zapcom.

INTERNSHIP REPORT

ADARSH GADEKAR
2065

Zapcom Solutions Pvt Ltd.

Goa University

zapcom.

1st June 2023

TO WHOMSOEVER IT MAY CONCERN

This is to inform you that **Mr. Adarsh Gadekar,** student of Master of Computer Applications (MCA) of Goa University, Goa, is currently undergoing his final semester project (Semester VI/V) at our company, **Zapcom Solutions Pvt. Ltd** from 4th January, 2023.

During his tenure he has met the expectations of his team lead/mentor/guide and found to be regular and sincere.

This letter is being issued on his request to be submitted with the project report at Goa University.

The final internship completion letter will be provided on completing his internship.

For Zapcom Solutions Pvt. Ltd.

Zapcom Solutions Pvt. Ltd. INDIA

Srinivas Reddy Kothakota
Chief Operating Officer

INDIA
Zapcom Solutions Pvt. Ltd
9th Floor, Gamma Tower, Sigma Soft Tech Park,
Whitefield, Bangalore - 560066
Ph: +91-80-67232300

www.zapcg.com

USA
Zapcom Group Inc.
105 Decker Court, Ste. 810 Irving,
TX 75062.
Ph: (972)441-2081

REPORT OF INTERNSHIP DONE AT ZAPCOM SOLUTIONS PVT. LTD.

SUBMITTED BY

ADARSH GADEKAR

2065

UNDER THE GUIDANCE OF

Mr. Suburamanian
(Delivery Manager, Zapcom)

Mr. Manjunath Sirur
(Data Scientist, Zapcom)

Sai Conda
(CTO, Zapcom)

GOA UNIVERSITY

GOA BUSINESS SCHOOL

CERTIFICATE OF EVALUATION

This is to certify that **Mr. Adarsh Gadekar** has been evaluated for the project work titled "**Report of Internship done at Zapcom Solutions Pvt Ltd.**" undertaken at **Zapcom Solutions Pvt. Ltd. , Bangalore** in partial fulfillment for the award of the degree in Master of Computer Application.

Examiner 1

Examiner 2

Place: Goa University

Date: 16th June 2023

Dean, Goa Business School

Acknowledgement

I am immensely grateful to Mr. Kishore Pallamereddy, the Executive Chairman & CEO of Zapcom, and Zapcom Solutions Pvt Ltd., for granting me the incredible opportunity to intern at their esteemed organization. This internship has been a transformative experience, allowing me to gain invaluable insights into the world of Information Technology.

I would also like to express my deepest appreciation to Mr. Suburamanian A, the Delivery Manager at Zapcom, and Mr. Manjunath Sirur, Data Scientist, for their exceptional mentorship and unwavering guidance throughout my internship. Their expertise, patience, and support have been truly invaluable, helping me navigate challenges and grow both personally and professionally.

Special thanks go to Mr. Manjunath Sirur, the Data Scientist at Zapcom, for his invaluable help and guidance in the technologies I worked on during my internship. His expertise and willingness to share knowledge have been instrumental in my skill development.

I would also like to acknowledge the immense contribution of Mr. Ramrao S. Wagh, Associate Professor and Head of Department of Computer Science and Technology, MCA at Goa Business School, Mr. Ramdas Karmali, Assistant Professor of Computer Science and Technology, MCA at Goa Business School, Goa University, and Mr. Hanumant Redkar, an Assistant Professor and TPO, MCA at Goa Business School, Goa University, as well as the entire faculty and non-teaching staff of the MCA department at Goa University. Their constant encouragement and support throughout my academic journey have laid a strong foundation for my professional success.

Finally, my heartfelt appreciation goes to the entire Zapcom family. The nurturing and inclusive environment provided by this exceptional team have played a significant role in shaping me into a more confident individual, ready to embrace any challenges that come my way.

I am immensely grateful to all who have contributed to my growth and success, and I look forward to paying forward the knowledge and experiences gained during this internship.

Adarsh Gadekar

INTERNSHIP REPORT

ADARSH GADEKAR

2065

Zapcom Solutions Pvt Ltd.

Goa University

REPORT OF INTERNSHIP DONE AT ZAPCOM SOLUTIONS PVT. LTD.

SUBMITTED BY

ADARSH GADEKAR

2065

UNDER THE GUIDANCE OF

Mr. Suburamanian A
(Delivery Manager, Zapcom)

Mr. Manjunath Sirur
(Data Scientist, Zapcom)

Sai Conda
(CTO, Zapcom)

Internship Letter

zapcom.

1st June 2023

TO WHOMSOEVER IT MAY CONCERN

This is to inform you that **Mr. Adarsh Gadekar,** student of Master of Computer Applications (MCA) of Goa University, Goa, is currently undergoing his final semester project (Semester VI/V) at our company, **Zapcom Solutions Pvt. Ltd** from 4th January, 2023.

