

## UE17CS490B - Capstone Project Phase - 2

### **SEMESTER - VIII**

### **END SEMESTER ASSESSMENT**

Project Title : Automatic generation of tutorials from a text file

Project ID : PW21NSK01

Project Guide : Prof. NS Kumar

Project Team : Shrutiya M PES1201700160

Tejaswini A PES1201700740

Shubha M PES1201701540

Kritika Kapoor PES1201701868

# IEEE conference acceptance information

---



## IEEE paper accepted at CONIT 2021

Paper ID / Submission ID :944

We are pleased to inform you that your paper entitled "Automatic Tutorial Generation from Input Text File" has been accepted for the Oral Presentation as a full paper for the- "The IEEE International Conference on Intelligent Technologies(CONIT 2021)" Hubballi, Karnataka, India.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication.

Abstracting & Indexing (A&I) Databases of IEEE Xplore:

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach (Capstone Phase - 1)
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Abstract

---



- Automation of generating a **full fledged tutorial** from a text file.
- Creation of **generator-subscriber model** .
- Generation of an easily navigable tutorial a **topic-subtopic hierarchy**.
- Inclusion of audio **voice-over** based on topics with an additional **summarized presentation**.
- User interaction with **assessments** and **user progress** with an easy to use web interface.
- **Handling of multimedia** in the input text files in the form of images.
- **Language agnosticism** has been attempted for Kannada and Hindi.

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach (Capstone Phase - 1)
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Suggestions

---



- Complete integration of language agnosticism for the tutorial to support indian regional languages too.
- Addition of a metric to measure accuracy of the tutorial.
- Take care of edge cases in the input file.
- Finish all the deliverables on time.

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach (Capstone Phase - 1)
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Progress so far

---



- Application enhanced to generate tutorial for regional languages.
- Questions generated in indian regional languages.
- Most of the corner cases handled to generate a fault proof tutorial.
- Topic-subtopic hierarchy refactored to support better navigation.



# Outline

---



- Abstract
- Suggestions
- Progress so far
- **Summary of Requirements and design**
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Summary of Requirements and Design

---



## Input Document Specification -

### Acceptable Input

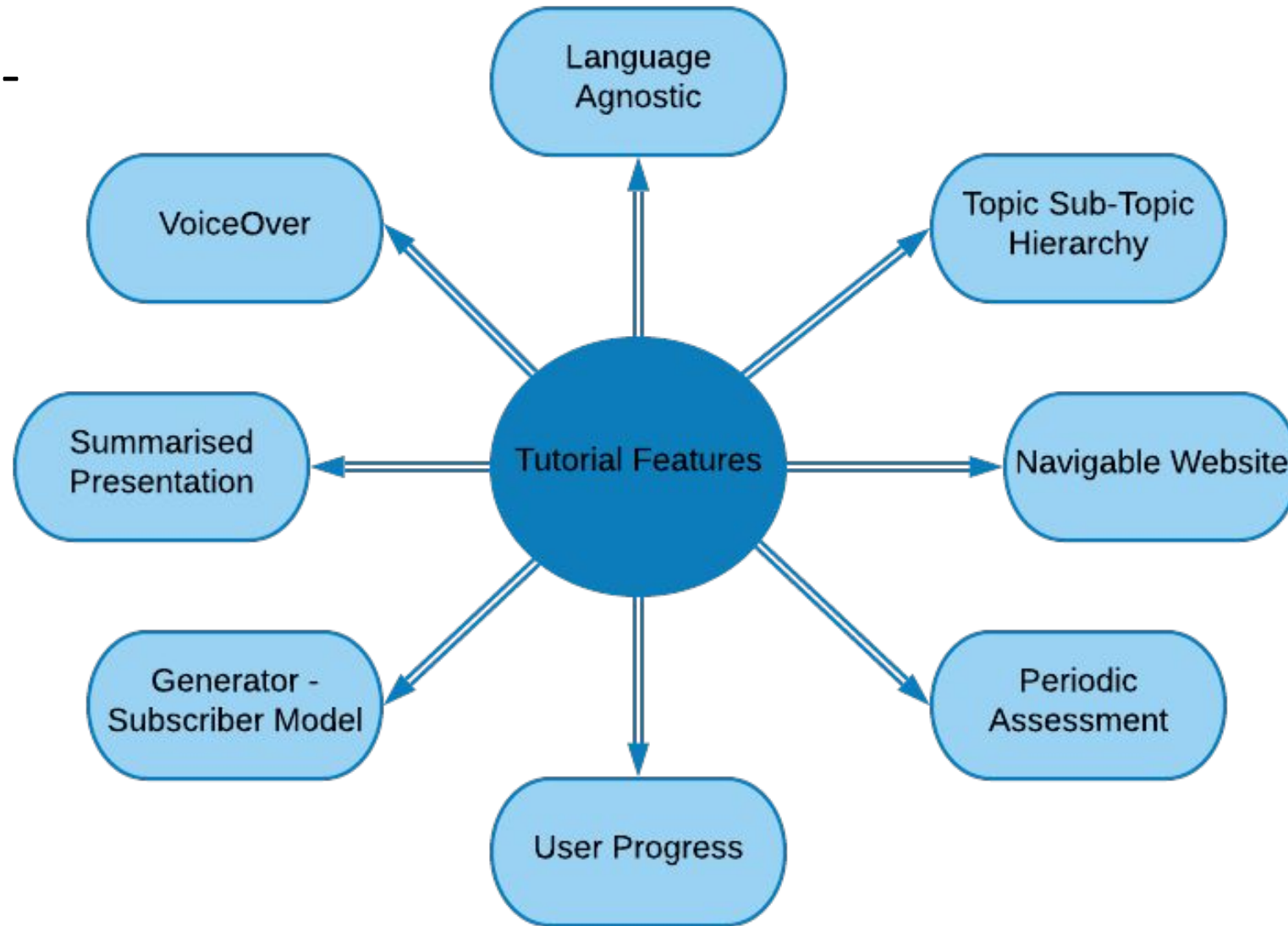
- Only PDF format documents accepted
- Proper formatting of the document is a prerequisite
- Text should be clearly formatted into a well structured document
- Should be formatted into proper hierarchy of headings according to font size
- Size of the text should be consistent
- Images should be annotated with captions titled “Figure/Fig.”.

