# KANTIPUR ENGINEERING COLLEGE

(Affiliated to Tribhuvan University)

Dhapakhel, Lalitpur



# [Subject Code: CT654] A MINOR PROJECT FINAL REPORT ON IOE RESULT AND NOTICE VIEWER WEBSITE

# **Submitted by:**

Aayush Shrestha [28002]

Aliz Shrestha [28007]

Ankit Kafle [28010]

Kashyap Ghimire [28033]

# A MINOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

#### **Submitted to:**

**Department of Computer and Electronics Engineering** 

March, 2023

# IOE RESULT AND NOTICE VIEWER WEBSITE

# **Submitted by:**

Aayush Shrestha [28002]

Aliz Shrestha [28007]

Ankit Kafle [28010]

Kashyap Ghimire [28033]

# A MINOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

#### **Submitted to:**

Department of Computer and Electronics Engineering
Kantipur Engineering College
Dhapakhel, Lalitpur

# **ACKNOWLEDGEMENT**

Firstly, we would like to give our most profound appreciation towards our project manager, Er. Bishal Thapa for his steady direction, support and proposals amid the course of our project development. We are moreover profoundly obliged to Er. Sahaj Shakya for his important recommendations and insights on web scraping and OCR. We would also like to thank our class teacher Er. Nishan Khanal who helped us during the course of our project when we encountered some errors and glitches. Besides, we would like to thank the Department of Computer and Electronics of Kantipur Engineering College for giving us resources for the completion of the project. Moreover, we are exceptionally thankful to Er. Suman Shrestha and Mr. Tek Narayan Adhikari for helping us by providing additional resources in the form of college emails which were used during the course of project. At long last, we would like to recognize and thank everybody who helped us specifically or in a roundabout way in this extend.

# **ABSTRACT**

In a real world scenario, such as a college campus, the information in the form of notice, hand-written manual, and verbal message, is being spread among the students. Today it is of the essence to not only use the predictable forms of statements, but also new forms such as cell phone technology and websites for faster and easier communication among the students. Viewing the results, notices immediately after the publication is a challenge for many students. due to which it consumes student's time and resources. To overcome this, a website will be designed and developed that allows the students of Kantipur Engineering College to login with a unique ID that allows them to view results, notices and send out notification regarding if the registered students have passed the exam or not. We use the concept of Optical Character Recognition(OCR) to scan the symbol number of the students and check their result status and send the user result status directly to their emails. For extracting the notices from Institute of Engineering (IOE), we use the algorithm of Web Scraping to fetch the important notices that are useful for the students.

Keywords—Optical Character Recognition, Web Scraping,

# **TABLE OF CONTENTS**

| A   | cknow                 | ledgeme                    | ent                         | i  |  |  |  |  |
|-----|-----------------------|----------------------------|-----------------------------|----|--|--|--|--|
| Al  | ostrac                | t                          |                             | ii |  |  |  |  |
| Li  | st of f               | igures                     |                             | v  |  |  |  |  |
| Li  | List of abbreviations |                            |                             |    |  |  |  |  |
| 1   | Introduction          |                            |                             |    |  |  |  |  |
|     | 1.1                   | Backgro                    | ound                        | 1  |  |  |  |  |
|     | 1.2                   | Problen                    | m Statement                 | 1  |  |  |  |  |
|     | 1.3 Obje              |                            | tives                       |    |  |  |  |  |
|     | 1.4                   | Applica                    | ation Scope                 | 2  |  |  |  |  |
|     | 1.5                   | Feature                    | s                           | 2  |  |  |  |  |
|     | 1.6                   | Feasibil                   | lity Study                  | 2  |  |  |  |  |
|     |                       | 1.6.1                      | Economic Feasibility        | 3  |  |  |  |  |
|     |                       | 1.6.2                      | Technical Feasibility       | 3  |  |  |  |  |
|     |                       | 1.6.3                      | Operational Feasibility     | 3  |  |  |  |  |
|     |                       | 1.6.4                      | Schedule Feasibility        | 4  |  |  |  |  |
|     | 1.7                   | System                     | Requirements                | 4  |  |  |  |  |
|     |                       | 1.7.1                      | Development Requirements    | 4  |  |  |  |  |
|     |                       | 1.7.2                      | Deployment Requirements     | 5  |  |  |  |  |
| 2   | Literature Review     |                            |                             |    |  |  |  |  |
|     | 2.1                   | Related                    | Projects                    | 6  |  |  |  |  |
|     | 2.2                   | Related                    | I research                  | 6  |  |  |  |  |
| 3   | ME                    | ГНОDО                      |                             | 8  |  |  |  |  |
|     | 3.1                   | Workin                     | g Mechanism                 | 8  |  |  |  |  |
|     | 3.2                   | UML D                      | Diagrams                    | 15 |  |  |  |  |
| 3.3 |                       | Software Development Model |                             |    |  |  |  |  |
|     | 3.4                   | Require                    | ment Analysis               | 18 |  |  |  |  |
|     |                       | 3.4.1                      | Functional Requirements     | 18 |  |  |  |  |
|     |                       | 3.4.2                      | Non-Functional requirements | 18 |  |  |  |  |
| 4   | Resi                  | ılt and D                  | Discussion                  | 19 |  |  |  |  |
|     | 4.1                   | Result                     |                             | 19 |  |  |  |  |