During his tenure he has met the expectations of his team lead/mentor/guide and found to be regular and sincere.

This letter is being issued on his request to be submitted with the project report at Goa University.

The final internship completion letter will be provided on completing his internship.

For Zapcom Solutions Pvt. Ltd.

Zapcom Solutions Pvt. Ltd. INDIA

Srinivas Reddy Kothakota
Chief Operating Officer

INDIA
Zapcom Solutions Pvt. Ltd
9th Floor, Gamma Tower, Sigma Soft Tech Park,
Whitefield, Bangalore - 560066
Ph: +91-80-67232300

www.zapcg.com

USA
Zapcom Group Inc.
105 Decker Court, Ste. 810 Irving,
TX 75062.
Ph: (972)441-2081

GOA UNIVERSITY

GOA BUSINESS SCHOOL

CERTIFICATE OF EVALUATION

This is to certify that **Mr. Adarsh Gadekar** has been evaluated for the project work titled "**Report of Internship done at Zapcom Solutions Pvt Ltd.**" undertaken at **Zapcom Solutions Pvt. Ltd. , Bangalore** in partial fulfillment for the award of the degree in Master of Computer Application.

Examiner 1

Examiner 2

Place: Goa University

Date: 16th June 2023

Dean, Goa Business School

Acknowledgement

I am immensely grateful to Mr. Kishore Pallamereddy, the Executive Chairman & CEO of Zapcom, and Zapcom Solutions Pvt Ltd., for granting me the incredible opportunity to intern at their esteemed organization. This internship has been a transformative experience, allowing me to gain invaluable insights into the world of Information Technology.

I would also like to express my deepest appreciation to Mr. Suburamanian A, the Delivery Manager at Zapcom, and Mr. Manjunath Sirur, Data Scientist, for their exceptional mentorship and unwavering guidance throughout my internship. Their expertise, patience, and support have been truly invaluable, helping me navigate challenges and grow both personally and professionally.

Special thanks go to Mr. Manjunath Sirur, the Data Scientist at Zapcom, for his invaluable help and guidance in the technologies I worked on during my internship. His expertise and willingness to share knowledge have been instrumental in my skill development.

I would also like to acknowledge the immense contribution of Mr. Ramrao S. Wagh, Associate Professor and Head of Department of Computer Science and Technology, MCA at Goa Business School, Mr. Ramdas Karmali, Assistant Professor of Computer Science and Technology, MCA at Goa Business School, Goa University, and Mr. Hanumant Redkar, an Assistant Professor and TPO, MCA at Goa Business School, Goa University, as well as the entire faculty and non-teaching staff of the MCA department at Goa University. Their constant encouragement and support throughout my academic journey have laid a strong foundation for my professional success.

Finally, my heartfelt appreciation goes to the entire Zapcom family. The nurturing and inclusive environment provided by this exceptional team have played a significant role in shaping me into a more confident individual, ready to embrace any challenges that come my way.

I am immensely grateful to all who have contributed to my growth and success, and I look forward to paying forward the knowledge and experiences gained during this internship.

Adarsh Gadekar

Table of Contents

Introduction

This report serves as a concise overview of my on-site internship experience at Zapcom in Bangalore. I have been an active member of the Zapcom team since January 4th, 2023, and during my internship, I have gained valuable insights into the company's operations and contributed to various tasks related to intellectual property (IP) development.

In this report, I aim to provide essential information about Zapcom, including its organizational structure, work environment, and overall culture. I will also detail the responsibilities I undertook during my internship and highlight the online courses I completed to enhance my skills in the field.

Throughout my internship, I had the opportunity to work with a range of tools and technologies that are relevant to the data science field. I will discuss these tools and technologies and how they were utilized in the context of my internship tasks. Additionally, I will provide a timeline of my activities and progress to give a comprehensive overview of my internship experience.

Finally, I will conclude this report by reflecting on my overall experience at Zapcom and how it has contributed to my personal and professional growth. I will discuss the knowledge and skills I have acquired and outline how this internship has shaped my career aspirations in the field of data science.

Throughout this report, I will ensure grammatical accuracy, correct any spelling errors, and present the information in a clear and concise manner.

Company Profile

ZapCom is a dynamic company dedicated to the development and delivery of innovative products. With its ISO 9001 and 27001 certifications, ZapCom places a strong emphasis on maintaining high-quality standards and ensuring robust information security measures. The company's leadership team boasts an impressive collective experience in the Travel and Hospitality domain, through collaboration and partnership, the company aims to drive positive change and shape the future of the Travel, Hospitality, Fin-tech and Retail sectors..

