



V. S. International College
(Affiliated to Tribhuvan University)
Minbhawan, Kathmandu

An Internship Report
On
Data Analysis and backend development
At
“sarbatra”

In partial fulfillment of the requirements for the Bachelor in Computer Application

Submitted to:
Department of Computer Application
V. S. International College

Submitted By
Saugat Khadka (6-2-1228-27-2021)

Under the Supervision of
Mr. Sheejan Adhikari
18th March 2026

V. S. International College
(Affiliated to Tribhuvan University)
Minbhawan, Kathmandu

An Internship Report
On
Data Analysis and backend development
At
“sarbatra”

In partial fulfillment of the requirements for the Bachelor in Computer Application

Submitted to:
Department of Computer Application
V. S. International College

Submitted By
Nawaraj Gautam (6-2-1228-16-2021)

Under the Supervision of
Mr. Sheejan Adhikari
18th March 2026

SUPERVISOR'S RECOMMENDATION

I hereby recommend that this internship project is prepared by **Mr. Arun Shrestha** as “Full stack web Developer” at **IntellectSprouts** in partial fulfillment of the requirements for the degree of Bachelor of Computer Application, he is hereby recommended for the final evaluation.

Mr.
Supervisor

Sheejan

Adhikari

LETTER OF APPROVAL

This is to certify that the project is prepared by **Mr. Arun Shrestha** entitled “**at IntellectSprouts**” in partial fulfillment of requirements for the degree of Bachelor Computer Application (**BCA**) has been well studied.

..... Mr. Sheejan Adhikari Supervisor	 Mr. Santosh Pokherel Coordinator
..... Mr. Raj Kumar Acharya Principal	 External Examiner

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Sarbatra Inc. Pvt. Ltd. for providing me with the opportunity to undertake my internship and gain hands-on experience in a professional software development environment. I am deeply thankful to the entire team for their constant support and for fostering a culture of learning and innovation.

I am particularly grateful to my mentor, Mr. Ashish Dahal (Founder), for his invaluable guidance, technical insights, and patience throughout the internship. His mentorship was instrumental in helping me understand the complexities of backend development and quality assurance, particularly during my work on the Data analyst project.

Finally, I would like to thank my college and the department faculty for their encouragement and for providing the academic foundation that allowed me to complete this internship successfully. This experience has significantly enhanced my technical skills and professional growth, preparing me for my future career in the IT industry.

Yours sincerely,

Nawaraj Gautam

ABSTRACT

This report summarizes my internship experience at Sarbatra Inc. Pvt. Ltd., a digital solutions provider in Kathmandu. During this period, I was integrated into the Software Development Unit, where I contributed to two distinct projects: the development of the Nilkantha Press web portal and the engineering of a specialized OCR Data Extraction tool. My core responsibilities included backend development using the Django framework for web services, designing OCR pipelines in Python to digitize physical documents, and conducting rigorous API testing using Postman to ensure data integrity.

Through this internship, I gained practical exposure to the Software Development Life Cycle (SDLC), specifically learning how to bridge the gap between physical records and digital systems. I worked extensively with Nepali Unicode extraction, database management, and the automated generation of analytical reports via Openpyxl. This experience significantly enhanced my technical proficiency in Python, my analytical approach to data processing, and my ability to solve complex backend challenges. The following report provides a detailed account of the organizational structure, working domains, weekly activities, and key learning outcomes achieved during my tenure.

Keywords: *Backend Development, Django Framework, OCR Engineering, Nepali Unicode, Data Analysis, API Validation, Excel Automation, Nilkantha Press.*

TABLE OF CONTENTS

SUPERVISOR'S RECOMMENDATION	3
LETTER OF APPROVAL	4
ACKNOWLEDGEMENT	5
ABSTRACT.....	6
LIST OF FIGURES	9
LIST OF TABLE	10
LIST OF ABBREVIATIONS.....	11
CHAPTER I INTRODUCTION.....	1
1.1 Introduction.....	1
1.2. Problem Statement.....	2
1.3 Objectives	2
1.3.1 General Objectives.....	2
1.3.2 Specific Objectives	2
1.4 Scope and Limitation	3
1.4.1 Scope.....	3
1.4.2 Limitation.....	3
1.5 Report Organization.....	4
CHAPTER II INTRODUCTION TO ORGANIZATION.....	5
2.1 Organization Details	5
2.3 Working Domains of Organization.....	6
2.4 Description of Intern Department/Unit	7
CHAPTER III BACKGROUND STUDY AND LITERATURE REVIEW	8
3.1 Background Study.....	8
3.2 Literature Review.....	9
CHAPTER IV INTERNSHIP ACTIVITIES.....	11

4.1 Roles and Responsibilities	11
4.1.1 Roles	11
4.2 Weekly Log.....	13
4.3 Description of Project Involved During Internship	16
4.4 Activities Performed	17
CHAPTER V CONCLUSION AND LEARNING OUTCOMES.....	17
5.1 Conclusion	19
5.2 Learning Outcomes.....	19
REFERENCES	21
APPENDICES	22