### Non-Acceptable Inputs

- Cannot handle Mathematical Formulae
- Cannot handle Tables currently.
- Cannot handle vertical text.
- Cannot handle multi-column

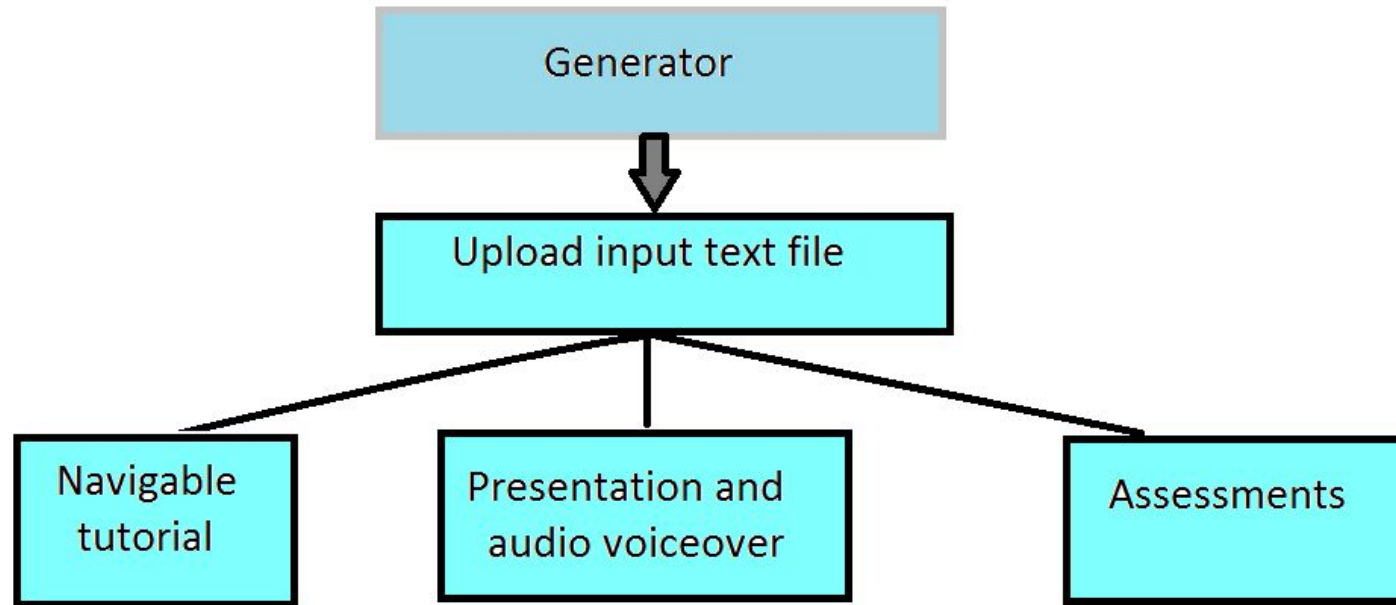
# Summary of Requirements and Design

Feature Model -



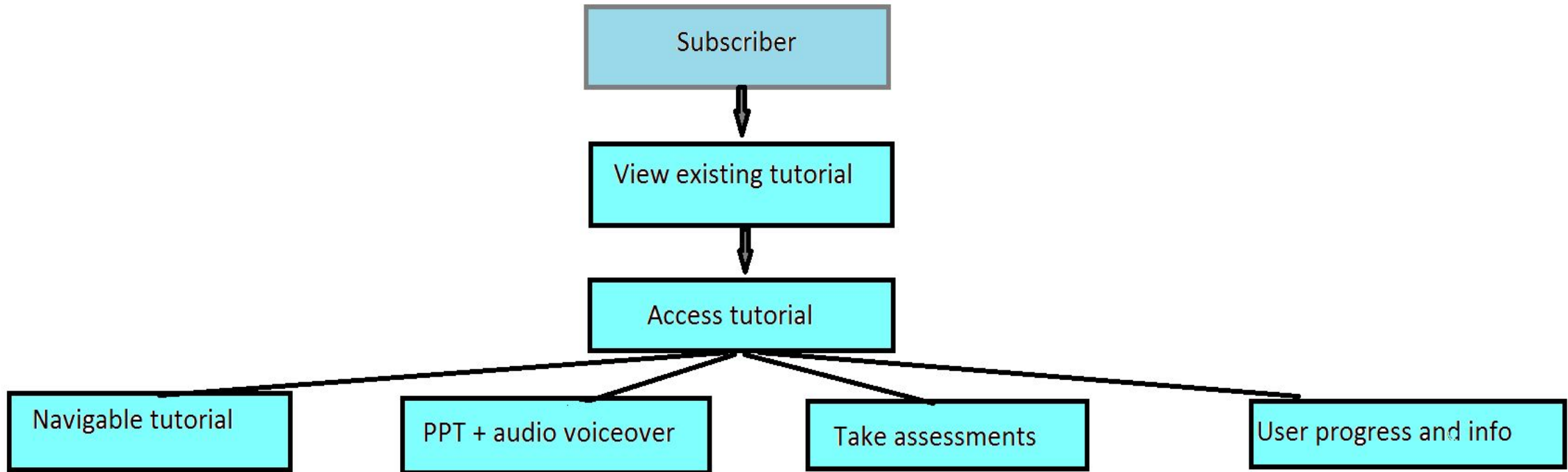
# Summary of Requirements and Design

Generator Model -



# Summary of Requirements and Design

Subscriber Model -



# Summary of Requirements and Design

---



## Other Requirements -

- **Assumption:**

A text file (.pdf) is provided by the Generator.

- **Dependencies:**

A robust system to handle the load of the application.

- **Limitations:**

- Handling different symbols in math and source code files
- Storage and scalability

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- **Team Roles and Responsibilities.**
- Summary of Methodology / Approach (Capstone Phase - 1)
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Team Roles and Responsibilities



Shrutiya M	<ol style="list-style-type: none"><li>1.Implementation of topic-subtopic hierarchy</li><li>2.Integration of frontend-backend</li><li>3.Inclusion of periodic assessments.</li><li>4.Multimedia handling.</li><li>5.Generator-subscriber module</li><li>6.User progress tracking</li></ol>
Tejaswini A	<ol style="list-style-type: none"><li>1.Extractive summarisation</li><li>2.Presentation generation</li><li>3.Audio voiceover</li><li>4.Language agnosticism</li><li>5.Question generation</li><li>6.Multimedia inclusion</li></ol>
Shubha M	<ol style="list-style-type: none"><li>1.Presentation generation</li><li>2.Frontend enhancement</li><li>3.Question generation</li><li>4.Topic subtopic hierarchy</li><li>5.Language agnosticism</li><li>6.Navigable tutorial development</li></ol>
Kritika Kapoor	<ol style="list-style-type: none"><li>1.Integration of React and flask.</li><li>2.Audio voiceover.</li><li>3.Topic subtopic hierarchy</li><li>4.User progress tracking</li><li>5.Generator-subscriber implementation</li><li>6.Assessment evaluation and formatting</li></ol>



# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- **Summary of Methodology / Approach**
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Summary of Methodology / Approach

---



- Summarisation and analysis of the given input text.
- Generation of a navigable tutorial with next and previous buttons after extracting the outline of the document using font size analysis
- Generation of short presentation based on the text summarized only to highlight the important points
- Generation of assessments after every two subtopics to ensure user interaction
- Enabling of Progress tracker.
- Explanation with audio in the presentation as an additional feature for thorough comprehension in the presentation.
- Inclusion of language agnostic features
- **Design Approach**
  - Client-server Architecture

