**MOOCs RECOMMENDER BASED ON LEARNING STYLES**

Project ID: 18-036

Project Proposal Report

Saugat Aryal

Porawagama A.S

40 mm

20 mm

20 mm

25 mm

Hasith M.G.S

Thoradeniya S.D

Bachelor of Science Special (Honors) in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

April 2018

**MOOCs RECOMMENDER BASED ON LEARNING STYLES**

Project ID: 18-036

Project Proposal Report

(Proposal documentation submitted in partial fulfilment of the requirement for the Degree of Bachelor of Science Special (honors)

40 mm

20 mm

20 mm

25 mm

In Information Technology)

Bachelor of Science Special (Honors) in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

April 2018

**DECLARATION**

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |
| --- | --- | --- |
| Name | Student ID | Signature |
| Saugat Aryal | IT 14 146602 |  |
| Porawagama A.S | IT 14 142024 |  |
| Hasith M.G.S | IT 14 140280 |  |
| Thoradeniya S.D | IT 14 138232 |  |

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Mr. Nuwan Kodagoda Date: 2018-04-05

Signature of the co-supervisor: Ms.Kushnara Suriyawansa Date: 2018-04-05

**Abstract**

Over the years, Massive Open Online Courses (MOOCs) has evolved as a learning phenomenon providing large number of courses in different domains to a wide range of users. However, learners mostly explore through different platforms and resources before they find the material that best fits their needs. It is a known and proven fact that different learners have different learning styles. As of yet, no system has been implemented which takes the learner’s style into consideration and recommends the best available MOOCs to the learner. The proposal aims at describing processes of building a learner-style centric MOOCs recommender that takes his/her learning style as a valuable input. Different video lecture styles (like talking head, presentation slide with voice-over, animation etc.), available quizzes, resources etc. are considered as significant factors when mapping learning style to MOOCs. We will be focusing only on Computer Science Programming courses on edX, Coursera and Futurelearn platforms. Among various available models, Felder-Silverman Learning Style Model (FSLSM) is selected based on literature review, to identify the learning style of users. Apart from looking for preferred courses, the user can also use the system as an advanced search engine to find any course of choice. The main objective of the system is to present most suitable resources to a learner so that learning and mastering a subject becomes effortless.

**TABLE OF CONTENTS**

[1 INTRODUCTION 1](#_Toc510689333)

[1.1 Background 1](#_Toc510689334)

[1.2 Literature Survey 2](#_Toc510689335)

[1.3 Existing Products 7](#_Toc510689336)

[1.4 Research Gap 9](#_Toc510689337)

[1.5 Research Problem 9](#_Toc510689338)

[2 OBJECTIVES 11](#_Toc510689339)

[2.1 General Objectives 11](#_Toc510689340)

[2.2 Specific Objectives 11](#_Toc510689341)

[3 METHODOLOGY 12](#_Toc510689342)

[3.1 System Design 12](#_Toc510689343)

[3.2 System Description 12](#_Toc510689344)

[3.2.1 MOOCs Scraping and Crawler 13](#_Toc510689345)

[3.2.2 Topic Modelling and Transcripts Complexity 14](#_Toc510689346)

[3.2.3 Video Style Classification 15](#_Toc510689347)

[3.2.4 Learner Profile Modelling and Recommendation 16](#_Toc510689348)

[3.3 Software Development Life Cycle 18](#_Toc510689349)

[3.4 Gantt Chart 19](#_Toc510689350)

[3.5 Work Breakdown Structure 20](#_Toc510689351)

[4 Tools and Technologies 21](#_Toc510689352)

[5 DESCRIPTION OF PERSONAL AND FACILITIES 21](#_Toc510689353)

[6 Business Potential 23](#_Toc510689354)

[7 REFERENCES 24](#_Toc510689355)

**LIST OF FIGURES**

[Figure 1: Class Central 7](#_Toc510725730)

[Figure 2: MyMooc 7](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725731)

[Figure 3: MoocLab 8](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725732)

[Figure 4: Coursetalk 8](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725733)

[Figure 5: High Level System Architecture 12](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725734)

[Figure 6: Web Scraping 14](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725735)

[Figure 7: Topic Modeling 15](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725736)

[Figure 8: Frame-based Video Classification using CNN 16](#_Toc510725737)

[Figure 9: Felder and Silverman Learning Style Characterstics 17](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725738)

[Figure 10: Agile Scrum Development Life Cycle 18](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725739)

[Figure 11: Gantt Chart 19](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725740)

[Figure 12: Work Breakdown Structure 20](file:///F:\Year%204\Research\18-036_ProjectProposal.docx#_Toc510725741)

**LIST OF TABLES**

Table 1: Comparison of existing products……………………………………………………….9

Table 2: Description of Personal and Facilities………………………………………………….20

**LIST OF ABBREVATIONS**

MOOCs Massive Open Online Courses

FSLSM Felder and Silverman Learning Style Model

ILS Index of Learning Style

CNN Convolutional Neural Network

# INTRODUCTION

## Background

Massive Open Online Courses (MOOCs) have taken a giant leap in the field of e-Learning, specifically open education. Over the years, various platforms such as edX[[1]](#footnote-1), Coursera[[2]](#footnote-2), Udacity[[3]](#footnote-3), Futurelearn[[4]](#footnote-4) etc. have emerged with the intent of providing massive educational resources to its users. One of the major and fundamental component in MOOCs is the video lectures and its production style. There are standard styles, like talking head, presentation slides with voice-over, animation, screencast etc. that are most commonly used. Along with the video lectures, quizzes and learning resources are also made available to help in the learning process.

In the recent years, there has been increasing attention towards the characteristics of learners such as learning styles. Different learners have the ability to learn in different ways and hence poses their own style of learning. Because of this behavior, the learner mostly explores through different MOOCs platform to find the most optimal learning resources that best fit their needs, preferences and learning style. Until now, no system has been developed that takes learner’s learning style into consideration when recommending the best available MOOCs across different platforms.

Hence, we will be focusing on two different aspects: learning styles and MOOCs. Our sole purpose throughout the research project is to integrate both the factors together and develop an educational platform where learners from any domain can benefit from personalized learning services.

## Literature Survey

MOOCs have evolved as big players in the field of online learning and its unique characteristics makes it an effective Technology-Enhanced Learning (TEL) model in the modern era of education and technology. Lots of prominent research have been carried out considering the state-of-art and describing an overview of MOOCs [2], [3]. According to the survey from Class Central [4], until 2017, around 81 million students are registered worldwide, where 23 million were new learners that signed up only in 2017. Similarly, the total number of MOOCs surpassed 9400 contributed from over 800 universities in 2017 [4]. This exponential growth rate is promising to believe that MOOCs are a new revolution that can help people distributed all over the world to gain access to diversified quality education in a more convenient way.