LIST OF FIGURES

Figure 1: Company Logo	5
Figure 2: Organizational Hierarchy	6
Figure 3: OCR Reader & Visualization	22
Figure 4: Contact Page.....	23
Figure 5: Dashboard.....	23
Figure 6: Admin Page	24
Figure 7; Exported Data.....	24

LIST OF TABLE

Table 1 : Department Information	7
Table 2: Weekly Log	13

LIST OF ABBREVIATIONS

API	-	Application Programming Interface
BCA	-	Bachelor in Computer Application
CMS	-	Content Management System
CRUD	-	Create, Read, Update, Delete
CSS	-	Cascading Style Sheets
DJANGO	-	A high-level Python web framework
HTTP	-	HyperText Transfer Protocol
PYTHON	-	high level language
OCR	-	optical character reader
REST	-	Representational State Transfer
UX	-	User Experience

CHAPTER I

INTRODUCTION

1.1 Introduction

This internship focuses on the design and development of specialized web applications and data automation tools while serving as a Backend Developer and Data Analyst. The main objective of this internship is to gain practical experience in the full Software Development Life Cycle (SDLC) by contributing to real-world projects that address digital communication and automated data management challenges using modern technologies like Python and Django.

The first phase of the internship involved the development of Nilkantha Press, a web-based platform designed for streamlined contact management and digital presence. Following this, a specialized data management application engineered to digitize physical records through advanced OCR (Optical Character Recognition). This system enables organizations to extract information from complex documents, such as voter lists and IDs, and automatically generate structured analytical reports, including voter demographics and registration trends. By automating Nepali Unicode extraction and maintaining structured records, the system reduces human error, ensures data accuracy, and enhances the reliability of administrative data.

This internship is highly valuable because it provides hands-on experience in building diverse business applications ranging from standard web portals to high-stakes data extraction tools. It has facilitated the development of critical skills in backend engineering, OCR pipeline integration, data visualization, and API validation. Furthermore, it demonstrates how technology can bridge the gap between legacy paper-based records and modern digital systems, improving efficiency and informed decision-making in organizational data management.

1.2. Problem Statement

Many organizations and businesses in Nepal still rely on manual or semi-digital methods, such as physical ledgers or simple spreadsheets, to manage their data and registration processes. These traditional approaches are often time-consuming, error-prone, and difficult to maintain, especially when dealing with complex scripts like Nepali Unicode. Manual data entry from physical documents like voter IDs or handwritten records frequently leads to transcription mistakes, inconsistent records, and significant challenges in tracking demographic or financial performance over time.

In addition, preparing essential analytical reports such as voter demographic summaries or registration trends usually requires massive manual effort to aggregate data from paper files. This process delays the availability of critical information, which negatively affects timely and informed decision-making for administrative bodies. Many organizations also lack a centralized digital system that ensures secure data storage, consistent record management, and real-time access to information across different departments.

Therefore, there is a strong need for an automated, reliable, and secure system that utilizes OCR technology to efficiently digitize physical records. By reducing human error through automated field extraction and providing dynamic data visualization, such a system helps organizations improve efficiency, maintain accurate records, and support effective, data-driven decision-making.

1.3 Objectives

1.3.1 General Objectives

- To design and implement a structured django-based web application for efficient contact management and digital presence.
- To develop a specialized Python-based tool that uses OCR pipelines to transition physical records (like voter IDs) into digital formats without manual entry and visualize it.

1.3.2 Specific Objectives

- To develop a fully functional competition platform python, django and its libraries to ensure sustainable and proper workflow efficiently.

1.4 Scope and Limitation

1.4.1 Scope

The scope of this internship includes but is not limited to:

- Focused on building a structured Django-based platform to manage organizational contacts and digital content through a secure backend system.
- Enabled the efficient collection and processing of physical records—such as voter IDs—by using specialized OCR pipelines to extract Nepali Unicode text.
- Supported the automatic generation of analytical reports, including Age and Gender Demographics, visualized through dynamic pie charts and bar graphs.
- Provided a robust "Import to Excel" feature using openpyxl, allowing for the seamless conversion of extracted OCR data into audit-ready spreadsheets.
- Ensured secure data storage and reliable API communication, verified through Postman to maintain 100% data integrity between the extraction engine and the reporting modules.
- Delivered a user-friendly, web-accessible platform that allows administrators to manage records and view registration trends from any location with an internet connection

1.4.2 Limitation

The internship has several limitations:

The application requires a stable internet connection for Django backend services and API communication, which may limit the real-time processing of OCR tasks in areas with poor connectivity.

- The accuracy of the Nepali Unicode extraction is highly dependent on the quality of the uploaded images or PDFs; low-resolution or heavily blurred documents may require manual verification.
- The project was for academic purpose which resulted in learning curve.
- Time constraint because of the limited time the project weren't able to go to rigorous testing.

1.5 Report Organization

This report is organized into five main chapters:

Chapter I: Introduction - Provides an overview of the internship project, problem statement, objectives, scope and limitations, and report organization.

Chapter II: Introduction to Organization - Describes sarbatra company, its organizational hierarchy, working domains, and internship department details.

Chapter III: Background Study and Literature Review - Covers fundamental concepts related to the internship project including development and data analysis, with technologies such as django, python libraries.