|                                     | 4.2 | Discussion         | 19 |  |  |  |  |
|-------------------------------------|-----|--------------------|----|--|--|--|--|
|                                     | 4.3 | Limitations        | 20 |  |  |  |  |
| 5 Conclusion and Future Enhancement |     |                    |    |  |  |  |  |
|                                     | 5.1 | Conclusion         | 21 |  |  |  |  |
|                                     | 5.2 | Future Enhancement | 21 |  |  |  |  |
| References                          |     |                    |    |  |  |  |  |
| Anney                               |     |                    |    |  |  |  |  |

# LIST OF FIGURES

| 1.1 | Gantt Chart                  | 4   |
|-----|------------------------------|-----|
| 3.1 | Block diagram of the system  | 8   |
| 3.2 | Flow Chart of Web Scraping   | . 1 |
| 3.3 | Flowchart based on VIPS      | 2   |
| 3.4 | Use Case Model of the System | 5   |
| 3.5 | DFD Level 0 of the System    | 6   |
| 3.6 | DFD Level 1 of the system    | .7  |
| 3.7 | Incremental Model            | 8   |

# LIST OF ABBREVIATIONS

CSS: Cascading Style Sheet

DFD: Data Flow Diagram

EMAIL: Electronic Mail

HTML: Hyper Text Markup Language

IOE: Institute of Engineering

KEC: Kantipur Engineering College

OCR: Optical Character Recognition

PHP: Hypertext Preprocessor

SQL: Structured Query Language

URL: Uniform Resource Locator

VIPS: Vision Based Page Segmentation

# CHAPTER 1 INTRODUCTION

#### 1.1 Background

With the advancement in time and technology, there is a need for faster dissemination of information. The increasing advantages of automated systems now are at their highest position thus many manual processes are automated. Since the automated system is demanded now-a-days, educational infrastructures like colleges needed their manual system to function on Mobile Computing Systems [1]. Today's world is the world of information and technology. Small and big tasks, meetings can be attended through different mobile website. Viewing a website frequently for notices is a hectic thing to do. Similarly verifying the results immediately after the publishment is also quite difficult as the result may come anytime and the particular user can be unaware about it.

So creating a user-friendly website that provides the user to view the important notices of IOE and providing notification every time a notice has been published is an ideal thing to do. The main goal of this project is helping the student to view the notices from IOE. Not only viewing the notices but also helping the students to view their result status immediately after the publishment is the main aim of the project. Not limited to a single stream, our website facilitates on different streams and every student of every field can gain access to our website which provides different facilities.

#### 1.2 Problem Statement

In today's context, time is a key factor in everyone's life. This statement is true for everyone including students. For a delicate topic like viewing exam results. The process should be quick as students are worried about their results. These processes are very time consuming and tedious. Most of the important notices are missed by the students because there is a lack of a proper notification system. So to solve these issues we have decided to develop the website in order to help the students to view their result status and be updated with the latest notices.

Furthermore, it is challenging to obtain manual records of the cars that utilized the parking lot in the future. The database built into our system can give historical records in an emergency or when a record is urgently needed to help with emergency control.

#### 1.3 Objectives

- I To implement easy result viewing for all the faculties.
- II To send the result status automatically via email to users after publishment.
- III To provide an effective notice display system

#### 1.4 Application Scope

The system being designed is economical with respect to the students' point of view. The goal is to extract useful information from unstructured data using the concept of information retrieval, filtering and secure random algorithms. Our basic approach attempts to develop a website which can be used to make this process easier, secure and less error prone. This project provides a platform to extract real time data from IOE websites. It can be used as a tool to see the result status of the students according to their semester result.

