

Prize Bond Checker

A system for matching the prize bond with the winning list quickly and saving our valuable time.

**A Project Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of
Science in Computer Science and Engineering at the University of Asia Pacific**

Submitted By

Nayem Mia

ID: 17201075

Mahamudul Hasan Ripath

ID: 17201091

Abu Hamed Saifullah Prince

ID: 17101133

Supervised By

Md. Imran Bin Azad

Senior Lecturer, BRAC University



Department of Computer Science & Engineering

University of Asia Pacific

June 2022

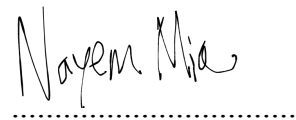
DECLARATION

We therefore declare that the work titled "Machine learning based OCR for text detection" given in this project is the result of our study conducted under the supervision of Md Imran Bin Azad, Senior Lecturer, BRAC University's Department of Computer Science and Engineering. We further state that no component of this project has been plagiarized from other sources or has been previously submitted. This is the only result of our investigation.

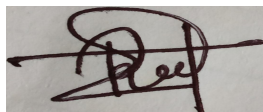
Countersigned

Signature

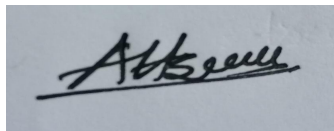
.....
(Md Imran Bin Azad)


.....

(Nayem Mia)



.....
(Mahamudul Hasan Riphath)



.....
(Abu Hamed Saifullah Prince)

ABSTRACT

Nowadays OCR is used in many areas in our day to day life .OCR stands for optical character recognition.

In our project we are using OCR for detecting text from images and the purpose is to check prize bonds .Usually people buy bonds and check them manually and it is time consuming and also irritating .To solve this problem we are approaching this system.

In our system there will be a winner list and users will just scan their bond and if their bond is from the winning bond they will get the result as winner otherwise not.

ACKNOWLEDGEMENT

We would like to convey our thankfulness to Allah, the Almighty, first and foremost. Since he supplied us with the ability, opportunity, and cooperative supervisor, we are now able to perform our project work with such ease.

We'd like to use this opportunity to express our gratitude to Md Imran Bin Azad, our excellent supervisor. Despite the fact that he was frequently preoccupied with other matters, he made sure we had enough time to complete our assignment. He not only gave us his time, but he also offered us fantastic assistance and sound advice whenever we had issues. His comments and recommendations were really helpful in the planning of our project. Last but not least, we are grateful to our family members and friends who have always been there for us through thick and thin.

Table of contents

Declaration.....	1
Abstract.....	2
Acknowledgement.....	3
Table of Contents.....	4
List of Figures.....	6
Chapter 1: Introduction.....	7
1.1 Overview.....	7
1.2 Background study.....	8
1.3 Motivation.....	9
1.4 Objective.....	10
Chapter 2: System Requirements.....	11
2.1 Overview.....	11
2.2 Software requirements.....	11
Chapter 3: Proposed System.....	13
3.1 Proposed system.....	13
3.2 Demo figures.....	14
Chapter 4: Methodology.....	16
4.1 Web page & server.....	17
4.2 Dataset & training.....	17
4.3 CSV.....	19

Chapter 5: Result & Scope of Further Work.....	21
5.1 Results.....	21
5.2 Discussion.....	21
5.3 Scopes of Further Work.....	21
5.4 Conclusion.....	22
Reference.....	22

List of Figures

Figure 1.2.1: OCR pipeline.....	8
Figure 3.2.1: Matching the winning prize.....	14
Figure 3.2.2: Matching the not winning prize.....	15
Figure 4.1: Prize bond.....	16
Figure 4.1: Frontend of the system.....	17
Figure 4.2.1: Annotation.....	19
Figure 4.3.1: Winners list.....	20

Chapter 1: Introduction

1.1 Overview

The government of a country issues a prize bond, which is a non-interest bearing investment, to raise funds and reduce borrowings. Individual savers are putting this money aside for themselves. It is an interest-free savings account, but the benefits are shared by a lottery that will be held after a set period of time. Any financial body that does not have government power cannot advertise or provide this.

The Bangladesh government developed the Prize Bond, a public savings plan, in 1974 with the goal of pooling local resources and offering incentives to small depositors. The prize draw takes place every quarter on January 31, April 30, July 31, and October 31. Prize bonds are government-issued financial products. These bonds are issued by the government to raise funds whenever it is required. The only way an investor can lose money is if their prize bond isn't picked. Prize bonds are drawn three times a year.