Chapter IV: Internship Activities - Details roles and responsibilities performed during the internship, provides a weekly log of technical activities based on git commits, describes projects involved, and lists tasks and activities performed.

Chapter V: Conclusion and Learning Outcomes - Summarizes work completed and provides learning outcomes.

CHAPTER II

INTRODUCTION TO ORGANIZATION

2.1 Organization Details

Sarbatra Inc. Pvt. Ltd. is a prominent digital solutions provider based in Koteshwor, Kathmandu. Established as a "one-stop" hub for digital transformation, the company specializes in bridging the gap between innovative technology and business needs. Under the leadership of founders like Ashish Dahal, Sarbatra Inc. has evolved from a local startup into a trusted brand serving both national and international clients. The organization is recognized for its commitment to quality, scalability, and fostering local talent through job-ready internship programs.



Figure 1: Company Logo

2.2 Organization Hierarchy

- **CEO:** Sets the company's strategic vision and drives long-term business growth by making high-level decisions that align the organization's mission with market opportunities.
- **Managing Director:** Focuses on overall business strategy, organizational leadership, and decision making. Responsible for managing operations, guiding company growth, and ensuring that business objectives are achieved efficiently.
- **HR:** Manages employee recruitment, handles staff needs, and maintains a positive work environment.
- **Development Team:** Builds, tests, and maintains the software or products of the company.

- **Design Team:** Creates visual designs and improves the look and user experience of products.
- **Technical Support:** Assists users by solving technical issues and providing help with systems or products.
- **Team:** A group of individuals who work together, share responsibilities, and collaborate to achieve common goals and complete tasks effectively.

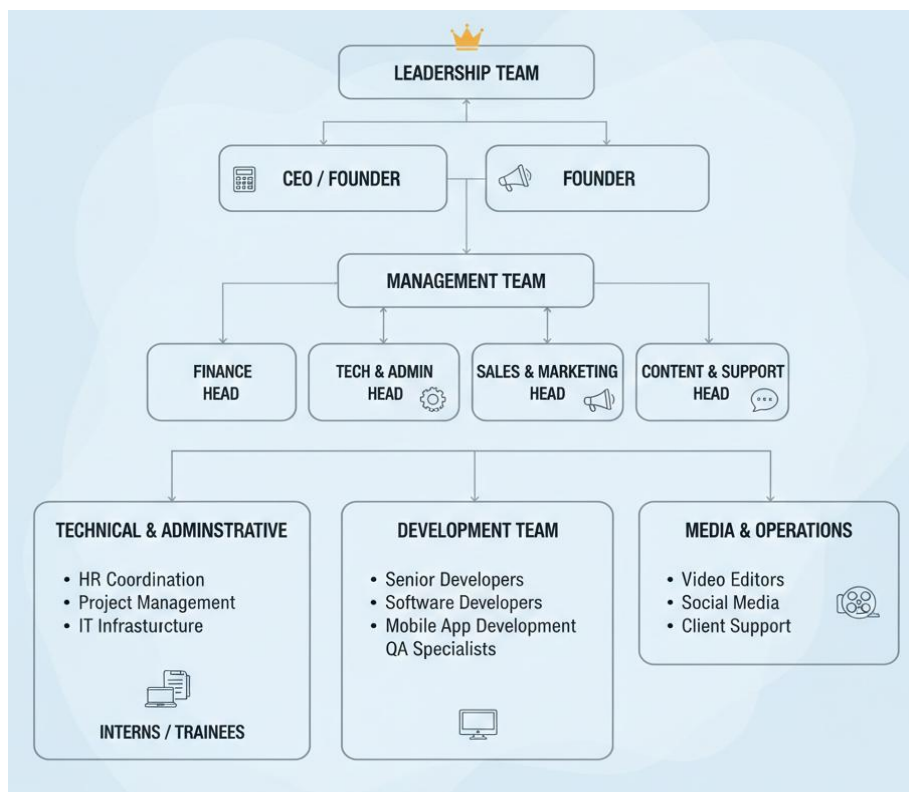


Figure 2: Organizational Hierarchy

2.3 Working Domains of Organization

Sarbatra Inc. Pvt. Ltd. operates across several specialized digital domains, focusing on high-quality delivery and innovation. The primary working domains are:

- **Web Development:** Designing and developing scalable, high-performance websites and enterprise web applications.
- **Mobile App Development:** Creating intuitive, cross-platform mobile applications for Android and iOS using modern frameworks like Flutter.

- **CMS:** Providing specialized, industry-specific CMS solutions for the travel, medical, and academic sectors.
- **Digital Marketing & SEO:** Offering a full suite of marketing services including SMM, Social Signals, and Search Engine Optimization.
- **Bulk SMS & Communication:** Managing large-scale business communication through integrated Bulk SMS platforms with local telecom providers.

2.4 Description of Intern Department/Unit

During the internship at Sarbatra Inc. Pvt. Ltd., placement was made within the Software Development department. This department is considered the technical backbone of the organization, where the focus is placed on building robust, scalable software solutions such as the Nilkantha Press web portal and the data extraction platform. A collaborative, agile environment is maintained within the unit under the direct mentorship of senior developers and the founder, Ashish Dahal.