# Outline

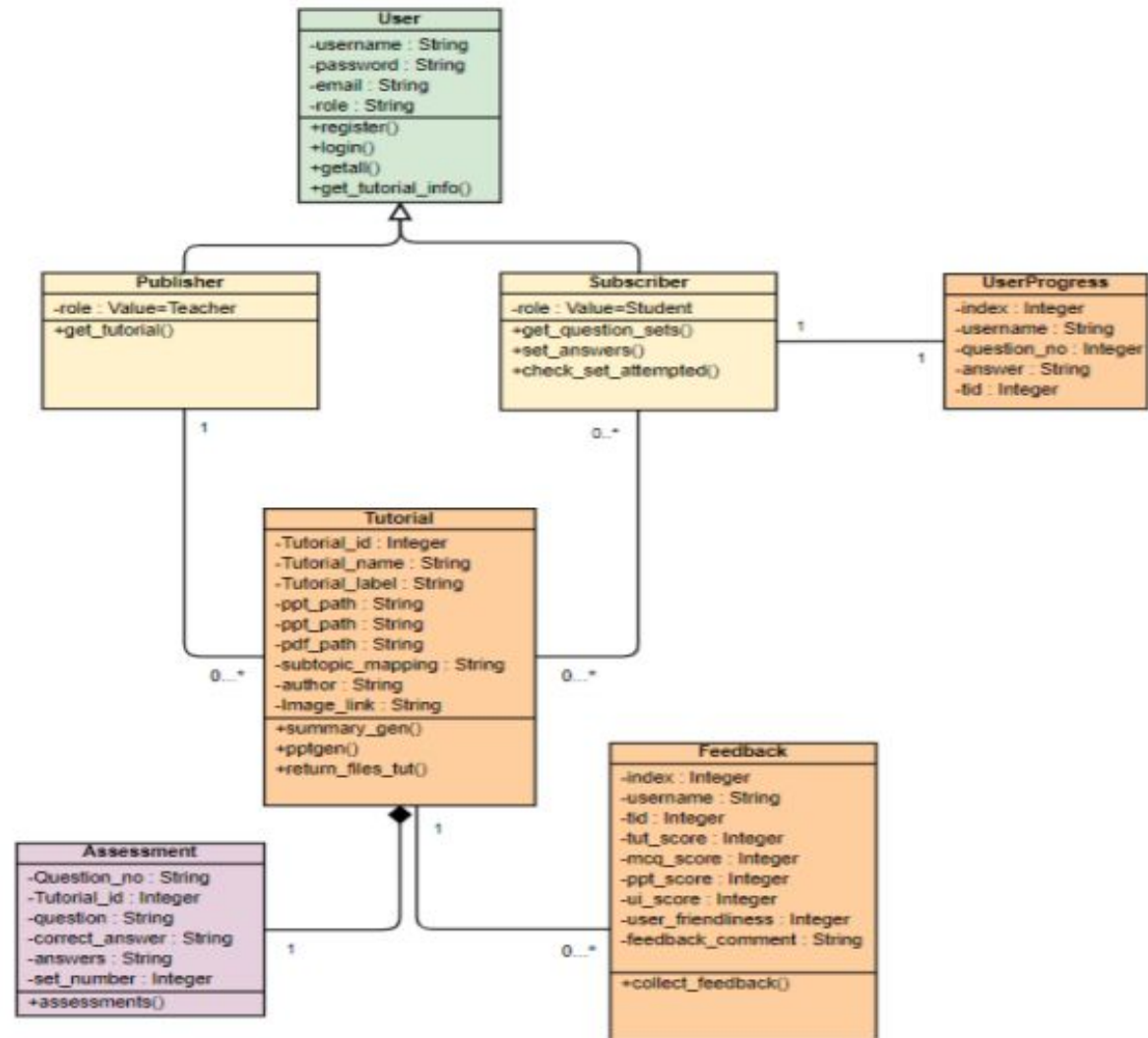
---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- **Design Description**
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