#### 1.5 Features

- 1. Provides an efficient and easy way to view notices.
- 2. Provides easy to use interface.
- 3. Provides subscription based result viewer.
- 4. Provides quick access to view result status of students

#### 1.6 Feasibility Study

The feasibility study is one of the main important things to be considered for the project development. The feasibility study must be done for different factors affecting the project. Here are some factors whose feasibility study should be done for our project.

#### 1.6.1 Economic Feasibility

Economic feasibility attempts to weigh costs of developing and implementing a new system, against the benefits that would accrue from having the new system in a place. This feasibility study gives the top management the economic justification for the new system. A simple economic analysis which gives the actual comparison of costs and benefits is much more meaningful meaning in this case. In addition, this proves to be a useful point of reference to compare actual costs as the project progresses. There could be various types of intangible benefits on account of automation. These could include increased customer satisfaction, improvement in product quality, better decision making, and timeliness of information, expediting activities, improved accuracy of operations, better documentation and record keeping, faster retrieval of information, better employee morale

#### 1.6.2 Technical Feasibility

Since the proposed system uses software technologies and tools which are freely available and technical skills required can be tricky but are manageable. There are many free machine learning libraries available for data analysis and predictions with proper documentation and courses. The hardware system in the project need not be highly computing but requires normal computing and the system server must be adequate and 4 manageable in the future. so it is seen that the hardware and software meet the needs of the system. So it's clear that the proposed project is technically feasible.

#### 1.6.3 Operational Feasibility

Though the advancement in technologies, any kind of system software or application is no longer hard to operate. Thus the user needs to be only a bit familiar with the software system backed with graphical explanations and that can be easily understood faster in time with usage. This system highly focuses on design-dependent parameters like reliability, maintainability, supportability, usability, predictability, sustainability, affordability, etc. So, the project is feasible in operation.

## 1.6.4 Schedule Feasibility

Schedule Feasibility is defined as the probability of a project to be completed within its scheduled time limits, by a planned due date. If a project has a high probability to be completed on-time, then its schedule feasibility is appraised as high. Schedule feasibility ensures that a project can be completed before the technology becomes unnecessary. Since there are many features in our project but can be implemented in a quality way it has a very high probability to be completed on time.

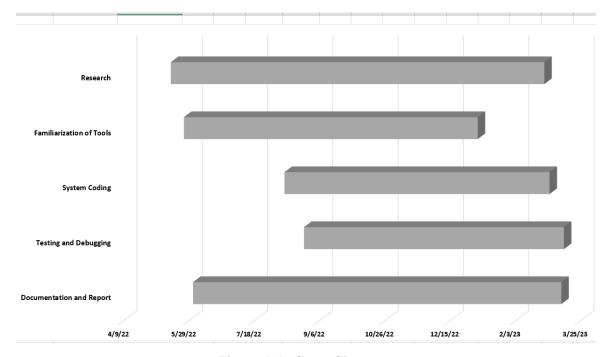


Figure 1.1: Gantt Chart

### 1.7 System Requirements

The system requirements for the project are as follows:

#### 1.7.1 Development Requirements

#### **Hardware Requirement(Minimum)**

PC with a minimum specification of 8GB RAM and a sixth generation i5 processor

# **Software Requirement**

• Microsoft Windows 7/8/10 (32 or 64 bit) or Mac OS

# 1.7.2 Deployment Requirements

# **Hardware Requirement(Minimum)**

- More than 1.5 GHz clock speed
- Minimum 4GB RAM

# **Software Requirement**

- Visual Studio Code
- Pycharm

# CHAPTER 2 LITERATURE REVIEW

#### 2.1 Related Projects

- IOE Syllabus: IOE Syllabus is a website in which students can see the notices, faculty syllabus, faculty notes, etc. Also students can get collection of past questions from this website.
- Onlineocr.net: A user friendly and free online OCR converter tool service to extract text from picture and that is available free of cost at the internet, help in doing Optical character recognition online.