A "client-first" approach is emphasized by the department, meaning real-world business logic and feature implementations were carried out that directly impacted end-users. The bridging of physical documentation with digital accounting was a key focus of this unit, and challenges related to Nepali Unicode processing and automated data analysis were addressed.

Department Information

Table 1 : Department Information

Particular	Time
Internship Start Date	2025/12/12 To 2026/3/12
Office Hours	11:00 AM to 8:00 PM
Mentor	Ashish Dahal
Supervisor	Mr. Santosh Pokherel (College)
Working Mode	Independent with mentor guidance

CHAPTER III

BACKGROUND STUDY AND LITERATURE REVIEW

3.1 Background Study

The first project, the Nilkantha Press website, was developed using the Django framework, with a strong emphasis on creating a scalable, maintainable, and user-friendly backend system. The technical foundation of the platform centered around leveraging Python's flexibility alongside Django's built-in administrative interface to simplify content and data management. This architecture was intentionally designed to be lightweight yet robust, enabling efficient handling of publishing-related data while ensuring that non-technical administrators could easily update, correct, and manage records without requiring developer intervention. Special attention was given to designing intuitive data models and streamlined workflows so that the system could support real-world operational needs in a practical and accessible manner. By prioritizing usability alongside performance, the platform successfully bridged the gap between technical complexity and everyday administrative tasks, making it both efficient for developers and approachable for end users.

In contrast, the second project shifted focus toward the domain of Computer Vision and automated data analysis, where the goal was to transform unstructured physical documents into structured, analyzable datasets. This system was built using Python-based OCR technologies such as Tesseract OCR and PaddleOCR, with a particular emphasis on accurately processing Nepali Unicode text an area that presents unique challenges due to script complexity and inconsistent print quality. The backend logic for this project was significantly different, focusing on intelligent field extraction and data parsing techniques to convert raw images, such as voter ID cards, into clean, structured records. Beyond extraction, the pipeline extended into data analysis and visualization, where processed information was transformed into meaningful insights through dynamic visual representations and dashboards. Additionally, the system supported exporting processed datasets into CSV format, enabling further use in external analytical tools. This project required careful handling of data accuracy, preprocessing, and validation to ensure reliable outputs, ultimately demonstrating how computer vision and automation can streamline traditionally manual and error-prone workflows into efficient, scalable digital solutions.

3.2 Literature Review

Modern software development emphasizes efficiency, scalability, and reliability, positioning Python as a leading choice for backend engineering due to its readability, extensive ecosystem, and strong community support (Rasulov, 2024; Srivastava, 2022). Research highlights that frameworks such as Django significantly streamline the development of data-driven applications by offering built-in modules for authentication, database ORM (Object Relational Mapping), session handling, and administrative interfaces (Holovaty & Kaplan-Moss, 2009; Chen et al., 2020). These features reduce development overhead while enforcing standardized design patterns that improve long-term maintainability (Idris et al., 2021). In the context of the Nilkantha Press platform, Django enabled a modular and scalable architecture where components such as contact management, content updates, and communication workflows could be handled efficiently within a unified system (Django Software Foundation, 2025). Furthermore, Django's auto-generated admin panel proved especially valuable in creating a system accessible to non-technical users, aligning with studies that emphasize the importance of usability and accessibility in administrative software systems (Chen et al., 2020). This approach reflects broader findings in web engineering literature that developer productivity and system reliability are significantly enhanced when using convention-driven frameworks that abstract repetitive backend tasks (Holovaty & Kaplan-Moss, 2009).

The development and implementation of RESTful API architecture serves as a foundational principle in modern web systems, enabling seamless communication between distributed components. APIs built using Python-based frameworks facilitate efficient data exchange through standardized HTTP methods and lightweight data formats such as JSON (Django REST Framework, 2025), making them suitable for both simple web applications and complex analytical tools. In the OCR and Data Analysis project, RESTful APIs played a critical role in transmitting extracted Nepali Unicode text and structured demographic data between the backend processing pipeline and the frontend visualization layer. This ensured interoperability across modules while maintaining data consistency and system responsiveness. Additionally, well-designed APIs reduce coupling between system components, allowing independent development and scaling of services an approach that aligns with the modular design of both the Nilkantha Press platform and the OCR-based data processing system (Idris et al., 2021).

For API testing and validation, tools like Postman are extensively documented as essential for ensuring backend reliability and correctness. Postman enables developers to isolate API endpoints, simulate requests, and verify responses independently of frontend systems, which is particularly useful in complex, multi-module architectures. Systematic API testing reduces integration errors and improves overall system robustness by identifying inconsistencies early in the development cycle. During the internship, this approach was crucial for debugging both the OCR data extraction pipeline and the CSV export functionality, where accurate data transmission and formatting were essential. Additionally, exporting structured data into formats such as CSV aligns with established data interoperability standards, allowing seamless integration with external tools like spreadsheet software and statistical analysis platforms. Rigorous testing practices, combined with modular system design, lead to higher reliability and reduced maintenance costs (Chen et al., 2020; Idris et al., 2021). Collectively, these practices ensured that both projects from simple content management in Nilkantha Press to complex demographic data analysis delivered outputs that were accurate, consistent, and dependable in real-world usage scenarios.