Optical Character Recognition (OCR) is a method of extracting data from written or printed text in an image or scanned document and converting it to machine-readable text for data processing such as scanning, editing, and searching. **OCR** can assist reduce the amount of physical space necessary to keep documents while also improving document workflows.

To convert a scanned image into text, **OCR** makes up the letters and numbers. It also compares scanned text to other scanned texts, documents, or data.

OCR is working in our system to determine whether or not a prize bond or lottery has won. A window or notification appears if the bond matches the outcome.

1.2 Background Study

Web application

The process of generating a web that may run on any web platform is known as web application development.

Framework

Gradio

Gradio is an open-source python library which allows you to create the frontend and backend of an ML model, any API, or any arbitrary function in just a few lines of code. You can integrate the GUI directly into your Python notebook, or you can share the link to anyone. Gradio is also the fastest way to create your machine learning model with a friendly web interface so that anyone can use it anywhere.

ML(Machine Learning)

Any Typical machine learning OCR pipeline follows the following steps :

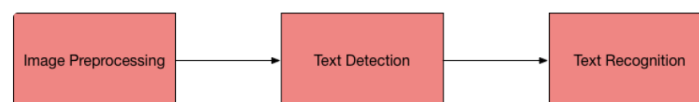


Figure 1.2.1: OCR pipeline

Machine learning (ML) is the study of computer algorithms that can learn and develop on their own with experience and data.

We used an optical character recognition (OCR) system for our project.

Text Detection is a technique in which the model is given a picture and the textual region is detected by plotting a bounding box around it.

Text Recognition: Following text detection, content recognition occurs, in which the discovered textual portions are further processed to determine what the text is.

It will show the result after recognizing the text, whether it matches the outcome or not.

Tesseract

PyTesseract is a OCR tool for python that can also be used as a wrapper for the tesseract OCR engine. It can read from an image and recognize text in images. It is commonly used in python OCR image to text use cases.

1.3 Motivation

Prize Bonds are a type of paper currency issued by the Bangladeshi government for the aim of saving. In 1972, the Department of National Savings introduced the "Bangladesh Prize Bond" as a prize bond to encourage people to save money. Lottery Bonds are another name for prize bonds. A Prize Bond lottery ticket can be purchased by anyone. Each year, the Bangladesh Bank Prize Bond Draw is held on the 31st of January, 30th of April, 30th of July, and 31st of October. There are around 26 series, each with 1000000 reward bonds. If a prize bond number is chosen, all prize bonds in that series are also chosen. Every prize bond has a unique number, and the numbers must be matched when the prize bond results are revealed. Manually checking it is inconvenient and time consuming for people, and it must be done every four months. People that check that number on a regular basis will benefit greatly from an automation system. The initial stage in any project development process is to select the appropriate method. It is critical to analyze the dynamics of selecting the suitable project development method before making a decision. There are various strategies to pick from depending on the types of people participating in the project. Choosing the appropriate methodology might help you get more done in less time.

1.4 Objectives

The major goal is to create a dynamic system that can modify lottery outcomes and check for matching by scanning an image. For our system, we'll establish a database. Those results will be stored in our system's database, and anyone who wants to check whether or not he won the lottery can do so by scanning the prize bond with our scanner.

Chapter 2: System Requirements

2.1 Overview

The examination of numerous processes conducted by the system, as well as the interaction between the inside and outside of the system, is known as system requirements analysis. This is the most crucial step in the project development process. Information is collected in an accessible file when analyzing system requirements, and decision points are processed by the online system. When analyzing system requirements, a standard approach for solving the problem using a computer should be used. Hardware support teams and project developers must build several lines of code to handle today's systems, which are quite complicated. We concluded that the final system should be created first, followed by a system requirements analysis. The investigation included a detailed examination of the current system, which led to the development of the new system. This resource will be the most beneficial to the project.

2.2 Software requirements

VS Code

We used VS Code for writing code and implementing our project.

Language: python

Python is a famous programming language for implementing machine learning projects. We used python for implementing this Prize Bond Checker project.

Anaconda

For annotation we have used anaconda labeling.py..... classes.txt and in labeling.py we have used yolo v4

Youlov4

YOLOv4 is a one-stage object detection model that improves on YOLOv3 with several bags of tricks and modules introduced in the literature.

