

OCR GRADING ASSISTANT

DATA FIELD EXTRACTION FROM MARKSHEET USING OCR



MAJOR PROJECT (2021-25)

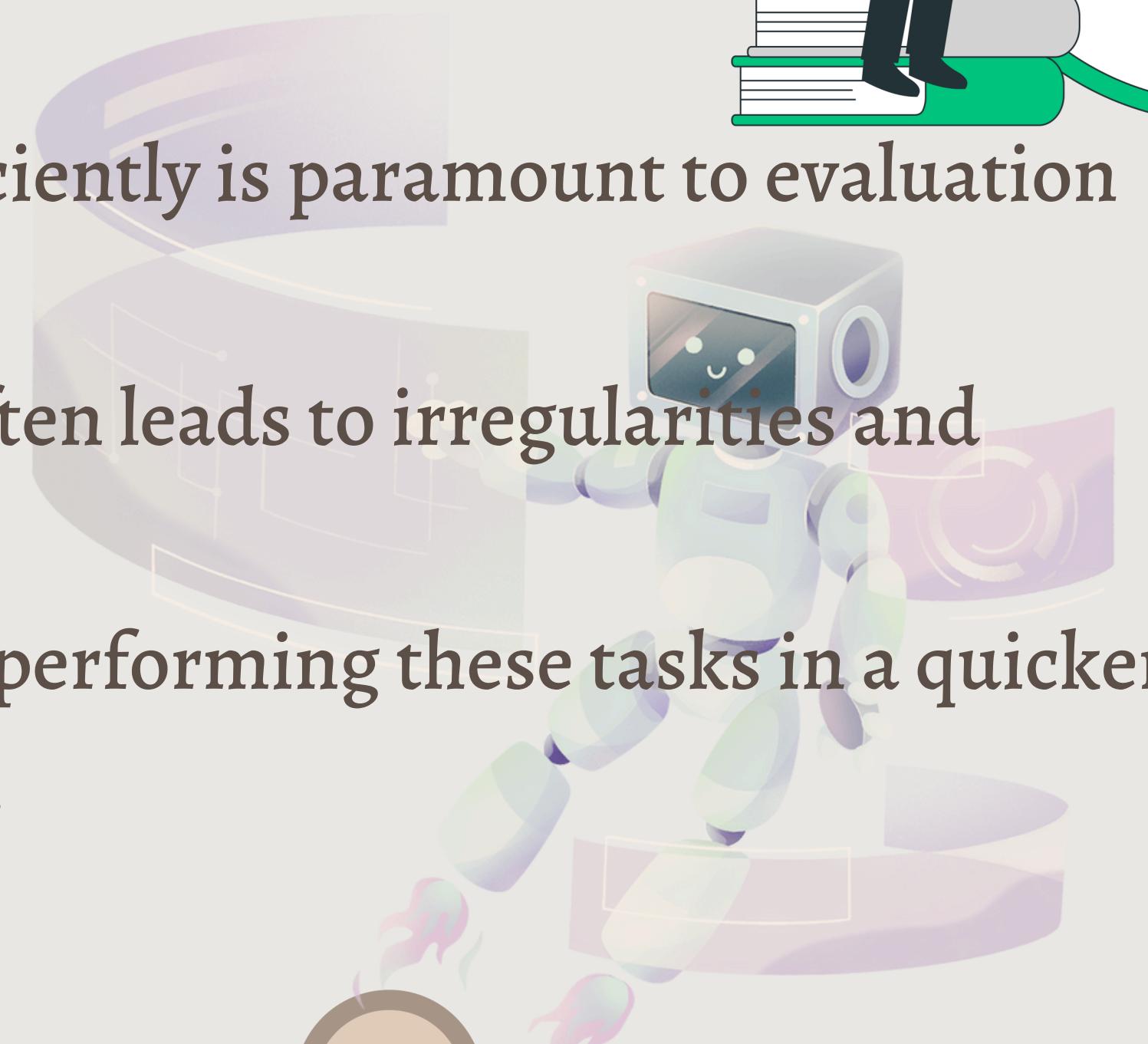
Group ID : PCS25-71

Group Member:	Shreya Sharma	Semester : 8	Department: CSE(C)
Group Member:	Mohammed Faaiz	Semester : 8	Department: CSE(B)
Group Member:	Samiksha	Semester : 8	Department: CSE(C)