CHAPTER IV

INTERNSHIP ACTIVITIES

4.1 Roles and Responsibilities

4.1.1 Roles

- Served as a Backend Engineer and Data Processing Specialist.
- Oversight and coordination of the full lifecycle of application features were performed.
- Development and optimization of backend architecture were carried out to maintain stable and efficient operations.
- Accuracy and reliability of data extraction, transformation, and processing were ensured across all modules.
- Performance, scalability, and system robustness were continuously monitored and enhanced during implementation.

4.1.2 Responsibilities

- A clear distinction between the two projects was established in terms of roles and contributions.
- End-to-end development of the Nilkantha Press web portal was carried out, with Django backend logic implemented to manage user contacts, site content, and administrative features.
- A specialized OCR pipeline was designed and implemented for the data extraction tool, enabling the conversion of physical records, such as voter IDs, into structured digital formats using Python libraries.
- Backend code was written to be maintainable and scalable, with a focus on optimizing the speed and accuracy of Nepali Unicode extraction and image processing.

An integrated dashboard was developed to transform raw extracted data into actionable insights, including visual reports such as demographic pie charts and registration bar graphs.

Postman was used to independently test backend APIs for both the web portal and the OCR tool, ensuring JSON data integrity across all endpoints.

API request-response cycles were monitored, and HTTP status codes were validated to ensure robust communication between the Django backend and frontend visualization modules.

Technical bugs across both projects were identified, documented, and tracked throughout the development lifecycle.

Root Cause Analysis (RCA) was performed to trace errors from the user interface back to specific Python functions or Django modules, ensuring accurate final Excel exports.

Database interactions were optimized and backend code refactored to improve the reliability and performance of data-heavy extraction features.

4.2 Weekly Log

Table 2: Weekly Log

Weeks	Phase	Agenda / Discussion Points	Progress / Outcomes
Week 1	Phase 1	- Transition Management - Architecture Analysis	- Joined intern at sarbatra (DEC 12) - Studied the transition from the Nilkanthapress.com deployment to the development of new internal data extraction tools.
Week 2	Phase 1	- OCR Pipeline Integration - Unicode Processing	- Enhanced a separate toolset by integrating OCR pipelines (using PaddleOCR and Tesseract) to read physical documents like voter IDs. - Refactored Python scripts to improve the accurate extraction of Nepali Unicode text from complex voter lists.
Week 3	Phase 1	- Analysis Engine Development - System Stability	Continued the development of the analysis engine to process and categorize extracted voter data.
Week 4	Phase 1	Environment Setup	Analyzed the backend requirements for the Django-based site architecture, focusing on contact management and site-wide communication

Week 5	Phase 1	Set up further environments	Configured user authentication and environment variables for the standalone web project to ensure secure administrative access.
Week 6	Phase 1	Backend Optimization	Developed and optimized the backend logic for the web application's core features, ensuring efficient data handling.
Week 7	Phase 1	API Refinement	Updated REST API endpoints to facilitate seamless communication between the new site's frontend and the Django backend.
Week 8	Phase 1	development Feature Engineering	Implemented a "Field Extraction Module" using AI models to structure raw OCR.
Week 9	Phase 2	Data Flow Optimization	Maintained the stability of the Django backend while managing large-scale data uploads and extraction requests.
Week 10	Phase 2	Data Flow Optimization	- Performed comprehensive testing of the "Import to Excel" feature using openpyxl to ensure correct data mapping from the OCR results. - Developed finalized REST API documentation for both the

			Nilkantha Press web project and the Data Extraction tool.
Week 11	Phase 3	• Data Visualization	• Added visualization features to the dashboard, creating pie charts for demographic distribution and bar graphs for registration trends. • Ensured data integrity by performing rigorous validation of extracted results against the original scanned images.
Week 12	Phase 3	• Production Refactoring • Quality Assurance	• Developed finalized REST API documentation for both the Nilkantha Press web project and the Data Extraction tool. • Refactored code across both projects to meet professional production standards for readability and maintainability.

4.3 Description of Project Involved During Internship

Project Name: Project Name: Data analysis and backend development

Platform: Web-based Data Extraction & Analysis Tool

Core Purpose: An automated solution was designed to digitize physical records and provide data-driven insights through advanced OCR and backend processing, thereby bridging the gap between legacy paper records and modern digital accounting.

Project Details: The application is designed as a specialized data management tool in which information is extracted from complex physical documents to generate structured digital records and analytical reports. The system is engineered for high-stakes environments, such as voter registration and administrative auditing, where data accuracy is considered critical. During the internship, focus was placed on the backend infrastructure and system stability of the following modules:

- **Automated Data Extraction:** Developed a specialized OCR pipeline to accurately extract Nepali Unicode text and unique identifiers from scanned voter documents and IDs using Python libraries.
- **Demographic Analysis:** Engineered backend logic to process extracted data and generate automated Age and Gender Demographics summaries for administrative review.
- **Data Visualization:** Built integrated dashboard features to visualize trends through dynamic pie charts (e.g., Gender & Relationship Distribution) and bar graphs (e.g., Registration Trends by District).
- **Reporting & Compliance:** Implemented a robust "Import to Excel" module using openpyxl to ensure extracted financial and demographic data is exported.