`Zapcom are passionate about building digital products and platforms that can bend revenue and cost curves. They build, operate and optimize technology for their clients by taking a data-centric approach to creating products, platforms, and teams that drive delightful experiences and measurable business value.

The approach that defines Zapcom with their work include Solution co-creation, Elasticity, Security first, Metrics Driven, Automate, Product Mindset, and Customized Pods.

Data Science Internship Experience at ZapCom: Exploring Real-World Applications and Enhancing Skills

- Gain Hands-on Experience: The primary objective of a data science internship is to gain practical experience in working with real-world data sets, tools, and technologies commonly used in the field. This includes developing proficiency in programming languages such as Python or R, utilizing data manipulation and analysis libraries, and implementing machine learning algorithms.

- Apply Data Science Techniques: Interns are expected to apply various data science techniques, such as data cleaning, data visualization, feature engineering, and predictive modeling, to solve business problems or address specific research questions. This involves understanding the problem statement, formulating an approach, and implementing and evaluating the effectiveness of different algorithms.

- Collaborate in a Team Environment: Internships provide opportunities to work collaboratively with other data scientists, engineers, and domain experts. Interns are expected to actively participate in team discussions, contribute ideas, and collaborate effectively to achieve project goals. This objective aims to develop skills in teamwork, communication, and the ability to work in a professional environment.

- Learn from Mentorship: Internships often provide mentorship opportunities where experienced data scientists guide and support interns in their learning and project work. Interns are expected to seek guidance, learn from feedback, and demonstrate an eagerness to acquire new skills and knowledge under the mentor's guidance.

- Deliver Results and Insights: The ultimate goal of an internship is to deliver meaningful results and insights through data analysis. Interns should aim to provide valuable findings, actionable recommendations, or predictive models that can contribute to the organization's decision-making processes or research outcomes.

Tools and Technologies

Programming Languages:

Python: Python is a versatile and widely used programming language in data science. It offers a rich ecosystem of libraries and frameworks such as Pandas, NumPy, and Scikit-learn, which enable efficient data manipulation, analysis, and machine learning tasks. Python's readability and extensive community support make it a popular choice among data scientists.

SQL: SQL (Structured Query Language) is the standard language for managing and querying relational databases. Data scientists often utilize SQL to extract and manipulate data stored in databases, enabling them to perform data transformations, aggregations, and joins. Proficiency in SQL is essential for efficient data retrieval and analysis, especially when working with large datasets.

Data Manipulation and Analysis:

Pandas: Pandas is a powerful Python library used for data manipulation and analysis. It provides a wide range of data structures, such as DataFrames and Series, which allow for efficient handling and processing of structured data. Pandas offers functionalities for data cleaning, transformation, merging, filtering, and aggregation. Its intuitive syntax and built-in methods make it an essential tool for tasks like data preprocessing, exploratory data analysis, and feature engineering.

NumPy: NumPy (Numerical Python) is a fundamental library in Python for scientific computing. It provides support for efficient numerical operations on multidimensional arrays and matrices. NumPy offers a wide range of mathematical functions, linear algebra operations, and tools for array manipulation. Data scientists often use NumPy in combination with Pandas for data preprocessing and performing numerical computations required in machine learning algorithms.

SQL: SQL (Structured Query Language) is a language used to communicate with relational databases. It allows data scientists to retrieve and manipulate data stored in databases using queries. SQL offers functionalities for data selection, filtering, sorting, aggregation, and joining of tables. It is particularly useful when working with large datasets stored in databases and when performing complex data manipulations required for analysis.

Data Visualization:

Matplotlib: Matplotlib is a widely used data visualization library in Python. It provides a flexible and comprehensive set of tools for creating static, animated, and interactive visualizations. Matplotlib allows data scientists to generate a wide variety of plots, including line plots, scatter plots, bar plots, histograms, and more. It offers fine-grained control over visual elements, such as colors, labels, titles, and axes, enabling the creation of publication-quality visualizations.

Seaborn: Seaborn is a high-level data visualization library built on top of Matplotlib. It provides a simplified interface and a wide range of pre-defined themes and color palettes, making it easy to create visually appealing statistical graphics. Seaborn offers specialized plots for visualizing statistical relationships, such as scatter plots with regression lines, box plots, violin plots, and heatmaps. It simplifies the process of generating complex visualizations and enhances the overall aesthetics of the plots.

Machine Learning Libraries:

Scikit-learn: Scikit-learn is a popular machine learning library in Python. It provides a comprehensive set of tools and algorithms for various machine learning tasks, including classification, regression, clustering, dimensionality reduction, and model selection. Scikit-learn offers a user-friendly interface and a consistent API, making it easy to implement and evaluate machine learning models. It includes a wide range of preprocessing techniques, feature selection methods, and evaluation metrics. Additionally, scikit-learn supports integration with other libraries and frameworks, making it a versatile choice for machine learning tasks.