HTML:

HTML, or Hypertext Markup Language, is a commonly used language for website front-end development. HTML is a markup language that is used to connect many web pages and to specify the structure of headings, lists, paragraphs, and other elements.

Bootstrap:

Bootstrap is a framework that makes it easier and faster to develop websites. It is used to make the website more responsive.

Chapter 3: Proposed System

The winning list will be in our system, and the user will scan their bond, and if the number of that bond matches up with any of the winning list numbers, the winner will be displayed.

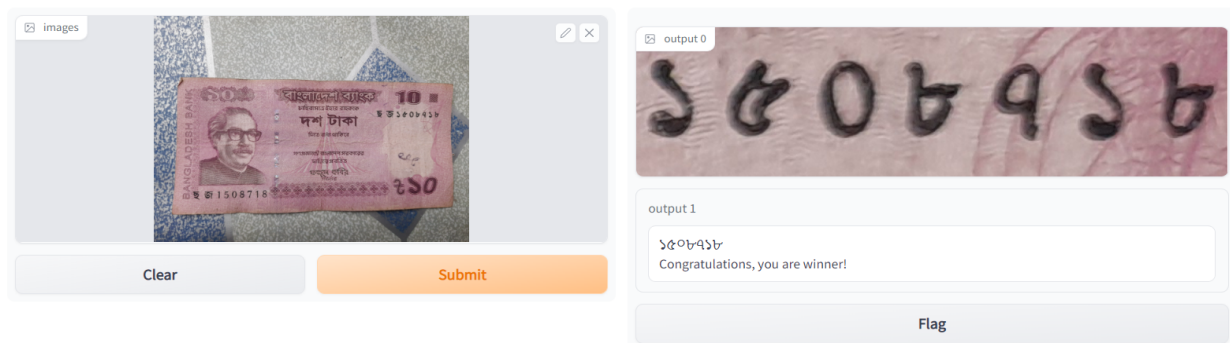
3.1 Proposed system

Prize Bonds are a type of paper currency issued by the Bangladeshi government for the aim of saving. In 1972, the Department of National Savings introduced the "Bangladesh Prize Bond" as a prize bond to encourage people to save money. Lottery Bonds are another name for prize bonds. A Prize Bond lottery ticket can be purchased by anyone. Each year, the Bangladesh Bank Prize Bond Draw is held on the 31st of January, 30th of April, 30th of July, and 31st of October. There are around 26 series, each with 1000000 reward bonds. If a prize bond number is chosen, all prize bonds in that series are also chosen. Every prize bond has a unique number, and the numbers must be matched when the prize bond results are revealed. Manually checking it is inconvenient and time consuming for people, and it must be done every four months. People that check that number on a regular basis will benefit greatly from an automation system.

3.2 Demo figures

For winning prize:

When the image is uploaded to our system, the primary role is to detect the numbers from the image and match the numbers with the winning numbers which are stored in the CSV file. When numbers are matched the system declares that the prize is the winner.



[view api](#) • built with [gradio](#)

Figure 3.2.1: Matching the winning prize

For not winning prize:

When the image is uploaded to our system, the primary role is to detect the numbers from the image and match the numbers with the winning numbers which are stored in the CSV file. When numbers are not matched the system declares that the prize is not the winner.