The role was to ensure that the logic for text extraction and data categorization was made resilient, maintaining high accuracy from the raw image input to the final exported report through rigorous Django development and API validation.

4.4 Activities Performed

In the capacity of Backend Engineer, primary focus was placed on the stabilization and modernization of the Nilkantha Press web portal. Using the Django Framework, a systematic overhaul of the site's core logic was conducted to ensure it could handle increasing traffic and complex API integrations.

One significant challenge addressed was maintaining a backend that remained decoupled yet highly performant. Robust system modules were constructed to facilitate clean communication between the database and the frontend. Modifications to the existing system architecture ensured that API integrations, ranging from user authentication to content delivery, were handled with strict validation and minimal latency. This transformed the portal from a static site into a dynamic, scalable platform capable of supporting real-time data flow, particularly for contact modules requiring immediate database commitment.

Parallel work included the development of a specialized Data Extraction and Analysis platform, centered on the implementation of Optical Character Recognition (OCR) pipelines. A modular engineering approach was prioritized to prevent "spaghetti code" arising from the variety of document formats.

Clean, reusable Python modules were developed to separate image pre-processing, text recognition, and post-processing stages. This modular design allowed multiple OCR engines and AI models to be integrated without disrupting the workflow. Component isolation improved testability and iterative development, particularly during complex field-mapping where specific headers and rows in scanned images were converted into structured formats.

A major technical achievement was the optimization of Nepali Unicode processing. Backend logic was refined to handle document images containing heavy Nepali text. Image-cleansing techniques, such as noise reduction and contrast adjustment, were implemented, and data-heavy report processing was fine-tuned to improve both extraction speed and accuracy. Linguistic integrity was preserved for official Nepali reporting and compliance with local regulatory standards.

RESTful APIs were designed and implemented to serve as the "nervous system" for both the Nilkantha Press portal and the data extraction dashboard. These APIs facilitated seamless communication between Django-managed data and frontend interfaces. Modules

for contact management and automated notifications were created for the Nilkantha Press project.

Specialized libraries were integrated for data visualization, allowing graphical representations of extracted data to be displayed directly within the dashboard. Trends over time and comparisons of extracted fields were made visually interpretable.

Export to CSV and Import to Excel modules were developed. Data structures were designed to ensure exported files remained aligned with database fields. Libraries such as openpyxl were used to manage Excel formatting, producing professional, error-free outputs suitable for business use.

A test-first methodology was adopted to ensure system reliability. API endpoints were subjected to rigorous protocol validation using Postman, verifying that OCR-extracted data matched expected frontend structures. Handling of incomplete document scans, low-resolution images, empty fields, and export latency under heavy server load was also tested.

Maintenance tasks included Root Cause Analysis (RCA) of persistent Django backend issues. Problems within the services layer, particularly in data mapping between OCR output and SQL databases, were identified and resolved. These efforts stabilized site functionality and ensured the accuracy of field extraction logic, significantly reducing data categorization errors.

A Performance Refactoring phase was initiated to optimize handling of large datasets. Backend code was optimized using advanced Django query methods, such as `select_related` and `prefetch_related`, reducing database hits and improving page load times. System responsiveness was maintained even when processing or exporting thousands of rows of analyzed data.

CHAPTER V

CONCLUSION AND LEARNING OUTCOMES

5.1 Conclusion

The internship at Sarbatra Inc. Pvt. Ltd. was a highly rewarding professional experience. During this period, work was conducted on data analysis and backend development, with primary focus placed on backend development, data analysis, and automated data extraction. The internship provided the opportunity for theoretical concepts learned in academic studies to be applied to a real-world enterprise application.

Under the guidance of mentor Ashish Dahal, contributions were made to the development of backend systems using Python and Django, and a specialized module for OCR-based data reading was engineered. Raw datasets were transformed into actionable insights through the creation of complex visualizations, such as graphs and pie charts, ensuring clarity and accessibility of the application's data. These tasks were directly aligned with the objectives of the internship, including understanding the software development lifecycle, improving backend functionality, and enhancing data accessibility through systematic analysis.

One of the most important lessons derived from this experience was the critical importance of data accuracy in financial systems, where even minor extraction errors can significantly affect financial reports, such as balance sheets and tax calculations. Technical skills were developed in Django, OCR integration, and Python-based data visualization, while analytical thinking and problem-solving abilities were strengthened.

Overall, the internship successfully fulfilled its objectives by providing practical exposure to enterprise software development and data-driven decision-making. The knowledge and skills gained while working on data analysis and backend development have prepared for future challenges in the field of software engineering and backend development.

5.2 Learning Outcomes

During the internship at Sarbatra Inc. Pvt. Ltd., valuable practical experience was gained in backend development and data analysis while working on data analysis and accounting backend systems.

Hands-on experience was acquired in developing and modifying backend business logic and architecture using the Django web framework. RESTful APIs for data-driven modules were designed, implemented, and updated, including request validation, HTTP methods, and structured error responses.