Cloud Platforms:

Amazon Web Services (AWS): AWS, short for Amazon Web Services, is a comprehensive cloud computing platform offered by Amazon. It provides a wide range of services and tools to facilitate the building, deployment, and management of applications and infrastructure in the cloud. During my internship, I had hands-on experience with one of its services called **Amazon RDS** (Relational Database Service) with **PostgreSQL**. Amazon RDS simplifies the setup and management of relational databases in the cloud, and using it allowed me to work with PostgreSQL, a robust open-source database management system, while benefiting from the scalability, security, and reliability provided by AWS.

Version Control:

Git and GitHub (Git Desktop): Git is a widely used distributed version control system that allows developers to track changes, collaborate, and manage codebase revisions efficiently. It provides features like branching, merging, and conflict resolution, enabling seamless collaboration among team members. Git Desktop is a user-friendly graphical interface for Git, simplifying the process of using Git commands and managing repositories.

Integrated Development Environments (IDEs):

Jupyter Notebook: Jupyter Notebook is an open-source web-based IDE widely used in data science and interactive computing. It supports multiple programming languages, including Python, R, and SQL, and provides an interactive and notebook-based interface for data analysis, machine learning, and visualizations.

Visual Studio Code: Visual Studio Code (VS Code) is a lightweight and extensible code editor developed by Microsoft. It offers features for coding, debugging, and version control integration, making it a versatile choice for developers, including data scientists.

pgAdmin 4: pgAdmin 4 is an open-source administration and development platform for PostgreSQL databases. It provides a web-based interface for managing, querying, and visualizing PostgreSQL databases.

Project: Exploring Data Science in the Field of Intellectual Property (IP) and OCR Automation

Title: Reward collection with OCR

Introduction:
During my internship, one of the key projects I worked on was the development of an Intellectual Property (IP) solution called Reward Collection with OCR. The project aimed to automate the process of collecting rewards by leveraging Optical Character Recognition (OCR) technology.
Traditionally, reward collection processes involved manual data entry, which was time-consuming and prone to errors. To address this challenge, the Reward Collection with OCR project aimed to streamline the process by automating the extraction of relevant information from reward images.

The use of OCR technology allowed for the extraction of text from images, enabling the automated retrieval of reward details such as customer names, reward codes, and expiration dates. By automating this process, the project aimed to enhance efficiency, reduce manual effort, and improve the overall reward collection experience.

My Role: Researching, Implementing, and Integrating AWS Textract for Document Scanning and Data Processing.
Throughout the project, my role was instrumental in researching, implementing, and integrating AWS Textract services API for document scanning and data processing. I took the lead in understanding the capabilities of AWS Textract and its integration with our project requirements. This involved conducting thorough research, exploring various OCR techniques, and selecting AWS Textract as the most suitable solution.

Moreover, I took charge of integrating AWS Textract with our back-end infrastructure, developing the necessary APIs and endpoints for seamless communication between the user interface and the OCR processing functionality. This integration enabled users to effortlessly upload reward images, initiate the OCR process, and retrieve the extracted reward information.

Furthermore, I designed and implemented a data storage system.

Tools and Technologies Utilized:

- **Version Control:** GitHub was used for version control, enabling efficient collaboration, code management, and tracking of project changes.

- **Programming Language:** Python served as the primary programming language for implementing various components of the project, including data processing, image preprocessing, and integration with AWS services.

- **Front-End Development:** Streamlit, a Python library, was employed for developing the user interface. It provided a seamless and interactive experience, allowing users to upload reward images and retrieve the extracted information with ease.

- **Data Storage:** PostgreSQL database served as the storage backend, ensuring efficient and secure

storage of the extracted reward information. It provided a reliable and scalable solution for managing the collected data.

- **AWS Integration:** Boto3, a Python library, played a vital role in integrating AWS services into the project. It facilitated communication with the AWS Textract services API, enabling the utilization of OCR capabilities for extracting text from reward images.