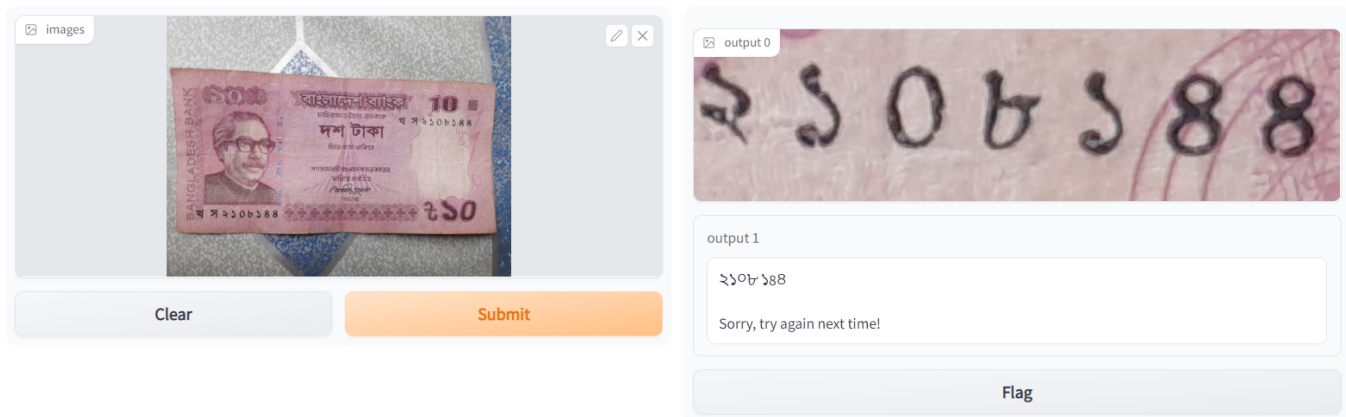


Figure 3.2.2: Matching the not winning prize

Chapter 4: Methodology

This project consists of Machine learning for scanning and detecting Text from an image. At first we thought and proposed that we will make/ create a Machine Learning for detecting text from Bangladeshi Prize bonds. But We have found that it's very tough to detect text from those. Because the images are very noisy to separate the bond number from it.



Figure 4.1: Prize bond

We have shifted the scanning from BD prize bond to Any English Text consists of lottery/ prize bond.

This System will scan and save the Scanned lottery number. And will get a winning lottery number from the CSV file. Then it will check whether it's a match or not.

Frontend of our system:

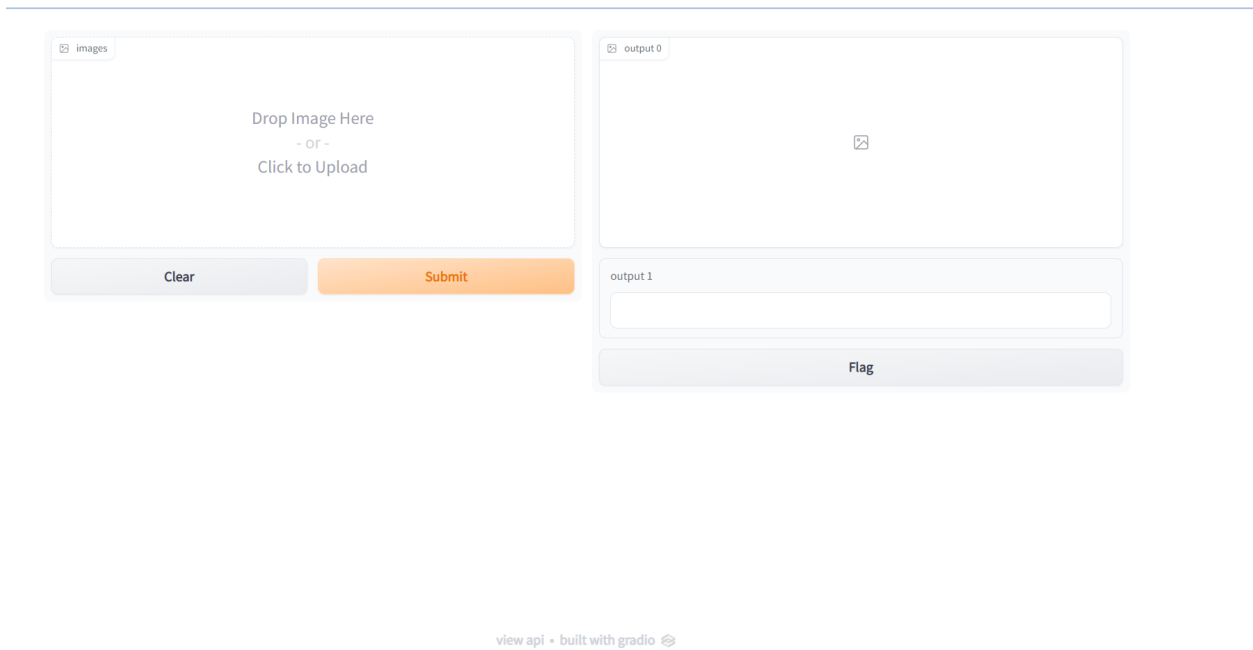


Figure 4.1: Frontend of the system

4.1 Web page & server

Gradio (Python)

Gradio is an open-source python library which allows you to quickly create easy to use, customizable UI components for your ML model, any API, or any arbitrary function in just a few lines of code. You can integrate the GUI directly into your Python notebook, or you can share the link to anyone.

4.2 Dataset & training

The ML Kit Text Recognition API can recognize text in any Latin-based character set. It can also be used to automate data-entry tasks such as processing credit cards, receipts, and business cards.