Specialized modules were engineered to read and process raw data from documents and images using OCR technologies, streamlining the data entry process. Scripts were developed to analyze complex datasets, and visual representations, including graphs and pie charts, were created to provide clear insights into system data.

Queries were optimized, and relational data was managed using the Django ORM to ensure efficient storage and retrieval of processed information. Issues within backend services and data pipelines were traced and resolved, improving the overall reliability of the extraction and reporting tools.

Backend logic and data processing scripts were enhanced to handle large datasets efficiently, reducing the time required to generate visual reports. Strict validation and error handling were implemented to ensure that data extracted via OCR remained accurate and consistent across the platform.

An understanding of how backend development and data analysis fit within the Software Development Life Cycle was gained, including version control and iterative system improvements. Collaboration with mentors and team members was maintained to resolve technical challenges and enhance the functionality of the data extraction and visualization tools.

REFERENCES

Chen, S., Ahmmed, S., Lal, K., & Deming, C. (2020). Django web development framework: Powering the modern web. *American Journal of Trade and Policy*, 7(3), 99–106.

Django REST Framework. (2025). *Django REST framework documentation*.

Django Software Foundation. (2025). *Django documentation*.

Holovaty, A., & Kaplan-Moss, J. (2009). *The definitive guide to Django: Web development done right*. Apress.

Idris, N., Mohd Foozy, C. F., & Shamala, P. (2021). A generic review of web technology: Django and Flask. *International Journal of Advanced Science Computing and Engineering*, 2(1), 34–40. <https://doi.org/>

Rasulov, H. R. (2024). Django: A comprehensive framework for Python web development. *International Journal of Scientific Researchers*, 9(1)

Srivastava, A. K. (2022). Django: The Python web framework. *International Journal of Scientific Research in Engineering and Management*, 6(5).

APPENDICES

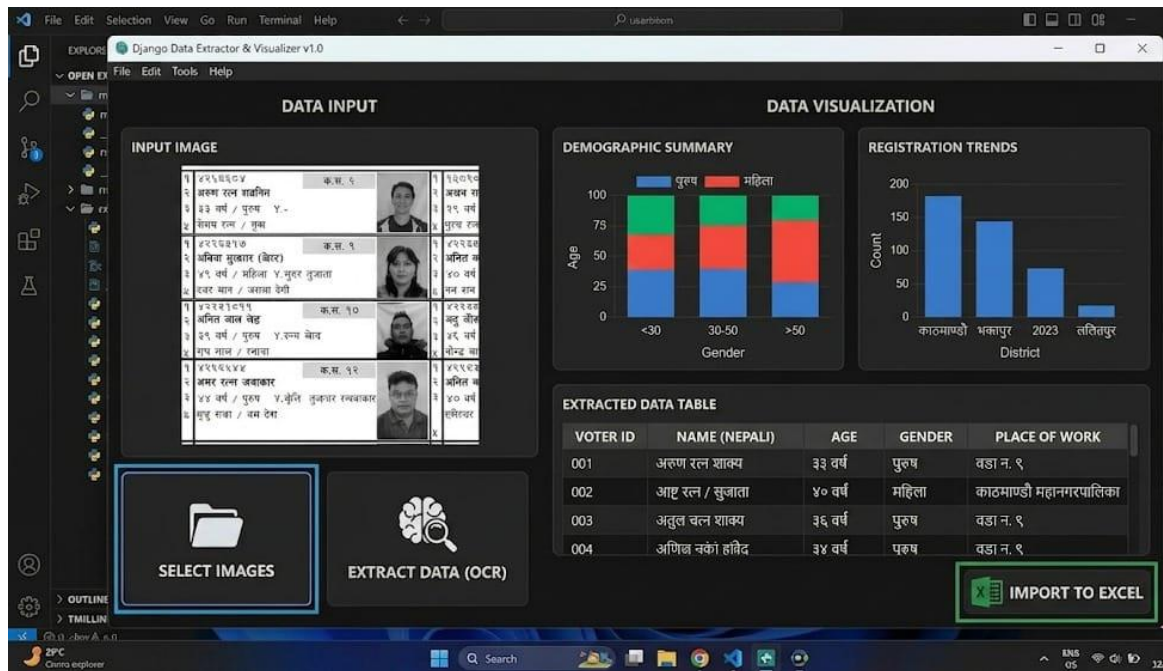


Figure 3: OCR Reader & Visualization

Send Your Message

Your Name *

Email *

Phone Number

Subject *

Message *

[Send Message](#)

Contact Information

NEW NILKANTHA OFFSET PRESS

Address
Nagagan-5, New Road,
Kathmandu, Nepal

Phone
+977-01-5103216
+977-01-5103217

Email
info@nilkanthapress.com

Business Hours
Sunday - Friday: 10:00 AM - 6:00 PM
holiday

Customer Support
Get help with your orders and
technical support
[Call Now](#)

Sales Inquiry
Discuss your printing
requirements with our sales team
[Email Us](#)

Find Our Location
Located in the heart of Kathmandu, easily accessible from all parts of the city