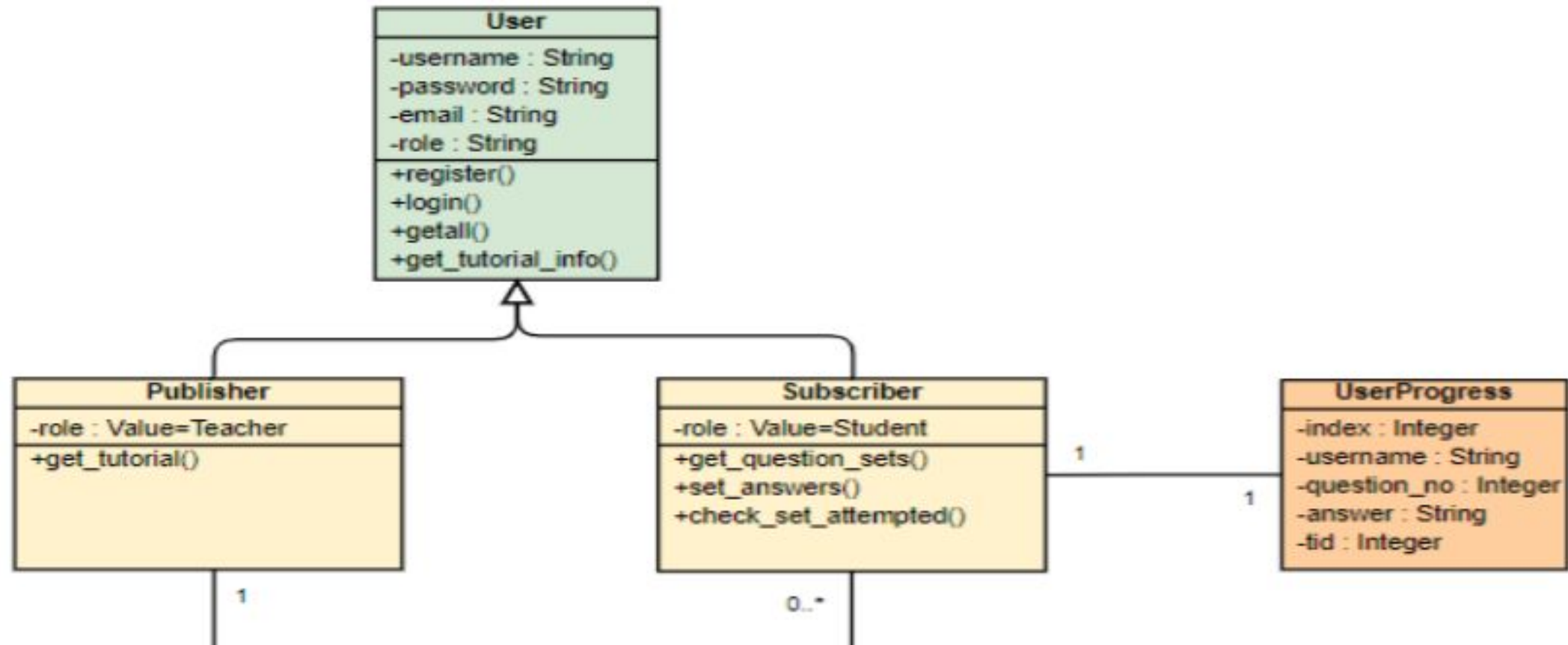
# Design Description

## Master class diagram -



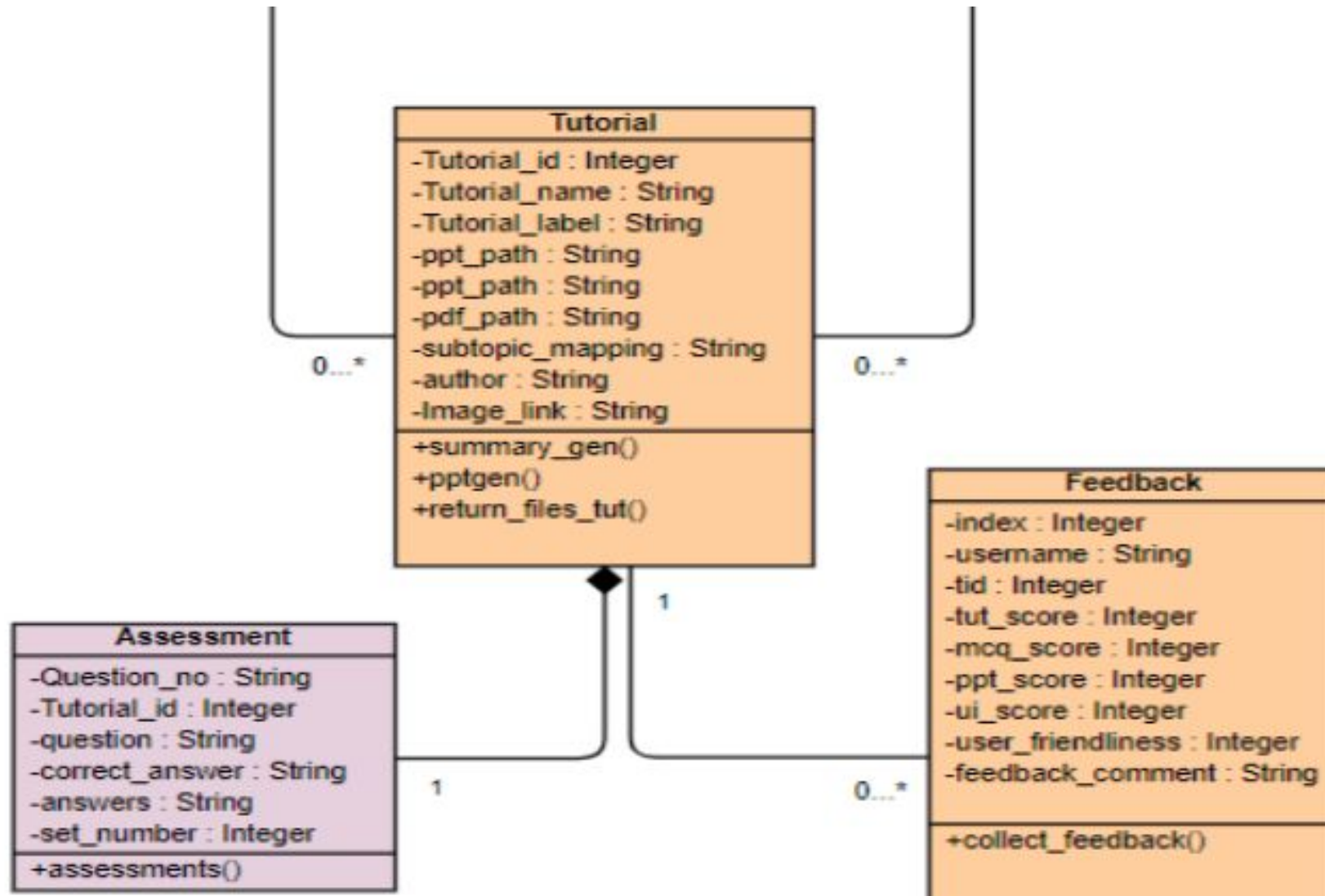
# Design Description

Master class diagram -



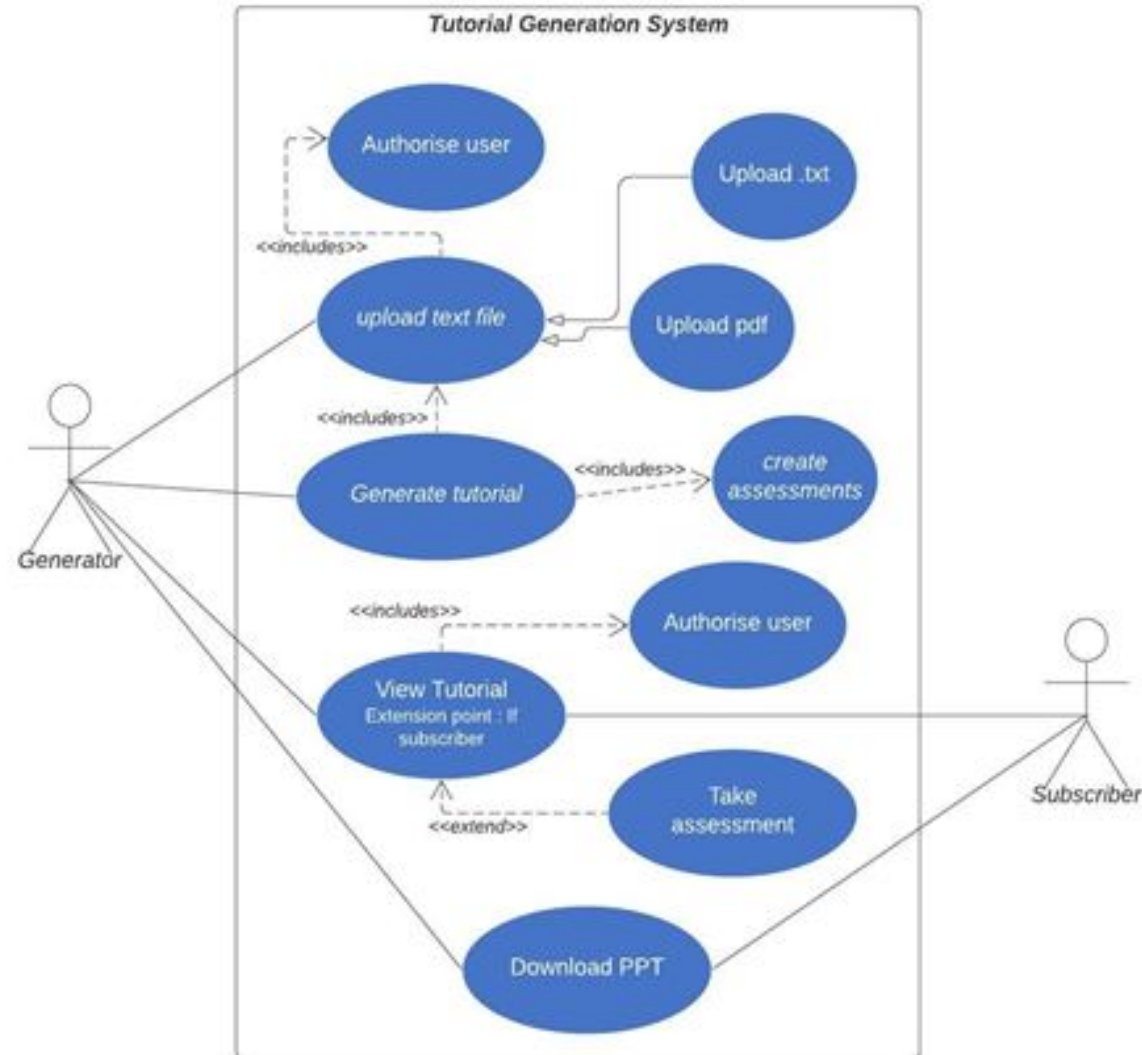
# Design Description

Master class diagram -



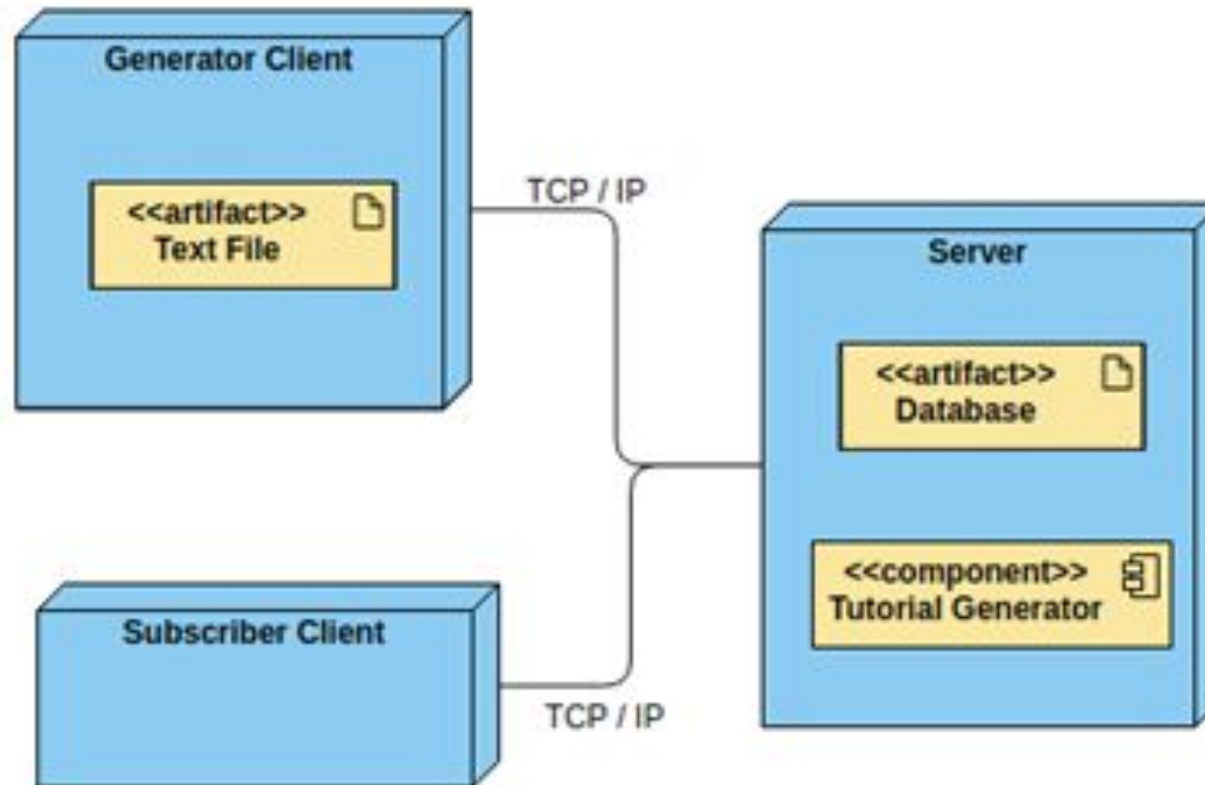
# Design Description

## Use case Diagram -



# Design Description

## Packaging and Deployment Diagram -





# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- **Modules and Implementation Details**
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Modules and Implementation Details

---



## Modules -

- **Presentation generation :**
  - Extractive text summarization through NLTK.
  - Generation of a concise presentation through Python-PPTx.
- **Topic - Subtopic hierarchy:**
  - Extraction of structure of document through Fitz.
  - Analyse the structure through extracted HTML tags.
  - Generation of a topic-subtopic hierarchy in a recursive fashion.
- **Navigable tutorial:**