Name of Guide : Prof. Vishal Kumar

PROBLEM STATEMENT

- Almost every college and university needs to maintain student marks in huge numbers.
- Storing these accurately and efficiently is paramount to evaluation of student grades etc.
- Manual handling of such tasks often leads to irregularities and discrepancies.
- An automated system capable of performing these tasks in a quicker, more efficient manner is needed.



PROJECT AIM

To develop a software GUI system to replace the traditional tiring way of manual writing of student's marks from answer sheet in database by college professors.

To Develop high accurate OCR model for character recognition.

Excel sheet generation containing roll-no, marks, etc. from answer sheet and storing it in the host computer.

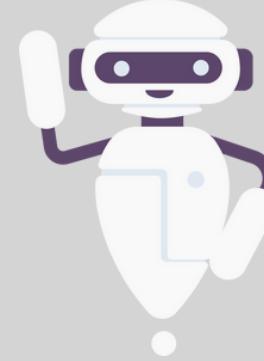


7	Class:	IX	Session:	2019-20		
8	Roll:	1	Date of Birth	1/5/2000		
9	Name :	Ron	Father's Name	Lukas		
10	Subjects	Total Marks	Passing Marks	Marks Obtained Grade		
11	English	100	35	15		
12	Math	100	35			
13	Physics	100	35			
14	Chemistry	100	35			
15	Biology	100	35			
16	Total Marks :		Obtained Marks:			
17	Passing Marks:		Percentage:			
18	Result:		Grade:			
19	exceldemy					
20						

PROJECT OBJECTIVES

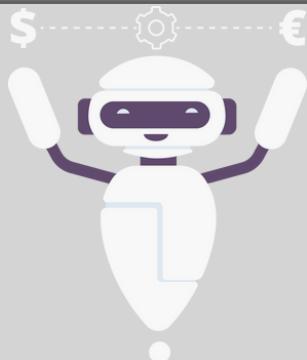
Improve accuracy

An OCR-based system can significantly reduce errors by accurately converting handwritten or printed text into digital format.



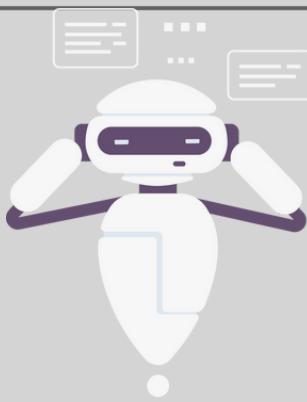
Reduced time and costs

Colleges and Universities can save money on manual labor costs and time associated with data entry and processing



Improve accessibility

System will make it easier to search, sort, and analyze student marks. Also share information with stakeholders.