Key capabilities

- Recognize text across Latin-based languages Supports recognizing text using Latin script
- Analyze structure of text Supports detection of words/elements, lines and paragraphs
- Identify language of text Identifies the language of the recognized text
- Small application footprint On Android, the API is offered as an unbundled library through Google Play Services
- Real-time recognition Can recognize text in real-time on a wide range of devices

Text structure

The Text Recognizer segments text into blocks, lines, and elements. Roughly speaking:

- a Block is a contiguous set of text lines, such as a paragraph or column,
- a Line is a contiguous set of words on the same axis, and
- an Element is a contiguous set of alphanumeric characters ("word") on the same axis in most Latin languages, or a character in others

The image below highlights examples of each of these in descending order. The first highlighted block, in cyan, is a Block of text. The second set of highlighted blocks, in blue, are Lines of text. Finally, the third set of highlighted blocks, in dark blue, are Words.



Figure 4.2.1: Annotation

For all detected blocks, lines and elements, the API returns the bounding boxes, corner points, recognized languages and recognized text.

4.3 CSV

For this system we need storage that could be used for our winning list. So that our system can match the given prize bond number with the winning number. Here, we are using a csv file to store the winning number in our system.

Currency ☆ 📁 ☁

File Edit View Insert Format Data Tools Extensions Help

↶ ↷ 🖨️ 🔍 | 100% ▾ | \$ % .0_ .00 123 ▾ | Default (Ari... ▾ | 10

A1 ▾	fx	Name
	A	B C D
1	Name	Number
2	IMG_20220521_143751.jpg	ፀፀፏፐፐፐፐ
3	IMG_20220521_143757.jpg	ፕፕፕፕፕፕፕፕ
4	IMG_20220521_143804.jpg	ፐፐፐፐፐፐፐፐ
5	IMG_20220521_144122.jpg	ፐፐፕፕፕፐፐፐ
6	IMG_20220521_144140.jpg	ፕፕፕፕፐፐፐፐ
7	IMG_20220521_144242.jpg	ፕፐፐፐፐፐፐፐ
8	IMG_20220521_143547.jpg	ፐፐፐፐፐፐፐፐ
9	IMG_20220521_143540.jpg	ፐፐፐፐፐፐፐፐ,
10	IMG_20220521_125431.jpg	ፐፐፐፐፐፐፐፐ.
11	IMG_20220521_125358.jpg	ፕፐፐፐፐፐፐፐፐ
12	IMG_20220521_125234.jpg	ፐፐፐፐፐፐፐፐፐ
13	IMG_20220521_125229.jpg	ፐፐፐፐፐፐፐፐፐ
14	IMG_20220521_125100.jpg	ፐፐፐፐፐፐፐፐፐ
15	IMG_20220521_125044.jpg	ፐፐፐፐፐፐፐፐፐ
16	IMG_20220521_125010.jpg	ፐፐፐፐፐፐፐፐፐ
17	IMG_20220521_124932.jpg	ፐፐፐፐፐፐፐፐፐ
18	IMG_20220521_124853.jpg	ፐፐፐፐፐፐፐፐፐ.
19	IMG_20220521_124825.jpg	ፐፐፐፐፐፐፐፐፐ
20	IMG_20220521_124659.jpg	ፐፐፐፐፐፐፐፐፐ
21	IMG_20220521_124552.jpg	ፐፐፐፐፐፐፐፐፐፐ
22	IMG_20220521_123311.jpg	ፐፐፐፐፐፐፐፐፐፐ
23	IMG_20220521_122908.jpg	ፐፐፐፐፐፐፐፐፐፐ
24	IMG_20220521_122815.jpg	ፐፐፐፐፐፐፐፐፐፐ
25	IMG_20220521_122506.jpg	ፐፐፐፐፐፐፐፐፐፐ
26	IMG_20220521_122445.jpg	ፐፐፐፐፐፐፐፐፐፐ
27	IMG_20220521_122253.jpg	ፐፐፐፐፐፐፐፐፐፐ

+ ≡ Sheet1 ▾

Chapter 5: Results and Scopes of Future Work

5.1 Results:

The money we've used here, our system has been able to detect the winning numbers with almost 100% accuracy.