  - Easily navigable tutorial through “Previous” and “Next” buttons .
  - Corresponding audio voice-over included at every stage.
  - Audio voice-over implemented through GTTS.
- **Generator - Subscriber model:**
  - Generator generates the tutorials through providing the input text file.
  - Subscriber accesses the created tutorials with the option to take up periodic assessments.

# Modules and Implementation Details

---



- **Assessment generation:**
  - Assessments in the form of MCQs generated through Questgen-AI.
  - Periodic assessments included to increase user interaction.
- **User progress tracking:**
  - Evaluation of assessments resulting in a user progress tab.
  - Consists of the tutorials completed and the corresponding progress in all the tutorials.
- **Multimedia inclusion:**
  - Generation of images based on the specific topic through web-scraping.
  - Inclusion of multimedia in the form of images in the navigable tutorial.
- **Language agnosticism:**
  - Language agnosticism attempted for Hindi and Kannada languages.
  - NLP techniques have been used for extractive summary generation,.
  - Google translate has been used to generate the questions.

# Modules and Implementation Details

---



## Technologies Used:

- **Backend-**
  - Python(NLTK, Python pptx, gTTs)
  - Flask
- **Frontend-**
  - ReactJS
  - HTML, CSS, Bootstrap
- **Database - SQLAlchemy (ORM)**

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- Modules and Implementation Details
- **Project Demonstration and Walkthrough**
- Results and Discussion
- Conclusion and Future Work
- References

## DEMO

- 
- A full fledged tutorial generation website.
  - Generator - subscriber model
  - Navigable tutorial
  - Topic-subtopic hierarchy
  - Concise presentation and audio voice-over
  - Periodic assessments
  - User progress tracking
  - Language agnosticism
-

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- **Results and Discussion**
- Conclusion and Future Work
- References