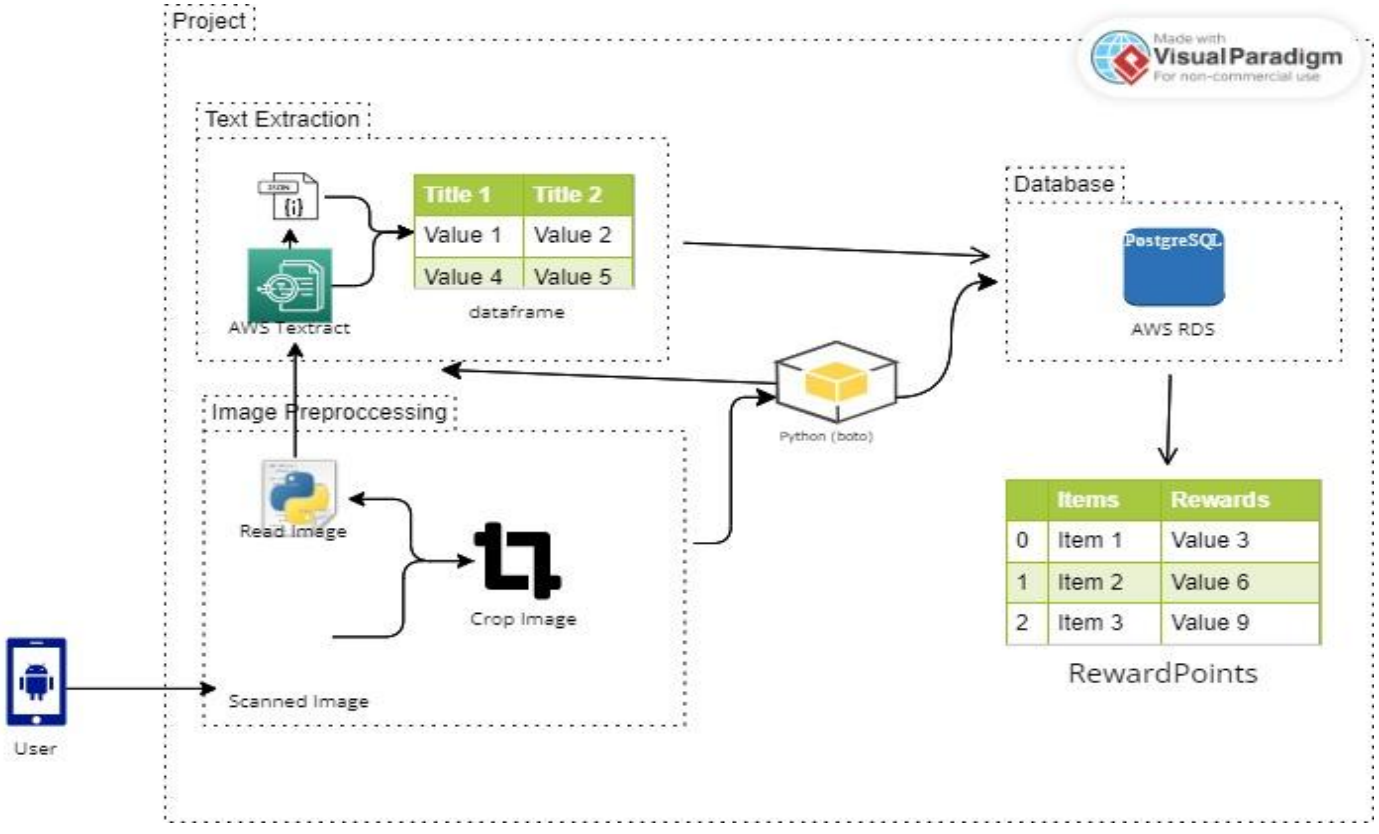
- **AWS Textract API:** The methods and functionalities offered by the AWS Textract API were extensively utilized for optical character recognition (OCR) on reward images. This allowed for the extraction of relevant information such as customer names, reward codes, and expiration dates from the uploaded images.

- **Integrated Development Environment (IDE):** Visual Studio Code was utilized as the IDE for developing and managing the project's codebase. Its powerful features and extensions provided an efficient development environment.

Business Objectives:

- OCR (Optical character recognition) - Putting OCR into right use and in multiple applications.
- Using OCR to convert unstructured data (Physical data) into structured format.
- Add-on feature to existing systems such as reward points collection by scanning a receipt, identifying a product from the receipt, and allocating a reward point for the purchase.
- Add-on feature to an existing ERP system, where the fields are pre populated using OCR once a receipt or a bill is uploaded

ARCHITECTURAL DIAGRAM:



Applicability Of Our Current Implementations:

- Implementing/Code reuse to the areas where NLP/OCR might be needed.
- Leveraging the system in reward collection platforms.
- Identifying data or specifics from a scanned document more precisely and present it as an add-on feature to an existing platform.