PROCESS FLOW GRAPH



Hardware mounted with Camera

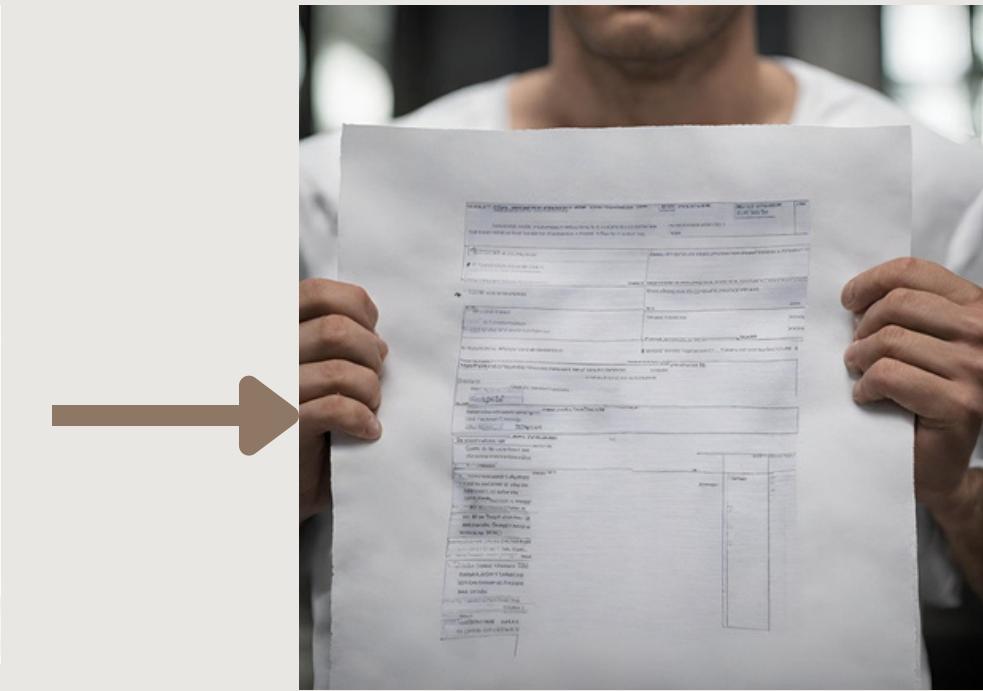
ROLL NUMBER

NAME _____

SEMESTER _____ SUBJECT CODE _____

1. Please bring your id-card at the time of examination
2. Examination will be as per the schedule given on webkiosk only
3. Reach atleast 5 minutes before exam
4. Student should be in proper dress at the time of examination

Answer Sheet Matrix



Professor holds the answer sheet
(softcopy of images)

