

BookGenics

Course name: Software Engineering (CSN – 302)

Supervisor: Prof. Mayank Gupta

Team Members (Group ID – 15):

- Shivam Raina (18103015, shivamraina.be18cse@pec.edu.in)
- Akshit Monga (18103074, akshitmonga.be18cse@pec.edu.in)
- Harshit Mittal (18103113, harshitmittal.be18cse@pec.edu.in)
- Neetish Kumar (18103009, neetishkumar.be18cse@pec.edu.in)

Project Introduction and Relevance:

The goal of the project is constructing a website/webapp which will have a user-friendly interface capable of predicting genres of short and long stories from their pdf/docx/json files. It is difficult for humans to classify the book based off its title or cover and quite tedious to manually read the whole book. So, this project is an attempt to provide the solution of the above problem. We didn't find any particular organized work regarding the same, so, we decided to inculcate this in our project.

Working/ Technology:

We plan to use NLP to achieve this. We will use a supervised classification model based on the concepts of topic modelling and topic classification. The concepts of LDA, LSA and multi label classification will be used to classify the text. The technology stack that will be used is MERN stack, we are aimed at making the website using the cutting-edge technologies like React.js, Node.js, Express.js, Tensorflow.js, Scikit Learn, etc. The languages used will most probably be JavaScript and Python. The database used will be mongo DB as we have to store json formats also.

Features:

It will contain the following features:

There will be some features that will be free and will require no login/signup, while some more features and the extension of free features will be available only after login/signup.

For users to use free features, he/she will just need to open the website.

Free Features:

- Free users can search for any book which is already present in database and check the genre of that book.

- User can also filter the books in database according to his genre interests.

For users to use premium features, they will have to Login or Sign up (if he doesn't have an existing account)

Premium Features:

- User will have access to basic features. Along with that there will be some extra features.
- User will be able to give the book pdf/doc as an input and the backend will predict its genre(s).
- User can give the summary of book as input and get genre of the book based off the summary as output.
- There will also be a feedback system for both type of Users to ensure the Web-Application is working properly and what more features could be added according to the requirements of the users.

Scaling:

A payments system can be setup in the future for premium usage.

Also, the input type will be extended to json/bson format files of the book.

Free Features:

- Free users can search for any book which is already present in database and read limited summary (if present).

Premium Features:

- User can give the title of book and get the summary (if present in database).
- User can give book as input and will get the summary of that book.
- User can also apply multi-filtering and multi-classify the books.
- Users can add books and their summary to the database.
- Based on the searches of the user, he will also be recommended books of similar genres.

Challenges:

The future challenges that can occur are: -

1. Database managing can be difficult.
2. Caching the search pattern of user.
3. Adding books by the user.
4. Integrating many technologies together.