Goals and Future Scope:

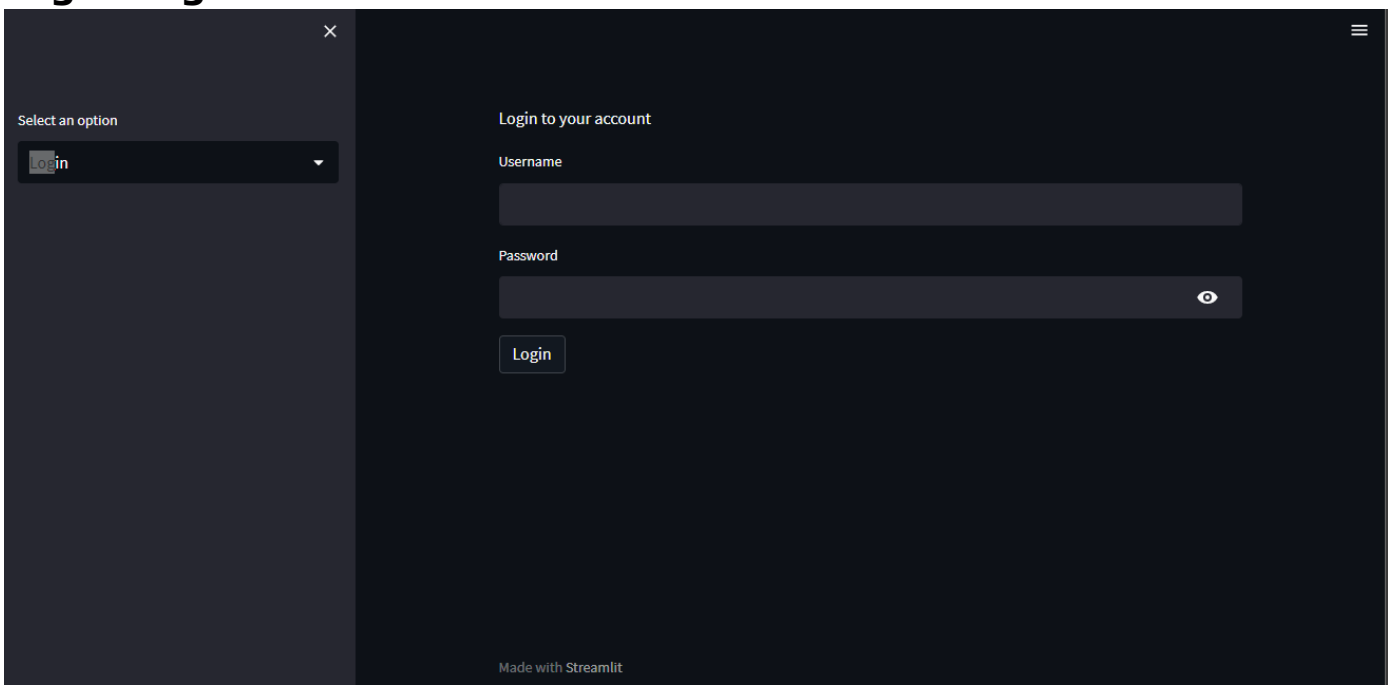
- Leveraging OCR to detect the type of liquor bottles and liquor specifications.
- Customer profile building, once a customer uploads a scanned copy of receipt, we can store and use the information for building customer profile using AI/ML.
- Building recommendation engines on the acquired data/purchase history to recommend products to the customer.

Summary:

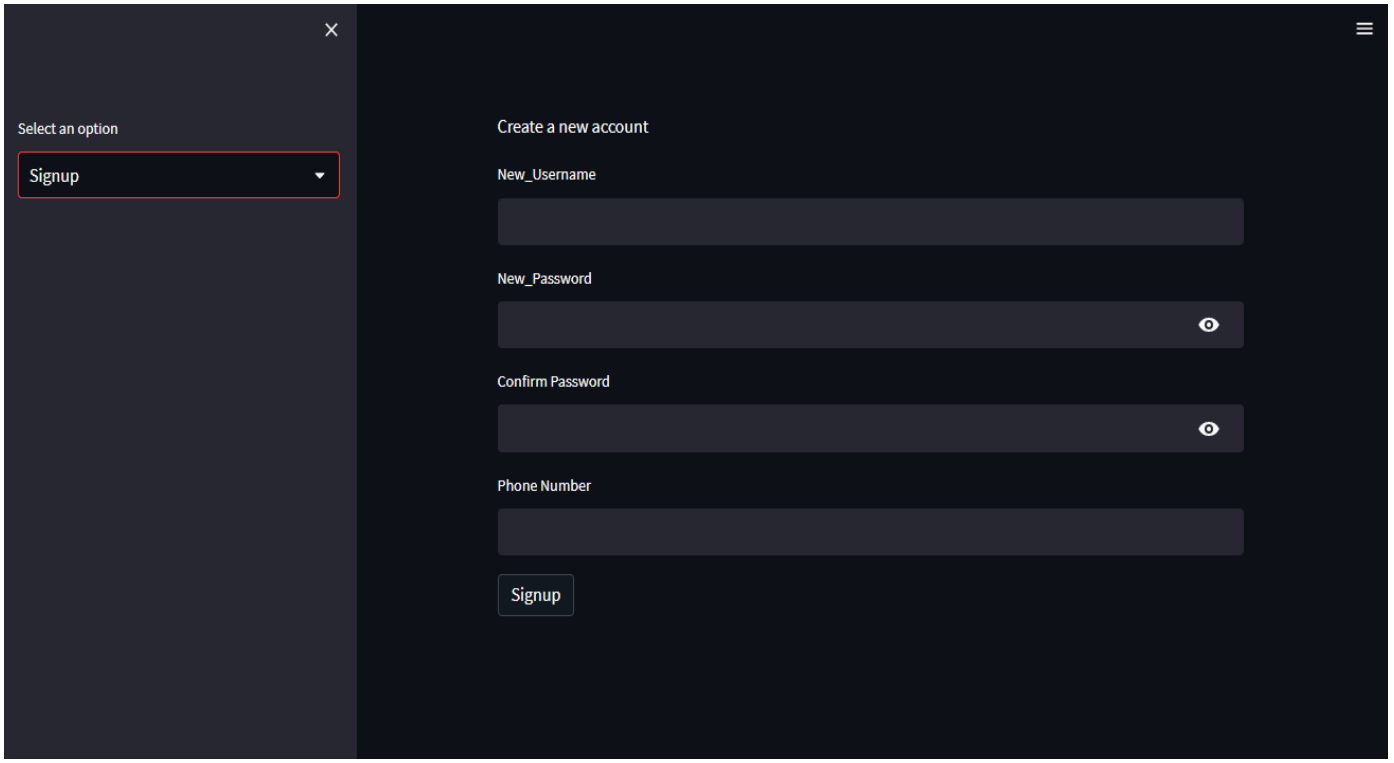
The Reward Collection with OCR project aimed to revolutionize the process of collecting rewards by automating the extraction and processing of relevant information from reward images. Through my research, implementation, and integration of AWS Textract, We successfully developed a comprehensive solution that streamlined the reward collection process.