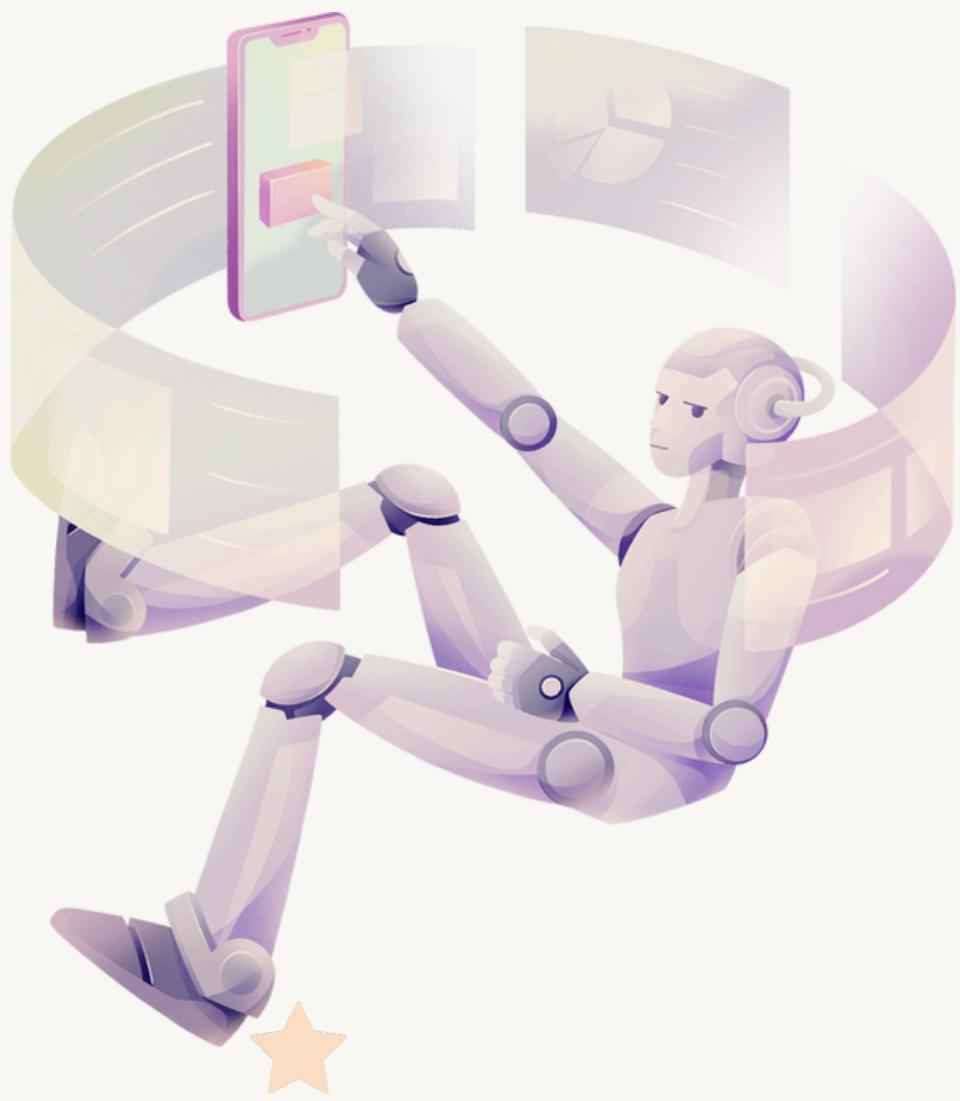
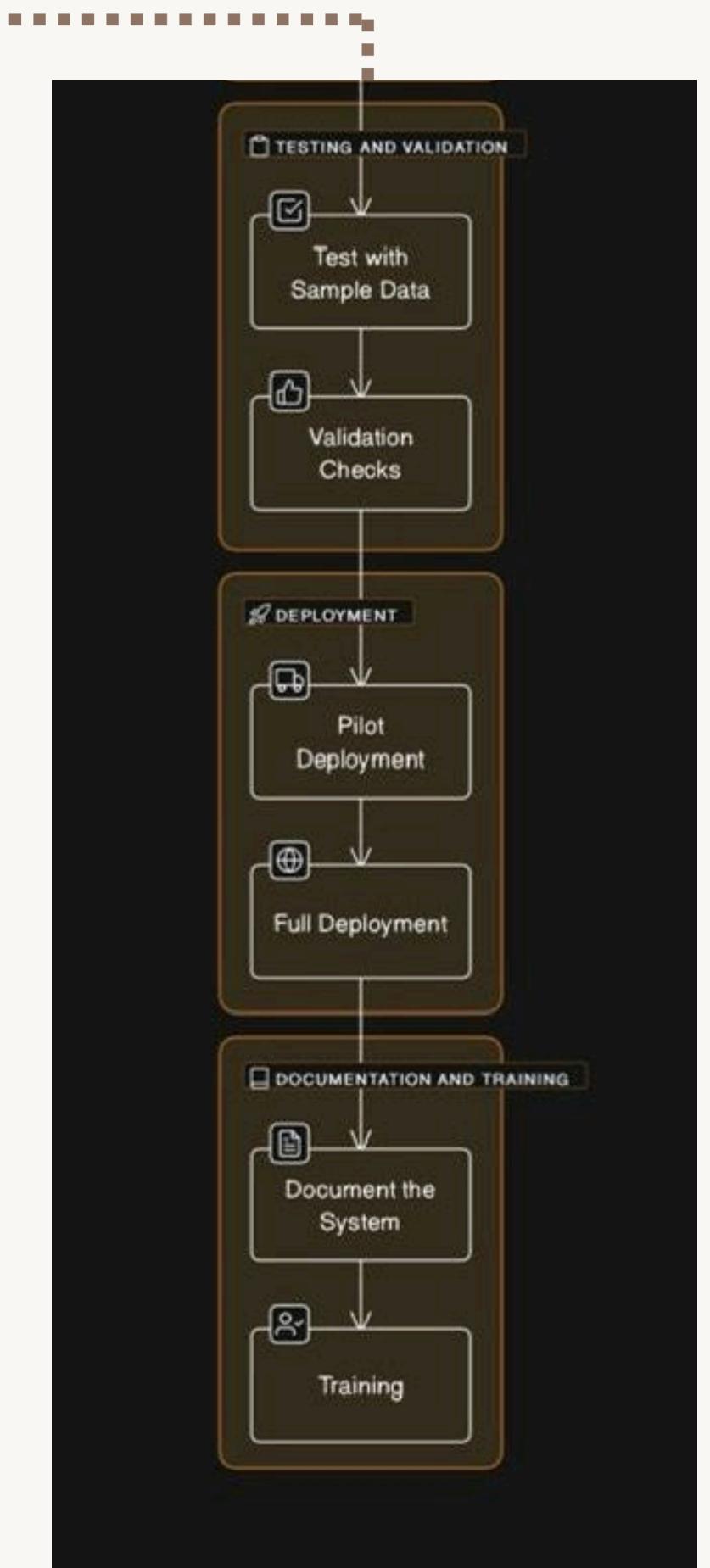
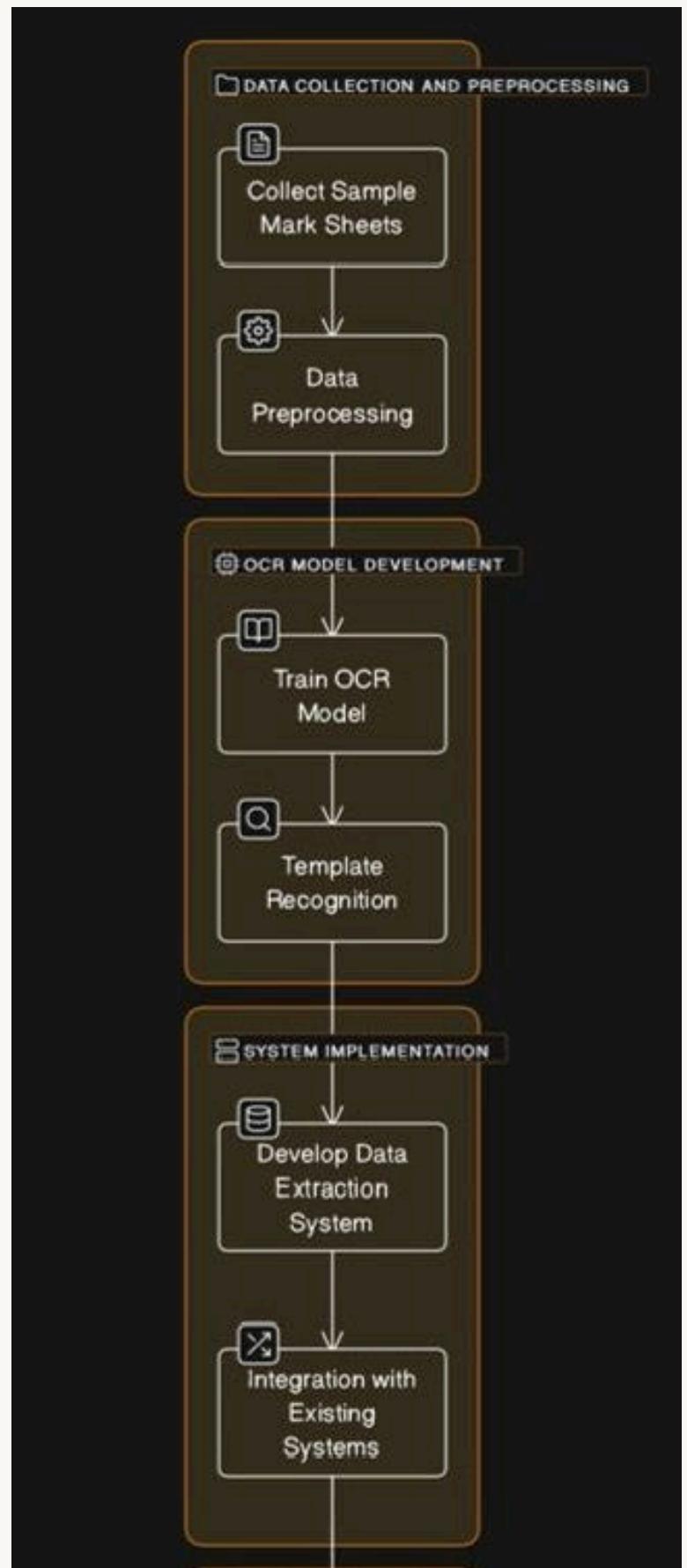
7	Class:	IX	Session:	2019-20
8	Roll:	1	Date of Birth	1/5/2000
9	Name :	Ron	Father's Name	Lukas
11	Subjects	Total Marks	Passing Marks	Marks Obtained
12	English	100	35	15
13	Math	100	35	
14	Physics	100	35	
15	Chemistry	100	35	
16	Biology	100	35	
18	Total Marks :		Obtained Marks:	
19	Passing Marks:		Percentage:	
20	Result:		Grade:	