Video lectures are the significant component of MOOCs and researches have been keen in improving the effectiveness of these videos. However, limited research has been carried out and more work needs to be carried out. Guo et al.[5] explores the relationship between video production style and student engagement. Videos were classified into 6 types of production styles. Another promising research by Hansch et al.[6] presented pros and cons of different video styles. One of the most extensive work is presented by Reutemann [7] which shows the comparison between different video styles in different platforms. Thorough statistics of video styles used in edX, Coursera, Futurelearn and Iversity is reported. The paper also raises an interesting argument to correlate the video styles with drop out rates.

Learning style refers to the way a learner receives and processes information [1]. Therefore, different learners have different learning styles[1]. Considering learners learning style when designing a course has been found effective and shown positive results. It has been stated that providing learners with learning materials and activities that suit their preferences and learning style makes learning easier for them[8][9]. More evidence to support the statement is provided by studies which showed students can achieve better learning outcomes and higher scores[10], and can also master the course in less time[11].

Several studies have proposed to integrate learning style into the open learning environment (MOOCs), to provide adaptive and personalized support for learning[9], [12], [13]. Other studies reveal the use of data mining and machine learning algorithms to automatically identify the learners’ learning styles. However, there has been no significant research to support the direct mapping of a specific learning style with MOOC based on characteristics such as, video lecture production style, available quizzes, resources etc.

Researches regarding MOOCs search engine and recommenders are also being carried out at a rapid rate. A recommender system using Case Based Reasoning (CBR) approach is proposed in [14]. User’s query is described by five attributes where each attribute is assigned a weight value based on user’s preference. “Courducate” is another system proposed for personalized search engine with two functionalities : multi-site search and multi-filed search [15]. Besides using BM25 ranking function, a novel ranking function is used to rank the sites upon query. A different approach is taken in [16] where the authors proposes to associate MOOCs with learning outcomes. Hence, allowing learners to discover the most suitable MOOCs for their learning objectives. [17] proposes two contributions: Using attribute and attribute value weight of resources to get specific user preferences; A new algorithm to overcome the shortcomings of the Collaborative Filtering (CF) and provide more accurate personalized recommendations on MOOCs. In the similar context, Content based and collaborative filtering recommendation approaches are used to accommodate several undergraduate characteristics when recommending MOOCs [18].

Finally, we can conclude that, there has been no system which addresses our research problem and hence a system that recommends the most suitable resources to the learner based on their learning style while at the same time, acting as an advanced search engine for MOOCs needs to be implemented.

In the following section, literature review in individual areas are presented.

* **Video Production Style Classification**

The popularity of videos is increasing exponentially because of its usage in multiple platforms. Also, with the advancement in deep learning (Convolutional Neural Networks) and computational power, research in video classification techniques and algorithms has spiked up. Initially, the standard approach to video classification included features extraction and passing it to an SVM classifier to distinguish the classes. [19] proposed CNNs as powerful class of models for video classification. The video is treated as a bag of short, fixed-sized clips. The authors also describe a multiresolution for addressing the computational efficiency. Comparison of various video classification techniques is presented in [20] and a new approach recognized as ‘Long Short Term memory Recurrent Neural Network’ is proposed to increase the performance of video classification with better time and space complexity. Two different methods: Feature Pooling and LSTM, capable of aggregating frame-level CNN outputs in video-level predictions is proposed in [21].

A similar work of classifying an educational video frame into a particular category is presented in [22]. The proposed system is composed of two components: Feature extractor and Feature processor. Feature extraction is carried out using AlexNet and the feature processing/classification is performed through SVM classifier. A given was not classified into any category rather than being incorrectly classified was the most common failure. Another limitation identified was that it classifies individual video frames whereas taking few adjacent video frames would be easier and efficient.

* **Topic Modelling and Transcript Complexity**

Discovering abstract topics in MOOCs video transcripts can further help to filter the required videos. Wide range of research have also been conducted in this area. [23] has suggested three different mechanisms for automatically detecting the topic of the text by reviewing the results and evaluation’s scores. It has been observed that the proposed three approaches are applicable to find out the appropriate topic for the single document short texts and they extract the Topic (title) that can concisely convey the message of the long text. The study concluded that the most relevant, frequent and suitable words are Nouns for suggesting the topic for a newspaper article. Similarly, the research in [24] presented an automated phrase mining framework with two novel techniques: the robust positive-only distant training and the POS-guided phrasal segmentation incorporating part-of-speech (POS) tags, for the development of an automated phrase mining frame-work AutoPhrase. Text readability is another research domain incorporated with the proposed system. [25] developed a readability measurement method that estimates readability scores for texts intended for EFL learners by examining the various linguistic features of the texts: lexical, syntactic and discourse features.

* **MOOC Scraping and Crawler**

Extracting information from three different MOOC platforms: edX, Coursera, Futurelearn and implementing a crawler to periodically look for new data is another research topic in our system. A similar research is carried out in [26] which designs a set of configurable news collection system based on web crawler, which can crawl news from target news website. It can crawl a variety of multi-source data and the crawler is customized highly. In addition, it can do corresponding processing to crawled news content in accordance with need. Another research on crawler [27] implements an incremental web crawler on the basis of the crawler architecture named “Scrapy”. It crawls the information of news on the website successfully and implements incremental crawling. But, one disadvantage of the web crawler is that it is not a general web crawler and we need write different crawling rules for different websites. A hybrid approach is proposed in [28] which employs the meta-search engines and a semantic-structure-based Web page analysis algorithm. Their design goal of the focused crawler is to collect web pages probably related to a specific topic as many as possible and reduce irrelevant pages as few as possible.

An intelligent web crawler is presented in [29] making the rule settings dynamic, at the same time, TF-IDF method is used to calculate the Web document correlation, and the automatic acquisition of data extraction rules is realized, which reduces the development cost and maintenance cost and improves the development efficiency of the crawler.

* **Learner Profile Modelling**

Various researches are being carried out to incorporate learning styles into technology enhanced learning. Personalization into e-learning has emerged as an interesting topic of discussion among the researchers. Ability to provide the right learning materials to the learners based on their preferences and interests is a big accomplishment in online learning. FSLSM model is adopted in the system as our major focus is in computer programming courses which falls under engineering education [1]. [30]provides an in-depth analysis of FSLSM based on data from the ILS questionnaire in order to get more information for a better application of learning styles in technology enhanced environment .Therefore, they associated each dimension of FSLSM with semantic groups (such as the preference for spoken language or the preference for concrete learning material), and analyzed the impact of each group for each learning style. A hybrid approach was used for detecting interesting features both from research and from application viewpoint. Algorithmic personalization is presented in [31] which involves two steps: first modeling the learning objects by matching each dimension of FSLSM with IEEE LOM meta-data elements and finally, recommending appropriate learning objects to the learners. Four weight values are calculated for each learning object. Similarly, [32] presents a system that uses Utility-Based Recommendation technique to recommend LOs based on three aspects: subject, personal preference and Learning Style. To validate the system, an experiment was conducted as well to show the degree of satisfaction of students with the recommended learning materials.