#### 2.2 Related research

In recent years, the use of technology and the popularity of e-commerce have led to a significant increase in the number of online shopping websites available. While this makes it convenient for people to shop online, it can also be time-consuming and require a lot of effort to search for the best deals and offers on these websites. In order to address this issue, the authors of the paper "Exploiting Filtering approach with Web Scraping for Smart Online Shopping" have developed a web scraping framework that allows users to efficiently gather product information and deals from multiple ecommerce websites. The framework is implemented using front-end technologies such as HTML (Hypertext Markup Language) and CSS (Cascading Style Sheet) as well as a back-end language, PHP (Hypertext Preprocessor). It uses Python libraries and HTML tags to write the scraping scripts, and the results are displayed dynamically to the user rather than being stored in a local database. According to the authors, this framework has a high accuracy rate of 93time required. The goal of the framework is to make it easier for people to find the best deals and offers on e-commerce websites and save them time and effort in the process.[2]

The paper "Barcode Character Defect Detection Method Based on Tesseract-OCR" aims to address the increasing quality requirements for barcodes as their use becomes more widespread with the advancement of information technology. However, during

the printing process, various defects can occur, such as flying ink, missing print, wrong print, black spots, and improper registration, which can be caused by factors such as poor typography, inadequate printing equipment, and imperfect printing technology. The traditional method of manually sorting defective barcodes is inefficient and prone to error due to the influence of various factors, leading to low precision in detection. In order to address these issues, the authors propose a method for detecting defects in barcodes using Tesseract-OCR (Optical Character Recognition) software. This method involves segmenting the barcodes using horizontal projection, recognizing the characters in the barcodes using Tesseract-OCR, and applying the Levenshtein Distance algorithm to detect character defects. The authors conducted experiments using 1000 barcode images, and the results showed that the proposed method had an accuracy of 94.3detecting barcode defects.[3]

The paper "NewsOne — an aggregation system for news using web scraping method" is an online platform that aggregates and summarizes the latest news updates from multiple national and international sources. It presents this information in a concise and easy-to-read format, and is designed to allow users to quickly and easily access the news without wasting time searching or waiting for content to load. To achieve this, NewsOne employs web scraping and crawling techniques to extract news content from various websites, which it then categorizes based on user interest. The platform is also service-oriented, allowing users to interact with each other from across the web. NewsOne uses a bot that dynamically extracts content from the stored URL's RSS feeds at set intervals, which are added to a database by the platform's administrators or subadministrators. This model of categorization is designed to extract useful information for classifying news articles into specific categories, such as Just-In, Technology, Health, Science, Sports, Business and Economics, and Entertainment. NewsOne's user experience is flexible, allowing readers to choose the categories of news that interest them and read the news for free and as quickly as possible. In addition, users have access to up-to-the-minute daily news coverage and headlines from over 100 fully licensed and trusted news sources from around the world. Finally, the platform also provides recommendations and thoughts for future development [4].

# CHAPTER 3 METHODOLOGY

#### 3.1 Working Mechanism

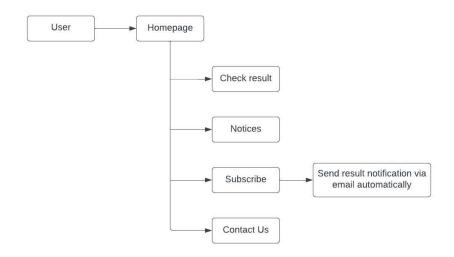


Figure 3.1: Block diagram of the system

- Homepage: At first, the user interacts with the homepage of our system where they will be able to select options such as Check Result, Notices, Subscribe and Contact Us section.
- 2. View Notices: One major feature of the website is to show Notices published by IOE. This feature now allows the user to view notices in real time after the IOE has published it. This feature can be implemented by the use of web scraping in which the website exam.ioe.edu.np will be scraped for notices. To view notices it is not necessary that the user must be subscribed.
- 3. View Results: Viewing results immediately after the publication with a proper interface is a hard thing to do. So with the webpage, the user can verify the result and see if they have passed the exams. To use this feature, the user doesn't require to be subscribed in into the 8 system. But to use the feature of getting automatically notified via email regarding the result status, the user must be subscribed using the email provided by KEC. The webpage admin automatically sends activation email to the subscriber and only after the subscription is activated, then after the publishment of result, the system automatically sends them an email regarding if they have passed or not. This can be achieved by using web scraping

and OCR scan.