# Results and Discussion

---



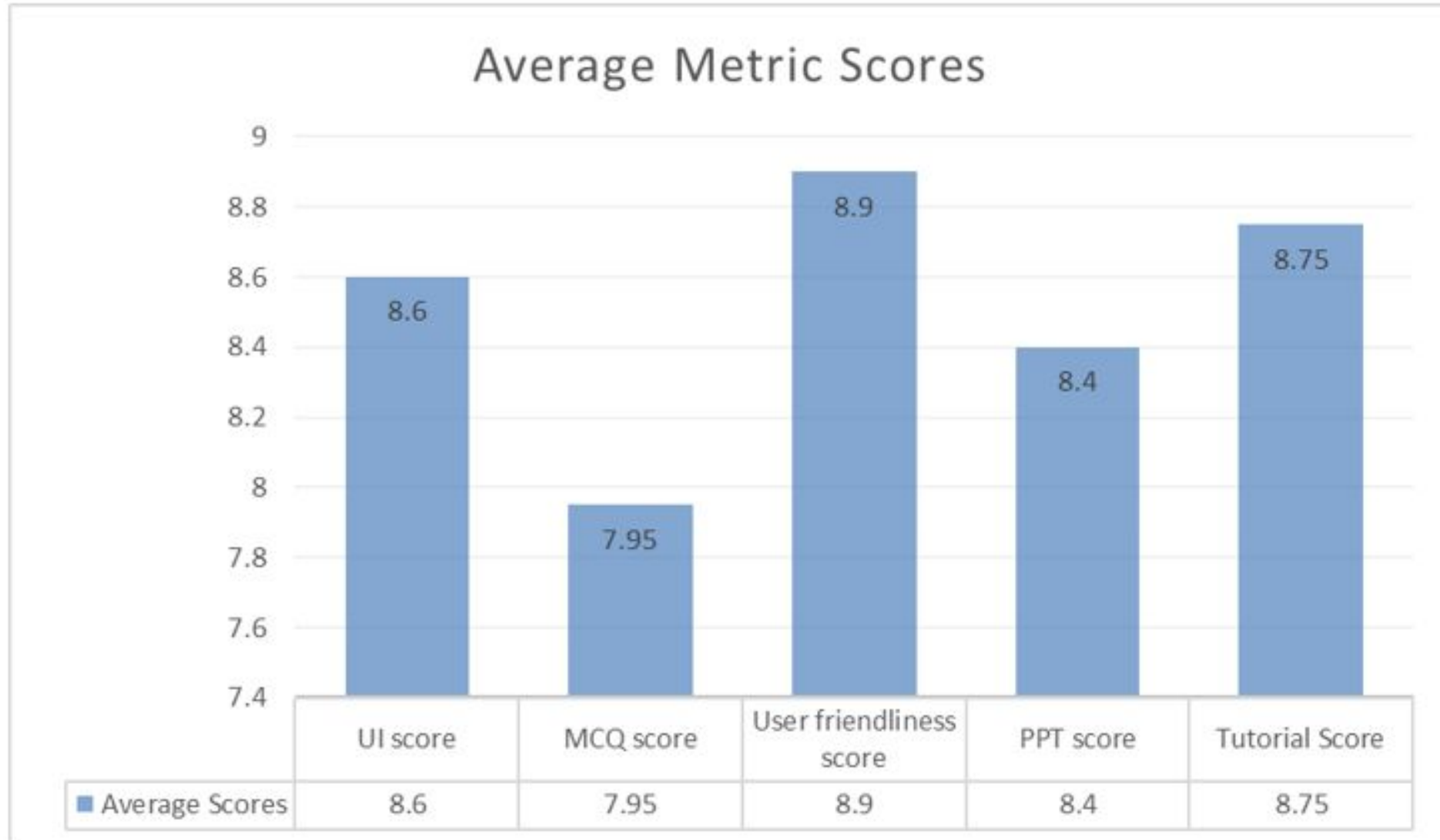
- The application was tested on multiple papers and varied pdfs.
- Pdfs following all the input restrictions successfully generated tutorials.
- Algorithm was altered to handle corner cases.
- Manual testing was carried out.
- Successful implementation of all the modules discussed.

Measures of accuracy -

- Rogue score - 0.67
- Average User feedback - 87.5%



## User Statistics



# Documentation

---



Show the evidences, status of the below documents:

- Project report

- IEEE Paper

- Github link -

<https://github.com/KritikaKapoor13/Tutorial-generation-from-text-file>

- Poster

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# Conclusion and Future work

---



## **Summary -**

Successfully built an end-to-end application where the Generator can provide an input text file and a full-fledged tutorial with assessments and progress tracking options is created and an end user, a subscriber can take advantage of the created tutorial to learn the topic.

## **Future Work -**

This project will be further extended to include features involving user recommendations. Handling of code segments, tables, mathematical formulae will also be attempted later.

# Outline

---



- Abstract
- Suggestions
- Progress so far
- Summary of Requirements and design
- Team Roles and Responsibilities.
- Summary of Methodology / Approach
- Design Description
- Modules and Implementation Details
- Project Demonstration and Walkthrough
- Results and Discussion
- Conclusion and Future Work
- References

# References

---



1. Datta, A., Ramabhadran, B., Emond, J., Kannan, A. and Roark, B., 2020, May. Language-agnostic multilingual modeling. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8239-8243). IEEE.
2. Sunitha, C., Jaya, A. and Ganesh, A., 2016. A study on abstractive summarization techniques in Indian languages. Procedia Computer Science, 87, pp.25-31.
3. Li, Y., Min, M., Shen, D., Carlson, D. and Carin, L., 2018, April. Video generation from text. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).
4. Kim, D., Joo, D. and Kim, J., 2020. TiVGAN: Text to Image to Video Generation With Step-by-Step Evolutionary Generator. IEEE Access, 8, pp.153113-153122.
5. Obeidat, R., Duwairi, R. and Al-Aiad, A., 2019, August. A collaborative recommendation system for online courses recommendations. In 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML) (pp. 49-54). IEEE.
6. Gulzar, Z., Leema, A.A. and Deepak, G., 2018. Pcrs: Personalized course recommender system based on hybrid approach. Procedia Computer Science, 125, pp.518-524.

**Thank  
You**