## Existing Products

1. **Class Central**

****

Figure 1: Class Central

It[33] is one of the most popular search engine and reviews site for MOOCs. It also provides various statistics or MOOC report on yearly basis. It allows filtering on factors such as: start date, language, certificate allowed, university and provider. It provides courses from 41 providers and 811 universities. Users can create account and are also recommended courses based on their activities and history.

1. **My Mooc**



Figure 2: MyMooc

My Mooc[34] is an international platform which several thousands of free courses available in French, English and Chinese. With more than 450 course providers listed, it helps to find the best courses according to the learners’ needs. As an European leader for online courses browsing, it maintains user profile, badges and leaderboard tables. It also allows filtering on basis of factors like category, duration, language, accessibility etc.

1. **MoocLab**



Figure 3: MoocLab

MoocLab[35] is a free community website which provides forums to connect about online learning and offers suggestions, help, reviews and guidance with online learning. Besides serving as a search engine for MOOCs it mainly acts as an e-Learning “hub” where people could interact about online learning in a centralized place. It provides unique features of finding a study buddy and jobs with MOOC-friendly employers.

1. **Coursetalk**



Figure 4: Coursetalk

Coursetalk[36] is another platform for MOOCs search engines. It is a well-organized website with thousands of courses from more than 30 subjects to choose from. Like other systems, users can create account and get recommendations for their courses of preference. It provides feature of course tracker which allows users to keep tabs of all their courses – past, present and future all in one convenient location.

## Research Gap

Even though there are existing products in the market with similar objective, they don’t address several problems that needs to be taken into consideration. From the literature survey and research activities carried out, we discovered that there is quite a research gap that needs to be addressed. The following table shows a comparison of features between existing products and the proposed solution.

*Table 1: Comparison of existing products*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | Proposed | | |  |
| Features | Class Central | MyMooc | MoocLab | Coursetalk |  |  | solution |  |  |
|  |  |  |  |  |  |  |  | |  |
| Learner’s Learning Style | ✗ | ✗ | ✗ | ✗ |  |  | ✔ | |  |
|  |  |  |  |  |  |  |  | |  |
| Video Production Style of MOOC videos | ✗ | ✗ | ✗ | ✗ |  |  | ✔ | |  |
|  |  |  |  |  |  |  |  | |  |
| Search filter based on English accent of speaker | ✗ | ✗ | ✗ | ✗ |  |  | ✔ | |  |
|  |  |  |  |  |  |  |  |  |  |
| Search filter based on specific keywords / topics | ✗ | ✗ | ✗ | ✗ |  |  | ✔ | |  |
|  |  |  |  |  |  |  |  | |  |
| User Profile and Dashboard | ✔ | ✔ | ✔ | ✔ |  |  | ✔ | |  |
|  |  |  |  |  |  |  |  | |  |

## Research Problem

With the rapid advancement of technology, learning resources available over the web are more distributed than ever. Specifically, MOOCs (Massive Open Online Courses) are a relatively recent learning phenomenon with different platforms available like, edX, Coursera, Udacity, Futurelearn, Khan Academy etc. which provides large number of courses to a massive audience in different domains. MOOCs has transformed the era of learning since its introduction and one of the main reason behind it is openness. Learners coming from different background has access to massive resources at any point of time. It truly has brought a revolution in field of e-Learning.

As the popularity of MOOCs increases, more platforms are stepping up into the play where each of them has unique format of displaying the course information. This will bring much inconvenience to the users and there is a high demand of a unified platform where learners can explore the courses based on their preference. An advanced MOOC search engine that allows the users to filter through multiple fields and parameters unlike other current general search engine is a key necessity and of high demand.

The concept of learning styles has acquired a great attention and influence within the educational field over the recent years. There is a thriving industry devoted to provide the best learning resources to the learner that fits his/her learning style. Many studies have also quoted that learning and mastering a subject becomes more effortless when the right resources are provided. MOOCs search engines that are available currently, do not provide the facility to search for courses depending on their learning style. Most of the users are even unaware of their style of learning and providing them insights on it, will definitely help them in the journey of learning.

Considering the above facts, we can conclude that there is a high demand of a unified platform for recommending MOOCs based on the learning style of user and at the same time, serve as a search engine allowing the users to filter through more advanced fields.

# OBJECTIVES

## General Objectives

The proposed system is a research project focused on several objectives. Upon the completion of the project, we are supposed to fulfil these research objectives. The general objectives of the research project are mentioned as follows:

* Develop a unified platform for searching MOOCs distributed on various platforms: edX, Coursera and Futurelearn using advanced filters.
* Recommend the most suitable courses from different platforms based on learners’ learning style and domain knowledge.
* Develop user-friendly interfaces so that users can easily get what they need.
* Help users identify their learning style based on Felder and Silver Learning Style Model.
* Explore and understand the application of deep learning techniques in various scenarios.

## Specific Objectives

The specific objectives of the research project are mentioned as follows:

* Classification of MOOCs videos into a particular production style category, such as talking head, presentation slide, tutorial/demonstration (code), discussion etc.
* Analyze the complexity of transcripts of videos, feedback from comments and topic modelling to extract abstract topics that occur in transcripts.
* Extract various information from several MOOC platforms using web scrapper.
* Implement a MOOC crawler to periodically look for new courses.
* Classify the audio of videos as native or non-native English speaker.
* Correlate learning style with MOOCs.
* Query and retrieve most relevant MOOCs from database depending on search filters specified by the user.
* Maintain user profile about their learning style and domain knowledge.