- 4. Subscribe: Viewing results immediately after the publication with a proper interface is a hard thing to do. So with the webpage, the user can verify the result and see if they have passed the exams. To use this feature, the user doesn't require to be logged in into the system. But to use the feature of getting automatically notified via email regarding the result status, the user must be logged in using the email provided by KEC. The webpage automatically verifies the user and sends them an email regarding if they have passed or not. This can be achieved by using web scraping and OCR scan.
- 5. Contact Us:If any problem arises then the user can directly contact the developer to solve the issue. The user must fill up the form where they have to insert their Name, Email, Subject and the message they want to send. The mail is now sent to the developers team email address

#### Web Scraping

Web scraping is the process of collecting structured web data in an automated fashion. It's also called web data extraction. Some of the main use cases of web scraping include price monitoring, price intelligence, news monitoring, lead generation, and market research among many others. In general, web data extraction is used by people and businesses who want to make use of the vast amount of publicly available web data to make smarter decisions. For our project we will use web scraping to extract the IOE notices and result from their website [5].

- Cropping the image: First, the image after the detection of the position of the license plate and the bounding box formulation is cropped according to the coordinates of the bounding ox so that we are resulted with the number plate from where the character extraction is to be done.
- Noise reduction and gray scaling: In order to male the image suitable for the character recognition, the image is filtered for the removal of noise and followed by the gray scaling of the cropped image so that the characters which are to be detected are separated from the background color. The conversion of image into

black and white that is, improving the contrast causes better performance of the model.

The processed image is then sent through the process of OCR where the characters that are present in the image are extracted. The extraction process or the character recognition process is done with the help of either pattern recognition or feature recognition.

# **Procedure for Web Scraping**

- 1. Start
- 2. Create scraping template
- 3. Browse website
- 4. If the content is found, go to step 5, else repeat step 3
- 5. Get the link post
- 6. Explore the link post
- 7. Get the required data (Data extracted)
- 8. Stop

# Flowchart of Web Scraping

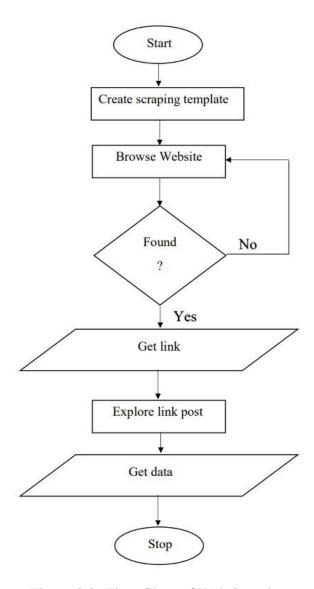


Figure 3.2: Flow Chart of Web Scraping

#### VIPS BASED SEARCHING

- 1. Start
- 2. Construct parse tree for the HTML tags
- 3. Isolate all the anchor tags
- 4. Search for the result and notices inside the anchor tag
- 5. If result and notices are present, go to step 6, else go to step 7
- 6. Extract the content of anchor tag
- 7. If the entire webpage is scraped, go to step 8, else go to step 4

#### Flowchart based on VIPS

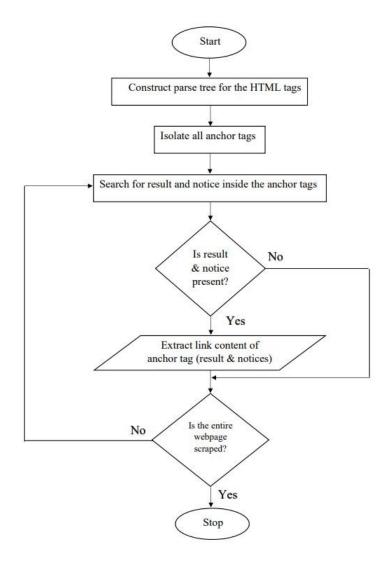


Figure 3.3: Flowchart based on VIPS

## **Implementation of Web Scraping**

- 1. At first, the website of IOE is visited. After this, the site is explored.
- 2. The HTML content of the link is explored up until 10 pages initially to gather data into the database.
- 3. Now the entire webpage is converted into a parse tree through the use of an HTML parser. As from the inspection of the webpage, we found that the data of the notices and results along with the date of publishment and the link of the PDF file

for the notices and results were stored by forming a table. The publishing date of the results and notices were stored in the third data cell for every table row. Also, every table row was considered for the extraction of the PDF file and the title for the notice and result. Only the notices and results of BE were taken in this stage.

