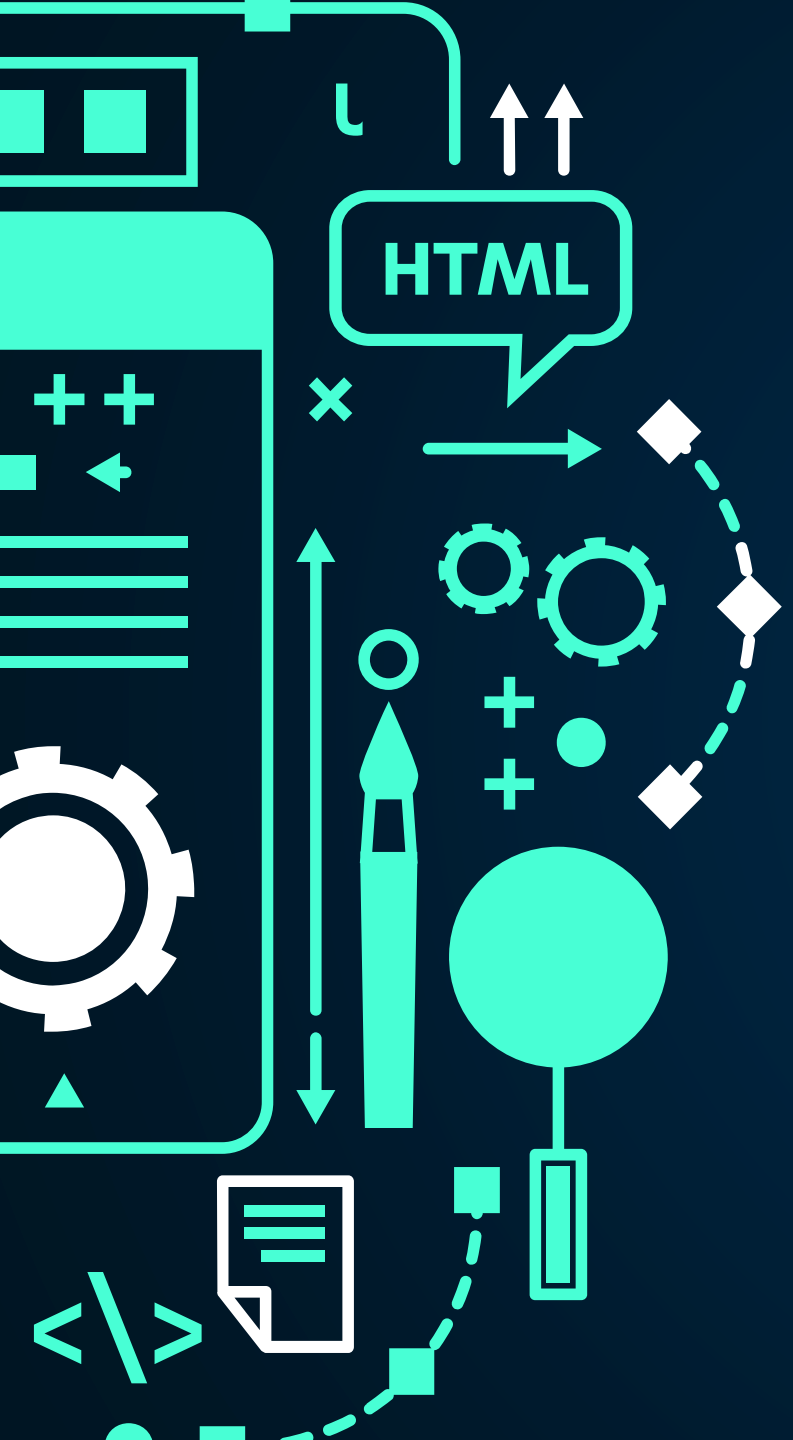
woman in red dress is sitting on the street

Conclusion and Future Scope

All of the key objectives of this project have been achieved successfully



- ☑ Exploring the use of different datasets to train the image captioning model
- ☑ Implementing and comparing the performance of different models for genre classification
- ☑ Integrating these models into a book recommendation system



THANK YOU