# METHODOLOGY

## System Design

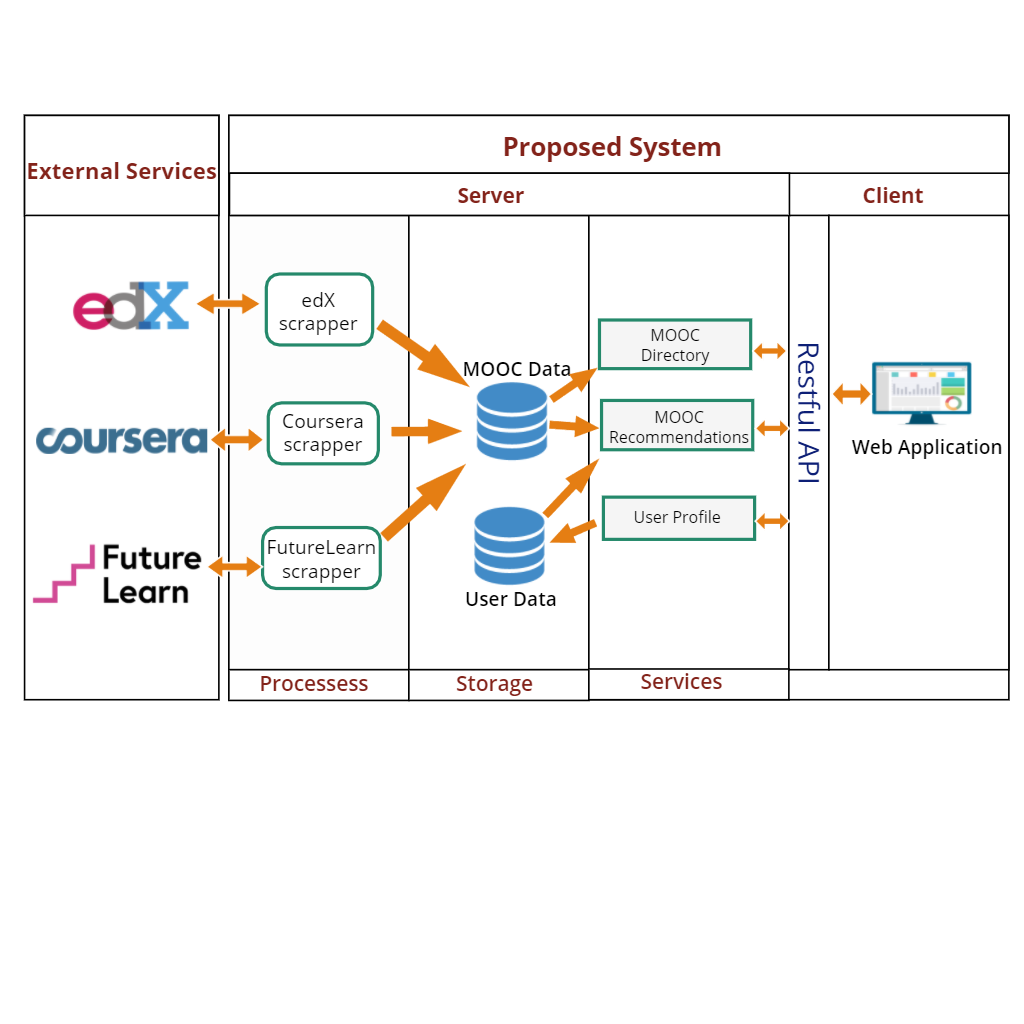


Figure 5: High Level System Architecture

## System Description

The proposed system is one of the first kind in the MOOCs domain, which first identifies the learner’s learning style. Out of several models defined, Felder and Silverman Learning Style Model (FSLSM) is adopted. Index of Learning Style (ILS) questionnaire is used as a mechanism to identify the learning styles. The FSLSM classifies learning styles into four dimensions and identifies two types of learners for each dimension[1]. The dimensions are perception, input, processing and understanding[1]. Video lectures of MOOCs from three platforms, edX, Coursera and Futurelearn are classified into various production styles using deep learning techniques. These classified video styles along with quizzes and resources available, MOOCs are labeled with a specific learning style, based on literature review. After the learning style has been identified, relevant MOOCs are suggested to the learner. Besides recommendation, the system can also be used as an advanced MOOC search engine by filtering using various parameters. Transcripts of video lectures, accent of speaker (English as First Language, FFL or English as Second Language, ESL), course description, provider details, ratings, comments etc. are also extracted, analyzed and stored which can be used for filtering MOOCs.

At the end of the project, we will be focusing on two measurable outcomes:

* **Accuracy:** The accuracy of the overall system can be measured by calculating what percentage of the MOOCs recommended by the system (considering the learning style) matches with the MOOCs that the learners would actually select in an open environment. In order to carry out the validation and measurement, experiments can be conducted taking students from same educational background and domain knowledge but with different learning styles. First, they will be freely allowed to select a specific course from any of three platforms. And finally, they will be given to use the system which also recommends a prioritized list of courses based on his/her learning styles. Hence, the course selected by the student can be validated with the course recommended by the system. With the process, an overall accuracy of the system can be measured.
* **Speed:** The overall speed of the system can also be measured by calculating how long does it take for MOOCs retrieval and recommendation. Our major goal will be to achieve a fast working system, hence accuracy might be tradeoff for speed.

As the above-mentioned outcomes, is measured for overall system, it can also be considered for each individual module of the system. The individual components of the system are described rationally as follows.

### MOOCs Scraping and Crawler

Gathering information about MOOCs from three different platforms: edX, Coursera and Futurelearn is the fundamental process in our system. We will be extracting all video lectures for different programming courses, transcripts, course description, rating, quizzes, user comments and any other resources available. Also, a MOOC crawler will be implemented to periodically can through the platforms to look for new resources. Python is by far the most popular language and provides powerful frameworks for web scraping and crawling.

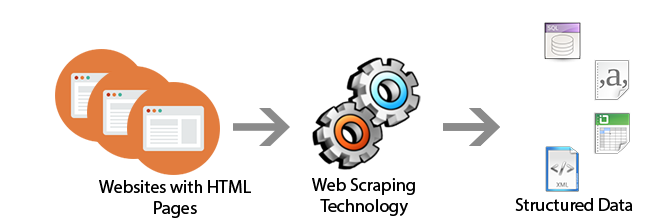


Figure 6: Web Scraping

The measurable outcomes for this component are:

* Accuracy: The accuracy for web scraping can be measured by validating the number of resources and information available in MOOC (video lectures, transcripts, course description, rating, comments, quizzes) against the information extracted by the scraping tool.
* Speed: The speed can be measured with the amount of time taken to extract/scrape information.

### Topic Modelling and Transcripts Complexity

For all the three selected MOOCs platform, transcripts are available for each video lecture. Specific topics can be extracting from the document which is useful in the scenario where the learner wants to search for MOOC with specific topics. Similarly, determining complexity of the language in transcript can be another important factor for filtering a MOOC by the learner. Phrase mining technique can be used to get short texts from the massive texts. Then for topic modeling, Nouns only approach can be used since it is the most relevant for topic suggesting as it has given a better result for the approach. Various linguistic features can be used to measure the complexity where features can be: lexical, syntactic or discourse.