Figure 4: Contact Page

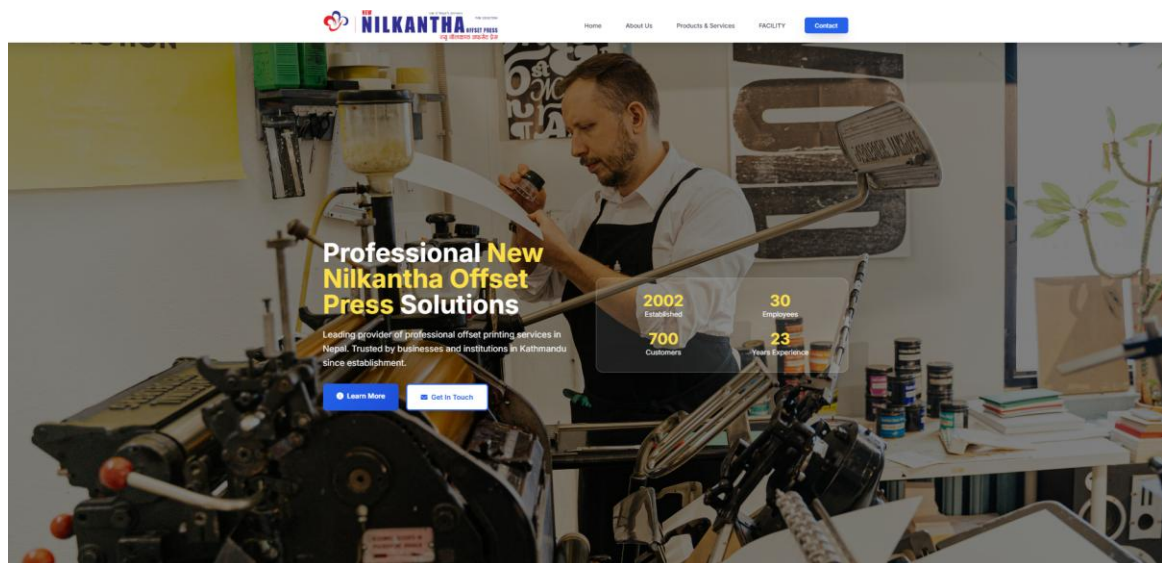


Figure 5: Dashboard

Welcome to Nilkantha Offset Press Content Management System

AUTHENTICATION AND AUTHORIZATION

Groups

Add

Change

Users

Add

Change

CORE

About Page CMS

Change

Contact Messages

Add

Change

Contact Page CMS

Change

Facility Page CMS

Change

Footer Items

Add

Change

Footer sections

Add

Change

Home Page CMS

Change

Navigation Links

Add

Change

Newsletter Subscribers

Add

Change

Service Items

Add

Change

Services

Add

Change

Services Page CMS

Change

Site Media

Add

Change

Site Settings

Change

Social links

Add

Change

Recent actions

My actions

About Page Content

About Page CMS

About Page Content

About Page CMS

Site Settings

Site setting

Home Page Content

Home Page CMS

Footer_logo

(ChatGPT_image_jan_28_2026_08_27_PM_v2XRCC7.png)

Site Media

Footer_logo

(ChatGPT_image_jan_28_2026_08_32_49_PM_vkGETv4.png)

Site Media

Footer_logo (Nilkantha-logo_45nNp4U_mcd4MUf.png)

Site Media

Footer_logo (ChatGPT_image_jan_29_2026_01_10_44_PM.png)

Site Media

Footer_logo (ChatGPT_image_jan_28_2026_08_32_49_PM.png)

Site Media

Footer_logo (ChatGPT_image_jan_28_2026_08_28_27_PM.png)

Site Media

Figure 6: Admin Page

Django Voter Data Extract extractor v1.0

File Edit Selection View Go Run Terminal Help

AutoSave 100

Voter_Data_Export.xlsx - Excel (Django-Extracted)

Userbilbon

Share

Share

INPUT IMAGE

SELECT IMAGES

EXTRACT DATA (OCR)

Export Successful: Voter_Data_Export.xlsx (12 Voters)

ID	PHOTO	NAME (NEPALI)	AGE (वर्ष)	GENDER	PLACE OF WORK	REGISTRATION DISTRICT	SERIAL NO.
001		अतुल धन शास्त्र	३३ वर्ष	पुरुष	काठमाण्डौ महानगरपालिका	काठमाण्डौ भाकापुर	1
002		अनघा मल्ल	३४ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	2
003		अतुल चन शास्त्र	३४ वर्ष	महिला	वडा ना न. ९	काठमाण्डौ भाकापुर	3
004		अश्वि कुल शास्त्र	३४ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	4
005		अनित देव सावर	३२ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	5
006		अनिता बाता बोदेत	३९ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	6
007		अमिता सला जाववर	३२ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	7
008		अतुल खन शास्त्र	३४ वर्ष	पुरुष	काठमाण्डौ महानगरपालिका	काठमाण्डौ भाकापुर	8
009		असुत वसी शास्त्र	३४ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	9
010		अनिता मल्ल	३४ वर्ष	पुरुष	वडा ना न. ९	काठमाण्डौ भाकापुर	10

Figure 7: Exported Data