- 4. Following the above step, the entire table rows of the table were explored and the title of the notices and results, which were present in the content of the <span> tag, was extracted. Also, the link for the PDF was present in the <href> part of the anchor tag, which was also extracted.
- 5. On observation of the actual PDF link for the notices and results, we encountered a problem in which the %20 part of the link was regarded as space which caused the actual link of the PDF to break. This was easily solved by replacing the space with %20 and joining the starting of the URL "https://exam.ioe.edu.np" with the link from the <href> tag.
- 6. Now the filtered title obtained from the content of <span> from step 4, the date of the PDF obtained from step 3, along with the link of the PDF file obtained at the end of step 5, was inserted into the built-in SQLite database with a constraint that if the link of the notices and results were already present in the top of the database, then the insertion process would stop.

#### **Optical Character Recognition(OCR)**

OCR also referred as text scan is the process that extracts and repurposes data from scanned documents, camera images and image-only pdfs. OCR systems use a combination of hardware and software to convert physical, printed documents into machine-readable text. For our project we will use OCR to extract the symbol number as a string and then we will compare the symbol number that has been obtained with the help of OCR scan to the symbol number of the students. If the symbol number 12 matches, then the backend of our project will send an email to the registered users automatically and then the students can know whether they have passed or not.

In our project the OCR scanning will be performed using a library in python.

# **Implementation Of OCR**

- 1. First, extract the PDF file's link from the database.
- 2. Using the link from step 1, extract the PDF file.
- 3. Convert the extracted PDF to an image format.
- 4. Scan the characters from the image and save the resulting string in the database.

# 3.2 UML Diagrams

# **Use Case Diagram**

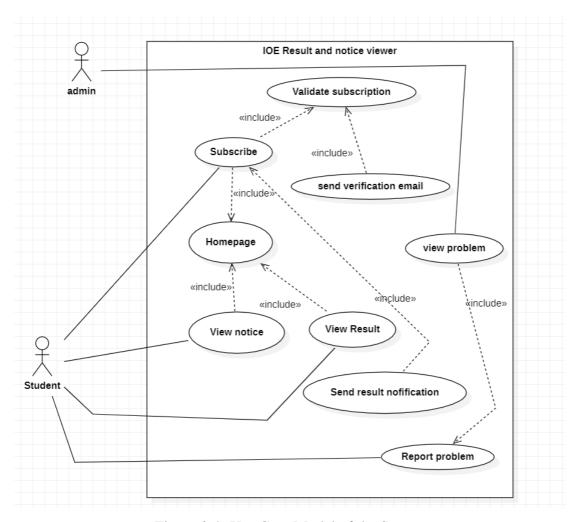


Figure 3.4: Use Case Model of the System

# **DFD Diagrams**

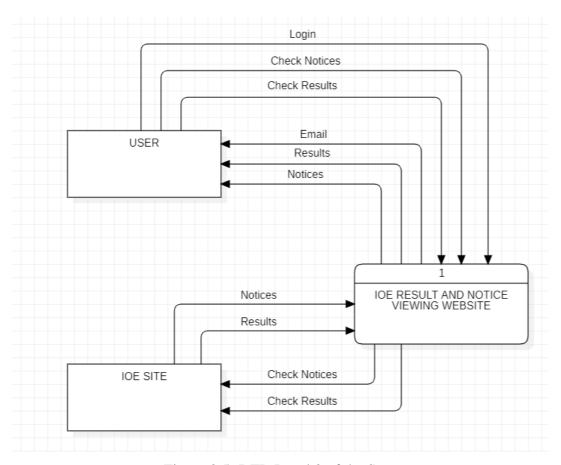


Figure 3.5: DFD Level 0 of the System

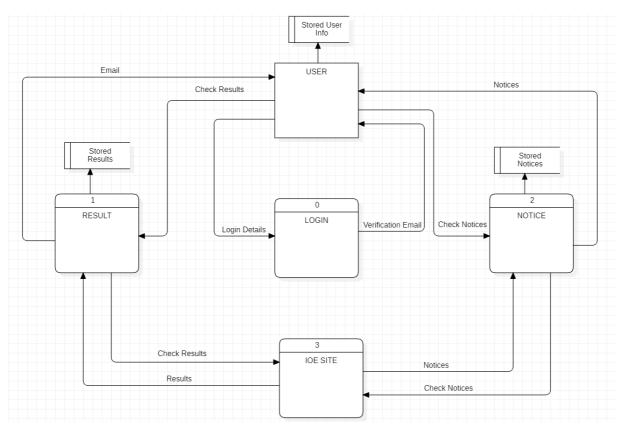


Figure 3.6: DFD Level 1 of the system

## 3.3 Software Development Model

The incremental model is one of the easiest to implement software development life cycle models. There are certain scenarios where the initial or the core software requirements are clearly defined, but the actual span or the full set of features of the project are unknown. Moreover, the development company might decide to not give the full functionality of the software in one go. Rather they prefer to give it out through periodic updates. Or the client requests some functionality enhancements during the process of development. In such cases, the incremental model is used.