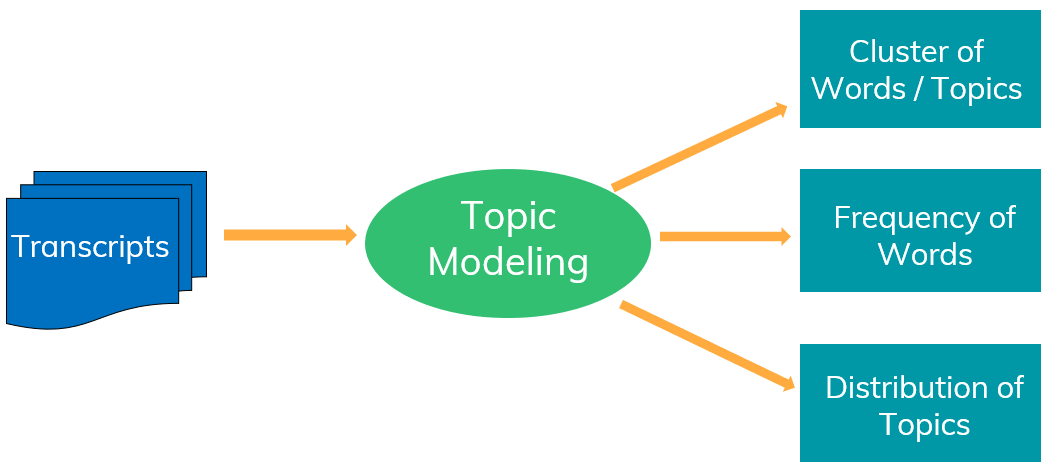
****

Figure 7: Topic Modeling

The measurable outcomes for this component are:

* Accuracy: The accuracy of topic modeling can be measured by comparing topics outputted by the topic modeling algorithm with the topics actually contained in the transcript document, which is extracted manually.
* Speed: The speed of the algorithm can be measured with the amount of time taken to show the topics in the transcripts.

### Video Style Classification

Video Style Classification into a particular category is an important component of the proposed system. Type of a video style along with other materials can be an indicator of a specific learning style characteristics. A MOOC vide is generally a combination of several video production styles. Hence, we need to process it by each frame. With the rapid advancement of deep neural networks increasing research in video classification techniques, Image-based Convolutional Neural Networks (CNN) classification can be a suitable approach for this scenario based on literature review. Videos are divided into set of frames/images and the images are passed to a CNN model. The neural network will automatically learn the features of the images and hence gives the prediction.

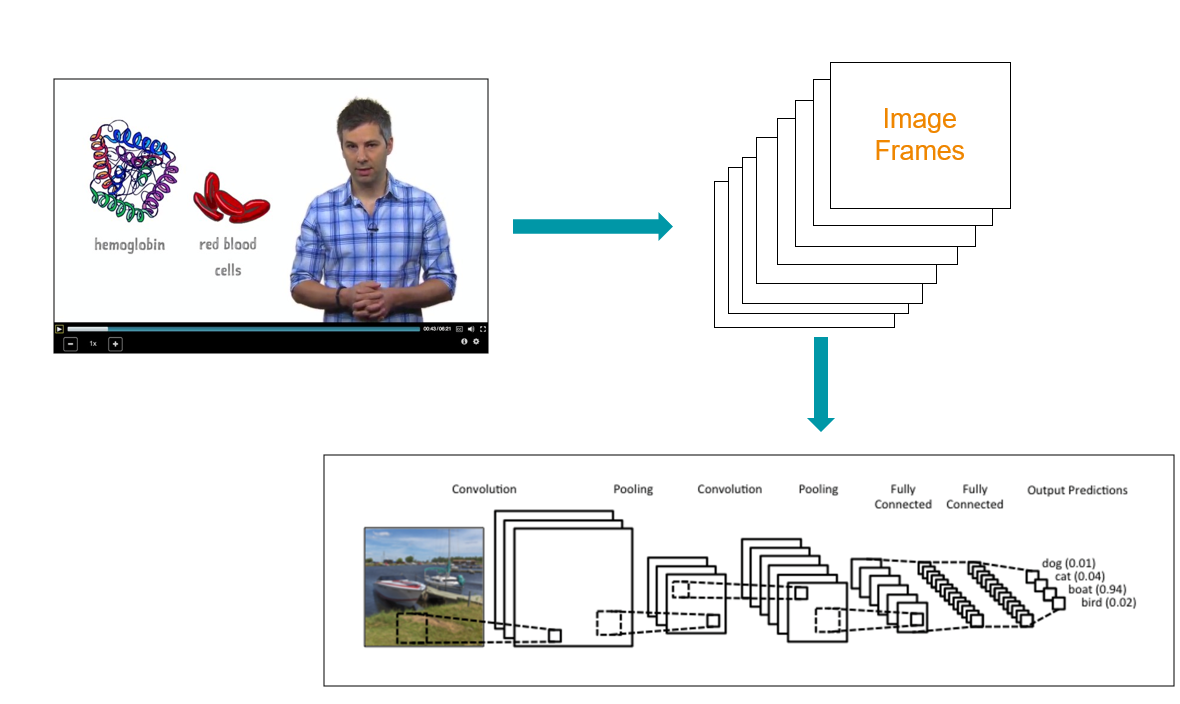


Figure 8: Frame-based Video Classification using CNN

The measurable outcomes for this component are:

* Accuracy: The accuracy of topic modeling can be measured by comparing the video styles classified by the neural network with the style actually contained in the given video, which can be determined manually.
* Speed: The speed of the algorithm can be measured with the amount of time taken to classify a given video(s) into a set of production style categories.

### Learner Profile Modelling and Recommendation

Using the ILS questionnaire, learner’s learning style will be captured and stored in database. Hence, the learner’s profile will be modelled accordingly. The MOOCs data extracted from the platforms are labelled with a particular learning style and given a ranking based on standard algorithms and literature review. And hence, based on the learning styles the most appropriate MOOCs are presented to the learner in a ranked order.

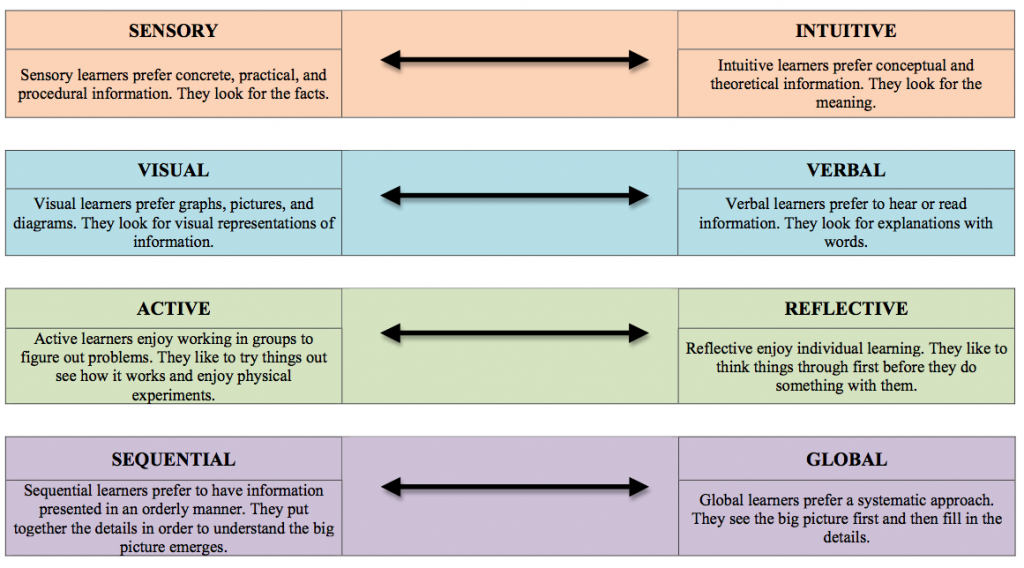


Figure 9: Felder and Silverman Learning Style Characterstics

The measurable outcomes of for this component are:

* Accuracy: The accuracy can be measured by validating the MOOCs recommended by the system against the MOOCs selected by the learner in an open environment. The learners should be from the same educational background with the same amount of domain knowledge for the given course.
* Speed: The speed can be measured with the amount of time taken to show the recommendations.

## Software Development Life Cycle

The software development methodology which is going to be used as the process of development of the system is Agile Scrum. Agile Scrum methodology facilitate more flexibility to software development as it allows for requirement change to be handled with ease throughout the process of development. Since research projects are associated with a lot of changes it is hard to follow a set sequence of development such as the waterfall model. Scrum concentrates on task management within a team based environment. As this project is done by a group of four members this methodology will help to team to perform all the formal and informal activities and to solve problems faces by individual members as well as facilitate common understanding of each area of the project by conducting regular meetings.

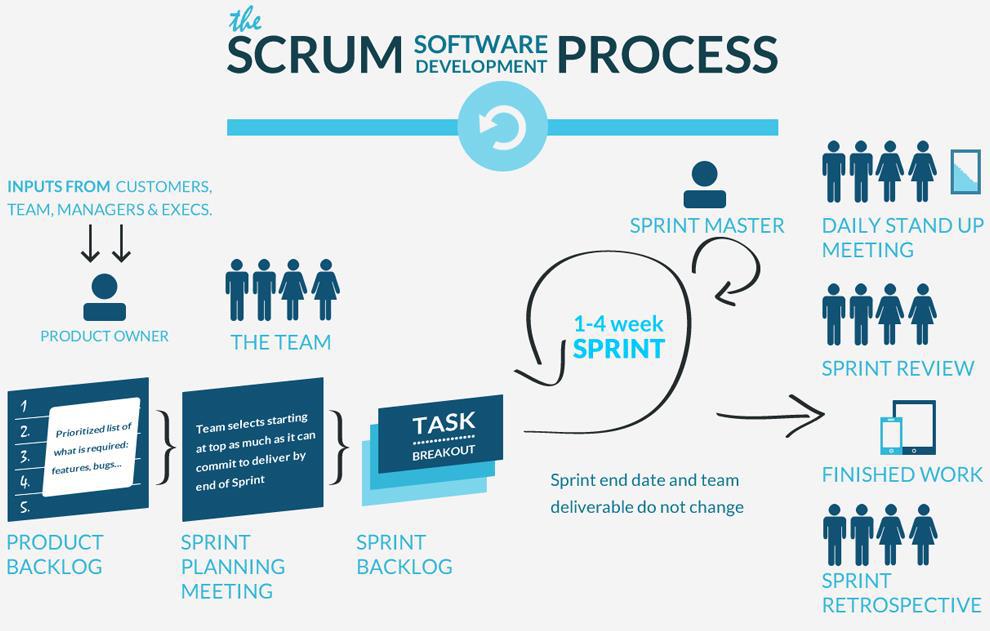


Figure 10: Agile Scrum Development Life Cycle

## Gantt Chart

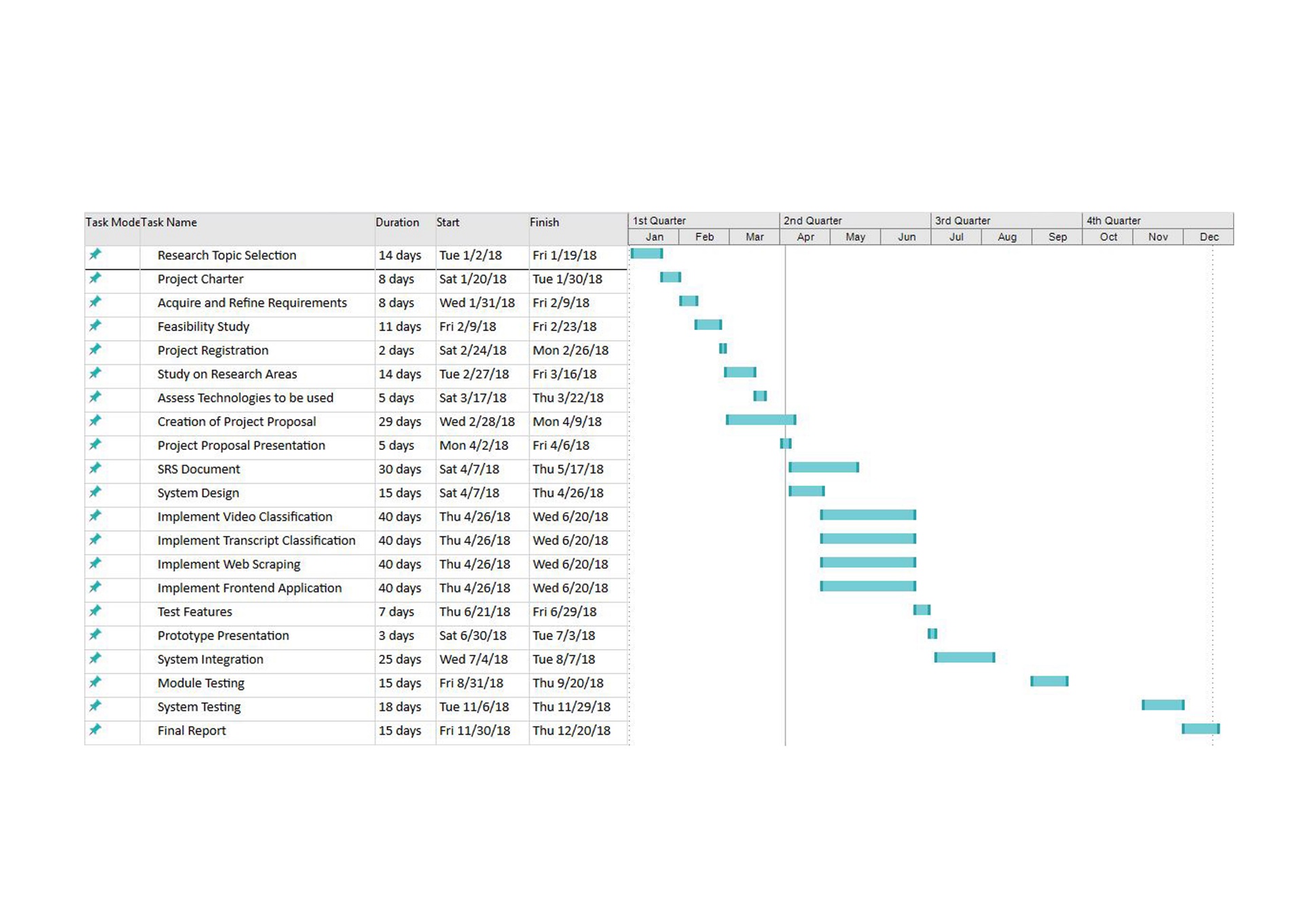


Figure 11: Gantt Chart

Figure 12: Gantt Chart

## Work Breakdown Structure

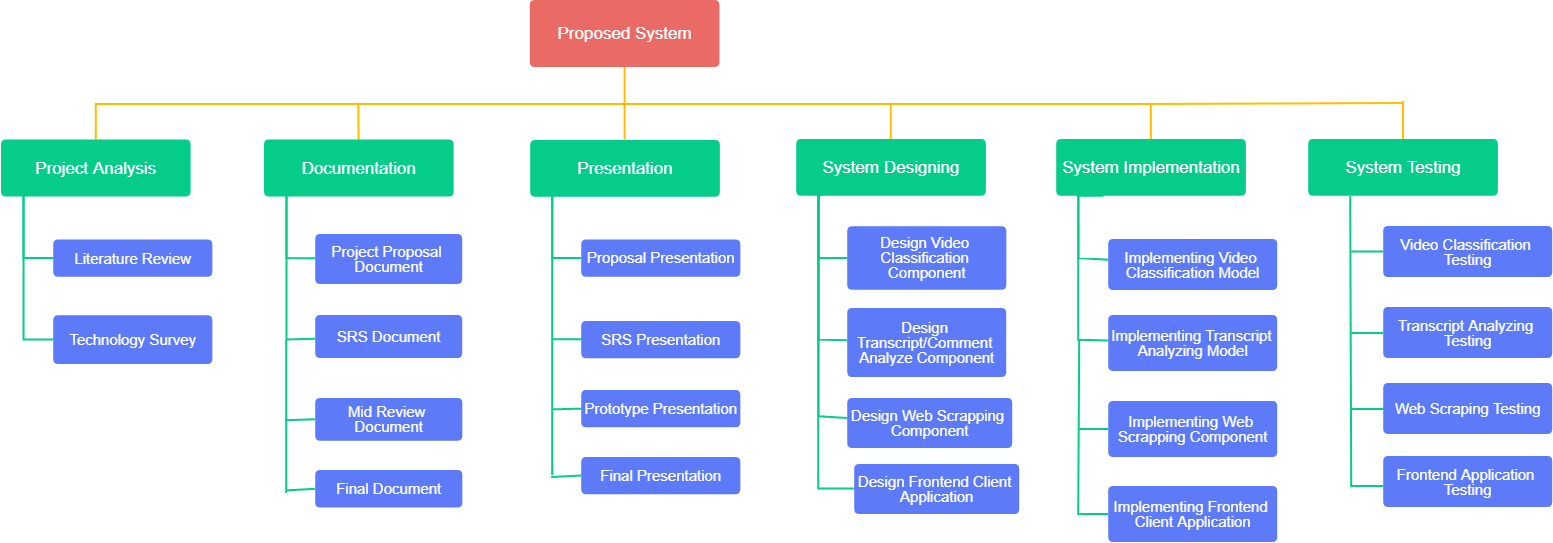


Figure 13: Work Breakdown Structure

# Tools and Technologies

Following tools and technologies are used while developing the proposed system:

**Tools**

* DBMS Tools
* Google Colaboratory
* PhpStorm
* Jupyter Notebook

**Technologies**

* Python, Scrapy
* Keras – Deep Learning Library
* Google TensorFlow
* MongoDB, MySQL
* PHP, HTML5, CSS3

# DESCRIPTION OF PERSONAL AND FACILITIES

|  |  |  |
| --- | --- | --- |
| Registration No | Name | Task Description |
| IT14146602 | Saugat Aryal | - Classify a given individual MOOC video into a specific video production class by breaking it down into frames and passing it into a trained model of neural networks.  - Extract information about what percentage of different video styles are contained in the video and save it to the database.  - Classify the audio of a given MOOC video as native or non-native English speaker.  - Documentation  - Testing |