5.2 Discussion

When we started to work on this project our main focus was to build such a system which can detect all the text properly. We have worked together with android and text recognizer. After building the whole project we have tested it in many ways and we are getting our expected results. We kept winner numbers in a csv file in our system and we were scanning the given bonds. Our system is much more able to read the numbers from the bonds and matching with the winning numbers and showing the expected result.

5.3 Scopes of future work

Future work is the idea that can be implemented on the project in near future ideas that we can implement on our project are:

1. We can add all the languages so that all kinds of languages can be detected by this system.
2. We can create an Android App which can do the work the same as our system.
3. If we can do the second step(Android App) we can easily and instantly scan the prize bond by capturing an image in the app. Just capture and submit.

5.4 Conclusion

In short ,The main purpose to build this project was to detect prize bonds , match them and get the winning result.Checking prize bonds manually is one kind of hassle and time consuming to avoid this situation worked.Hopefully our project will reduce the waste of time and make it easy to check.

Reference

1. DeepLearning_by_PhDScholar, "Realtime Text Detection in Images using Tesseract | OpenCV | Python | Tutorial for beginners." youtube.com.
<https://www.youtube.com/watch?v=Y7XBsFzByTQ> (accessed apr 13, 2022)
2. kiyoshiiriemon, "GitHub - kiyoshiiriemon/yolov4_darknet: YOLOv4 - Neural Network." github.com.
https://github.com/kiyoshiiriemon/yolov4_darknet(accessed apr 25, 2022)
3. abidlabs, "GitHub - gradio-app/gradio: Create UIs for your machine learning models." github.com. <https://github.com/gradio-app/gradio> (accessed june10, 2022)
4. onlineocr, "Free Online OCR - Image to text and PDF to Doc converter." onlineocr.net. <https://www.onlineocr.net/> (accessed may 2, 2022)
5. Sunidhi Pandey, "Install Anaconda Python, Jupyter Notebook, and Spyder on Windows 11 [2021 Update] Anaconda Navigator." youtube.com.
<https://www.youtube.com/watch?v=F2jR6YD1tzA> (accessed apr 20, 2022)
6. codebasics, "Object detection using YOLO v4 and pre trained model | Deep Learning Tutorial 32 (Tensorflow)" youtube.com.
<https://www.youtube.com/watch?v=IfRMV2MY9n0> (accessed apr 27, 2022)
7. w3schools, "HTML Tutorial" w3schools.com.
<https://www.youtube.com/watch?v=IfRMV2MY9n0> (accessed june10, 2022)

8. visualstudio, "Download Visual Studio Code - Mac, Linux, Windows"
code.visualstudio.com. <https://code.visualstudio.com/download> (accessed Mar 10, 2022)
9. getbootstrap, "Bootstrap - The most popular HTML, CSS and JS library in the world."getbootstrap.com. <https://getbootstrap.com/> (accessed june10, 2022)
10. anaconda, "Anaconda | Anaconda Distribution."anaconda.com.
<https://www.anaconda.com/products/distribution> (accessed apr 20, 2022)

CO, PO Mapping:

How Ps are addressed through the project and mapping among Ps, COs, and POs

Ps	Attribute	How Ps are addressed through The project	COs	POs
P1	Depth of Knowledge Requirement	Our depth of knowledge will be knowing about web development, knowing about OCR. OCR stands for "Optical Character Recognition." It is a technology that recognizes text within a digital image. It is commonly used to recognize text in scanned documents and images.	CO1, CO2	a, e, i
P2	Conflicting requirements	There are no conflicting requirements in our system	CO1, CO5	e, f
P3	Depth of analysis required	Depth of analysis of requirements stakeholders – (users and authors) authority should ensure that the users are clear about how to use this system and whether they are satisfied with that system or not.	CO8	c, e
P4	Familiarity of issue	Checking the prize bond one by one is very annoying but it is very easy to check by using our system.	CO3	e, f, k

How As are addressed through the project and mapping among As, COs, and POs

As	Attribute	How As is addressed through the project	COs	POs
A1	Range of Resources	Throughout the development stage, the project requires the use of diverse resources including stakeholders (users and monitoring authority), information, and modern technologies.	CO6	A, b, c, e
A2	Level of interaction	The project requires a significant number of interactions among the users, monitoring authority, and professionals from the software industry.	CO7	D, f, j
A3	Innovation	There is no other existing system so, it's an innovation.	CO8	b, i
A4	Familiarity	Matching the bond with the results manually is time consuming and one kind of hassle. Our system will significantly reduce the time.	CO4, CO6	a, c, i