Excel sheet generation
(containing roll-no., marks, etc)

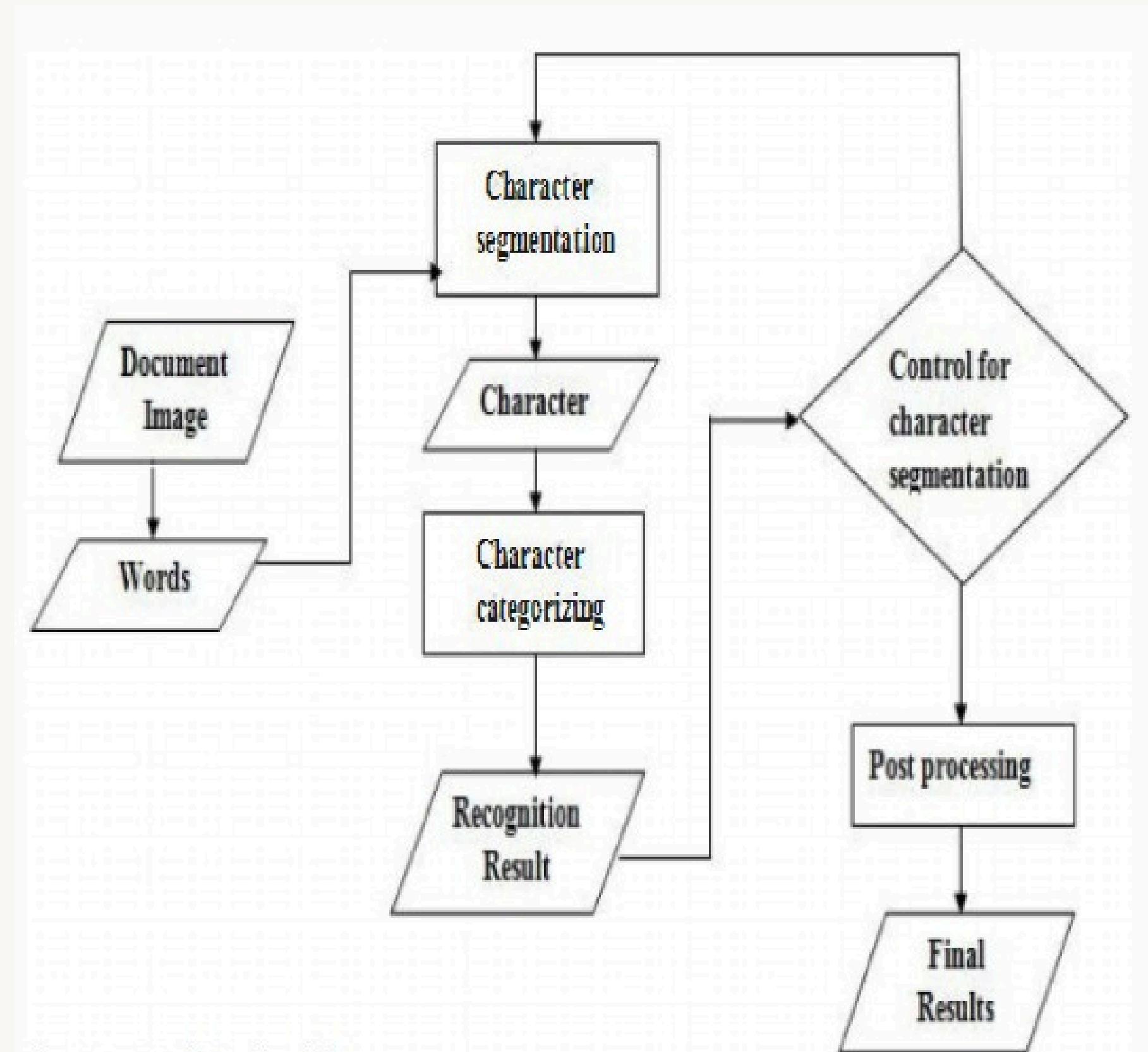


Software scanning the matrix page
(exam paper reader)

METHODOLOGY

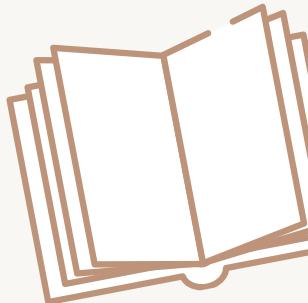


OCR: OPTICAL CHARACTER RECOGNITION



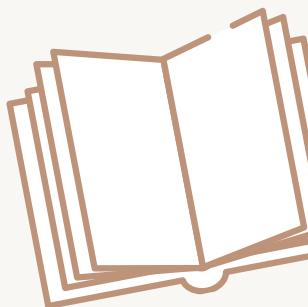
CO/PO Mapping

Course Outcome (CO)	Program Outcome (PO)
CO1: Role of image processing & OCR	PO1: Apply engineering knowledge
CO2: Design OCR-based grading system	PO2: Problem analysis
CO3: Evaluate recognition performance	PO4: Investigations & data analysis
CO4: Integrate software with hardware	PO5: Modern tool usage
CO5: Teamwork & presentation	PO9: Communication & team work



SDG Mapping

SDG Goal	Contribution through Project
SDG 4 – Quality Education	Fair, faster evaluation using automation
SDG 9 – Innovation & Infrastructure	EdTech solution with OCR & automation
SDG 12 – Responsible Consumption	Paper digitization & error reduction



Project Outcome

Automated Answer Sheet Evaluation System:

- Developed a hardware-software integrated solution to automate extraction of roll numbers and marks from student exam sheets.
- Implemented Adaptive Threshold for Color Detection (ATCD) for accurate Region of Interest (ROI) extraction under varying conditions.
- Employed Optical Character Recognition (OCR) using template matching and a dataset of 50 numeral samples (0–9).

Key Results:

- Achieved average OCR accuracy of 81% under various lighting and paper conditions.
- Replaced manual entry with automated Excel sheet generation, reducing time and human error.
- System supports voice instructions and automatic database updates for enhanced user experience.

Conclusion & Future Scope

Conclusion:

The proposed Exam Paper Reader (EPR) system successfully automates marks entry from physical exam sheets using image processing and OCR.

The method adapts to environmental noise and varying image qualities, ensuring robust performance.

Future Scope:

Enhance accuracy by training OCR with more diverse and extensive datasets.

Expand use beyond education—e.g., postal systems, automated form reading, etc.

Integrate with cloud services for real-time analytics and centralized data management.

Research Publication

Title: Data Extraction from Exam Answer Sheets using OCR

Authors: Mohammad Faaiz, Shreya Sharma, Samiksha Gupta

Institution: KIET Group of Institutions, Ghaziabad, India

**Keywords: OCR, Image Processing, Adaptive Thresholding,
Automated Grading**

THANKYOU