| IT14142024 | Porawagama A.S | - For a given MOOC video, perform topic modelling to discover the abstract topics for the specific video.  - Analyze comments for a single and overall videos in MOOC.  - Calculate the complexity of the transcript or the English language in terms of percentage.  -Prediction models evaluation.  -Documentation  -Testing |
| IT14140280 | Hasith M.G.S | - Develop different web scraper to extract MOOCs information from different platforms: edX, Coursera, Futurelearn.  - Implement a web crawler to periodically scan through given platforms to check for new courses.  - Documentation  - Testing |
| IT141382832 | Thoradeniya S.D | - Development of frontend client application.  - Identify learners’ learning style based on the ILS questionnaire.  - Mapping MOOCs and Learning Styles and assigning them weight, based on literature review and algorithms.  - Recommend the most suitable MOOCs based on learners’ learning style.  - Query for relevant and appropriate MOOCs based on users filter parameters.  - Documentation  - Testing |

# Business Potential

The proposed system presented in the paper has many potentials to be commercialized. The business strategy of the system is listed as follows:

* Ability to collaborate and partner with different MOOCs platforms to provide the noble services of our system.
* It can provide handful of information to various MOOCs developers to design a course being cost effective.
* Developed as a web application, Ads can be incorporated in the system.
* Different packages of services can be provided to the learners at standard charge.
* The system can also be used by various learning institutions to provide best learning environments to their students.

# REFERENCES

[1] R. M. Felder, R. M. Felder, and L. K. Silverman, “Learning and teaching styles in engineering education,” *Eng. Educ.*, 1988.

[2] M. H. Baturay, “An Overview of the World of MOOCs,” *Procedia - Soc. Behav. Sci.*, vol. 174, pp. 427–433, Feb. 2015.

[3] A. Mohamed, F. Yousef, M. A. Chatti, U. Schroeder, M. Wosnitza, and H. Jakobs, “A Review of the State-of-the-Art.”

[4] “By The Numbers: MOOCS in 2017 — Class Central.” [Online]. Available: https://www.class-central.com/report/mooc-stats-2017/. [Accessed: 26-Mar-2018].

[5] P. J. Guo, J. Kim, and R. Rubin, “How video production affects student engagement,” in *Proceedings of the first ACM conference on Learning @ scale conference - L@S ’14*, 2014, pp. 41–50.