In the first phase we performed Web Scraping and extracted all the results and notices from the website. In the second phase we used Database using MYSQL and inserted all the notices and results in the database. In the third phase we implemented the concept of OCR. In the fourth phase we performed mailing methods to send the email of result status to the subscribed students

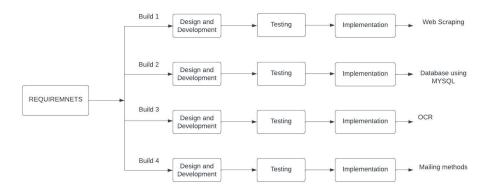


Figure 3.7: Incremental Model

# 3.4 Requirment Analysis

# 3.4.1 Functional Requirements

- The system should view the notices, results..
- The system should provide subscription feature for students of KEC only.
- The system should send result status email only to the subscribed accounts automatically after the publishment of result.
- The system should provide a feature to contact with admin of the website to report problems and error.

#### 3.4.2 Non-Functional requirements

- The system should be user friendly and easy to use.
- The system should be compatible and should run on almost all the device.
- The system should be reliable and responsive.

# CHAPTER 4 RESULT AND DISCUSSION

#### 4.1 Result

We have completed the design and development of the project along with obtaining the desired output of the project. The project currently takes in the symbol number of the guest as well as subscribed users and shows the users the status of their current semester's result. Along with this feature, the project also sends the status of the result automatically via email to the registered users, allows the user to view the notices published by IOE, allows the user to ask queries and report issues via "contact us" section and finally sends the activation email to the subscriber's email to confirm the validity of the entered data.

#### 4.2 Discussion

IOE Result and Notice Viewer website is a website that takes in the symbol number of the guest as well as subscribed users and shows the users the status of their current semester's result. Along with this feature, the project also sends the status of the result automatically via email to the registered users, allows the user to view the notices published by IOE, allows the user to ask queries and report issues via "contact us" section and finally sends the activation email to the subscriber's email to confirm the validity of the entered data. For the development of the website we used bootstrap, HTML, CSS, JavaScript to design the front end and of the website. The concept of web scraping was used to scrape IOE's website for accessing the results and notices immediately after the publishment, OCR scan was also used to scan the symbol numbers of the passed student from the pdf file uploaded in the IOE's website. These methods were implemented using python and Django along with MySQL as framework and database respectively. Finally, after the completion of the project, the notices and results uploaded in the IOE's website up until the tenth page were scraped and stored in the database. And among the scraped notices and results, only results were selected and the OCR scan was used in the selected results to obtain the symbol numbers. After this, the symbol numbers along with the faculty, exam date, semester information were inserted into the database and

19 using this data, the result status of the student was displayed to the guest users and the email about the status of the result was sent to the subscribers account automatically after the publishment of the result.

#### 4.3 Limitations

- Only the results that are published in BS Year 2079 can be viewed. .
- After subscribing in case of any error form the user must contact the administrator as they cannot change it themselves.
- The user must fill up the form after every semester as the details will be changed.

# CHAPTER 5 CONCLUSION AND FUTURE ENHANCEMENT

#### 5.1 Conclusion

In this project, we were able to design and build a website that helped the students of IOE to view the notices and result status after publishment. Also, only the students of Kantipur Engineering College could subscribe to the website to get the result status immediately via email after the publishment. The core of this system was designed using the web scraping and OCR.

#### **5.2** Future Enhancement

As our project still lacks in some sectors, further enhancements and improvements can be made for the system to be more efficient and accurate, the following enhancements can be done:

- This project can be further extended into mobile applications so that a user can have one touch mobile experience.
- Implementation of login system for individual users for security and privacy.
- This project can be further extended to upload the mark sheet and store it in the database.

# **ANNEX**