[6] A. Hansch, L. Hillers, K. McConachie, C. Newman, T. Schildhauer, and P. Schmidt, “Video and Online Learning: Critical Reflections and Findings from the Field,” *SSRN Electron. J.*, Mar. 2015.

[7] J. Reutemann, “Differences and Commonalities – A comparative report of video styles and course descriptions on edX, Coursera, Futurelearn and Iversity,” in *European MOOCs Stakeholders Summit*, 2016.

[8] S. Graf and T.-C. Liu, “Supporting Teachers in Identifying Students’ Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach,” *Educ. Technol. Soc.*, vol. 12, no. 4, pp. 3–14, 2009.

[9] C. Li and H. Zhou, “Enhancing the Efficiency of Massive Online Learning by Integrating Intelligent Analysis into MOOCs with an Application to Education of Sustainability,” *Sustainability*, vol. 10, no. 2, p. 468, Feb. 2018.

[10] N. Bajraktarevic1, W. Hall1, and P. Fullick2, “Incorporating learning styles in hypermedia environment: Empirical evaluation.”

[11] K. K. Graf Sabine, “Providing adaptive courses in learning management systems with respect to learning styles,” in *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Association for the Advancement of Computing in Education (AACE)*, 2007, pp. 2576–2583.

[12] H. A. Fasihuddin, G. D. Skinner, and R. I. Athauda, “Boosting the Opportunities of Open Learning (MOOCs) through Learning Theories,” *GSTF J. Comput.*, vol. 3, no. 3, p. 31, Dec. 2013.

[13] H. A. Fasihuddin, G. D. Skinner, and R. I. Athauda, “Personalizing Open Learning Environments through the adaptation to Learning Styles.”

[14] F. Bousbahi and H. Chorfi, “MOOC-Rec: A Case Based Recommender System for MOOCs,” *Procedia - Soc. Behav. Sci.*, vol. 195, pp. 1813–1822, Jul. 2015.

[15] Q. Cheng and Y. Gao, “Courducate --An MOOC Search and Recommendation System.”

[16] C. D.-K. Israel Gutiérrez-Rojas, Derick Leony, Carlos Alario-Hoyos, Mar Pérez-Sanagustín, “Towards an Outcome-based Discovery and Recommendation of MOOCs using moocrank,” in *European MOOCs Stakeholders Summit*, 2014.

[17] Y. Wang, B. Liang, W. Ji, S. Wang, and Y. Chen, “An improved algorithm for personalized recommendation on MOOCs,” *Int. J. Crowd Sci.*, vol. 1, no. 3, pp. 186–196, Sep. 2017.

[18] D. Fu, Q. Liu, S. Zhang, and J. Wang, “The Undergraduate-Oriented Framework of MOOCs Recommender System,” in *2015 International Symposium on Educational Technology (ISET)*, 2015, pp. 115–119.

[19] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, “Large-Scale Video Classification with Convolutional Neural Networks,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732.

[20] Lokhande B Gharde S, “A Review on Large-scale Video Classification with Recurrent Neural Network (RNN),” *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 2, pp. 1774–1778, 2015.

[21] Joe Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, “Beyond short snippets: Deep networks for video classification,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 4694–4702.

[22] M. Pratusevich, “EdVidParse: Detecting People and Content in Educational Videos,” *Massachusetts Inst. Technol.*, 2013.

[23] A. Sajid, S. Jan, and I. A. Shah, “Automatic Topic Modeling for Single Document Short Texts,” in *2017 International Conference on Frontiers of Information Technology (FIT)*, 2017, pp. 70–75.

[24] J. Shang, J. Liu, M. Jiang, X. Ren, C. R. Voss, and J. Han, “Automated Phrase Mining from Massive Text Corpora,” *IEEE Trans. Knowl. Data Eng.*, pp. 1–1, 2018.

[25] K. Kotani and T. Yoshimi, “A Machine Learning Approach to Measurement of Text Readability for EFL Learners Using Various Linguistic Features,” *US-China Educ. Rev. B*, vol. 6, pp. 767–777, 2011.

[26] M. Lu, S. Wen, Y. Xiao, P. Tian, and F. Wang, “The design and implementation of configurable news collection system based on web crawler,” in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017, pp. 2812–2816.

[27] Z. Shi, M. Shi, and W. Lin, “The Implementation of Crawling News Page Based on Incremental Web Crawler,” in *2016 4th Intl Conf on Applied Computing and Information Technology/3rd Intl Conf on Computational Science/Intelligence and Applied Informatics/1st Intl Conf on Big Data, Cloud Computing, Data Science & Engineering (ACIT-CSII-BCD)*, 2016, pp. 348–351.

[28] Y. Sun, P. Jin, and L. Yue, “A Framework of a Hybrid Focused Web Crawler,” in *2008 Second International Conference on Future Generation Communication and Networking Symposia*, 2008, pp. 50–53.

[29] Z. Guojun, J. Wenchao, S. Jihui, S. Fan, Z. Hao, and L. Jiang, “Design and application of intelligent dynamic crawler for web data mining,” in *2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2017, pp. 1098–1105.

[30] S. Graf, S. R. Viola, T. Leo, and Kinshuk, “In-Depth Analysis of the Felder-Silverman Learning Style Dimensions,” *J. Res. Technol. Educ.*, vol. 40, no. 1, pp. 79–93, Sep. 2007.

[31] B. Ciloglugil and M. M. Inceoglu, “A Felder and Silverman Learning Styles Model Based Personalization Approach to Recommend Learning Objects,” Springer, Cham, 2016, pp. 386–397.

[32] G. Borges and I. Stiubiener, “Recommending learning objects based on utility and learning style,” in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 2014, pp. 1–9.

[33] “Class Central • #1 Search Engine for Free Online Courses &amp; MOOCs.” [Online]. Available: https://www.class-central.com/. [Accessed: 03-Apr-2018].

[34] “Browse the best MOOC and free online trainings in the world - My Mooc.” [Online]. Available: https://www.my-mooc.com/en/. [Accessed: 03-Apr-2018].

[35] “Home | MoocLab - Connecting People to Online Learning.” [Online]. Available: https://www.mooclab.club/. [Accessed: 03-Apr-2018].

[36] “Student reviews of online courses on Computer Science, Business, Design, Data Science, Humanities and more | CourseTalk.” [Online]. Available: https://www.coursetalk.com/. [Accessed: 03-Apr-2018].

1. <https://www.edx.org/> [↑](#footnote-ref-1)
2. <https://www.coursera.org/> [↑](#footnote-ref-2)
3. <https://www.udacity.com/> [↑](#footnote-ref-3)
4. <https://www.futurelearn.com/> [↑](#footnote-ref-4)