Visual Representations:

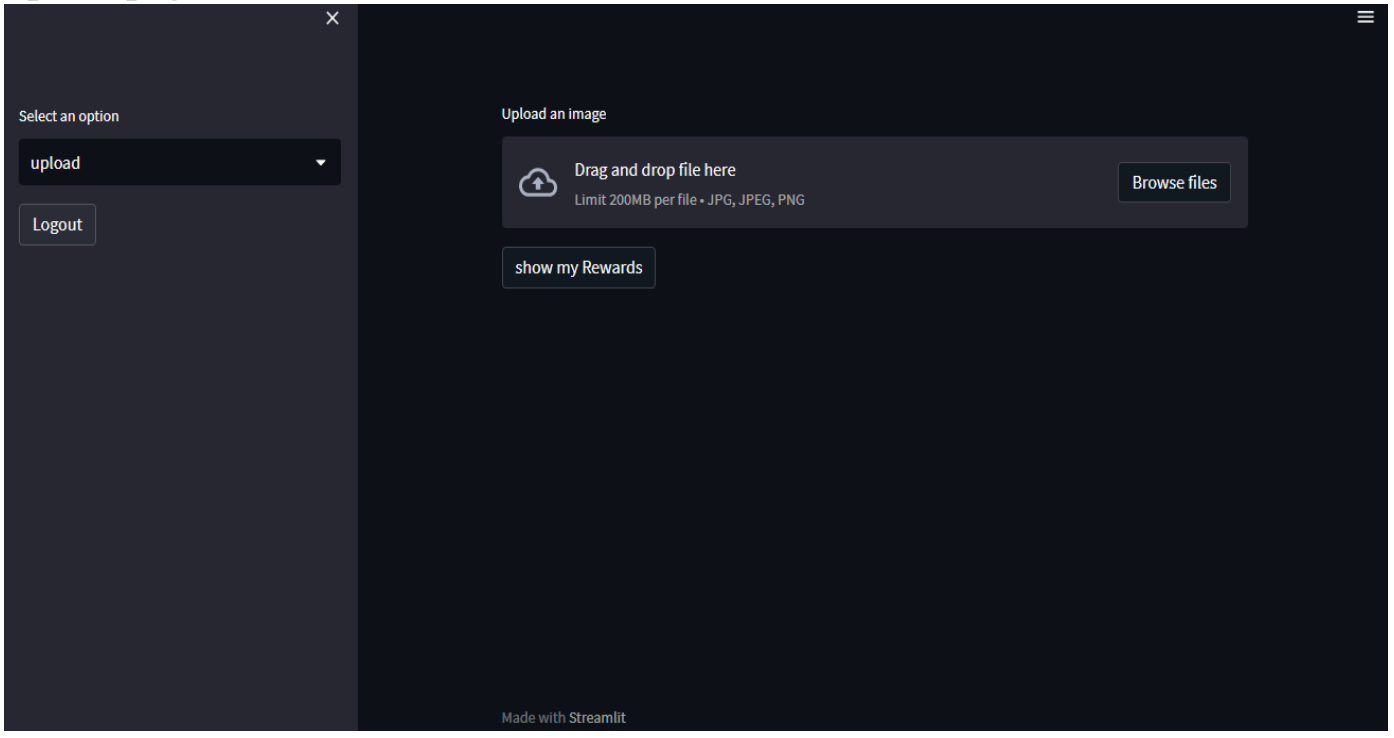
Login Page



Signup Page



Upload page:



Rewards Page

INTERNSHIP TIMELINE

Week 1: Orientation and Onboarding

Introduction to the company and the data science team
Familiarization with the company's policies and procedures
Setting up the development environment and necessary tools

Week 2-3: Sick Leave (Medical Absence)

Underwent surgery and took leave for recovery

Week 4-5: Python Basics and Advanced Concepts

Utilized the time for self-study and furthering my understanding of Python programming language
Explored advanced Python concepts and best practices

Week 6-7: SQL Training

Received training in SQL for data manipulation and management
Acquired knowledge of writing queries and interacting with databases

Week 7-10: Project Introduction And Research

Introduction to the Intellectual Property and OCR project
Conducted research on OCR techniques and AWS Textract services
Explored relevant libraries and technologies for implementation
Learned Version Control with Github

Week 11-14: Designing the System Architecture

Collaborated with the team to design the system architecture
Identified the required components and technologies
Planned the integration of AWS Textract and other tools into the project

Week 15-16: Data Integration and Storage

Set up a PostgreSQL database for secure storage of extracted reward information
Designed the database schema to store relevant data fields
Implemented data storage and retrieval functionalities

Week 17-18: Integration and Presentation with Senior Lead Team (Prototype)

Integrated the AWS Textract API for OCR capabilities
Tested and refined the OCR process for accuracy and efficiency
Final presentation with SLT (prototype)

Week 19-20: Introduction of Chatbot Project

During this period, a new project was introduced involving the production of a chatbot for a client requirement. The objective of this project was to develop a conversational AI solution that could provide automated customer support and assistance.

- Research on Chatbots
- NLP Research and Study.
- Research on libraries like RASA

Skills and Knowledge Acquired in Internship:

Programming Languages:

Through hands-on project work and continuous learning, I developed a strong command over Python and SQL. These languages served as the foundation for data manipulation, analysis, and visualization tasks.

Data Manipulation and Analysis :

I gained expertise in utilizing libraries such as Pandas and NumPy for efficient data manipulation, transformation, and cleaning

Cloud Computing:

Through the utilization of Amazon Web Services (AWS), specifically AWS RDS (Relational Database Service), I acquired knowledge of cloud-based infrastructure and gained experience in leveraging cloud resources for data storage and processing.

Collaborations and Teamwork:

Throughout my internship at Zapcom, I had numerous opportunities to collaborate and work as part of a team. These experiences not only enhanced my technical skills but also taught me valuable lessons about effective teamwork.

I had the privilege of working closely with a team of developers, data scientists, and project managers on the Intellectual Property (IP) project. Together, we brainstormed ideas, shared insights, and worked collaboratively to design and implement the OCR automation solution. I actively participated in team meetings, contributing my expertise in project development. Through this collaboration, I learned the importance of effective communication, adaptability, and embracing diverse perspectives to achieve project goals.

Personal Growth:

During my internship at Zapcom, I experienced significant personal growth and professional development. Through hands-on projects, I expanded my technical skills in Python, SQL, and data analysis techniques. Engaging with colleagues from diverse backgrounds enhanced my networking and communication abilities. I also sharpened my problem-solving and critical thinking skills by tackling real-world challenges. This internship has provided me with valuable experiences and insights that have shaped my future career goals in the field of data science. I am grateful for the support and mentorship received at Zapcom, and I am confident that the skills and knowledge gained will propel me towards a successful future in the industry.

My Internship Experience at Zapcom

At Zapcom, I had the privilege of working alongside experienced professionals and having a dedicated mentor who guided me throughout the internship. The supportive and collaborative work culture fostered an environment of continuous learning and growth. Regular demo presentations provided an opportunity to showcase my learnings and receive feedback from my mentor and the delivery head, enabling me to refine my skills and enhance my project outcomes.

Overall Experience and Appreciation: The internship experience at Zapcom exceeded my expectations, not only in terms of technical learning but also in the personal connections and interactions with colleagues. The company's inclusive culture and various events, such as happy hours, Fun Fridays, and celebrations, contributed to a vibrant and enjoyable work atmosphere. I am grateful for the knowledge gained, the relationships formed, and the positive experiences that will shape my future endeavors.

I extend my best wishes to everyone at Zapcom and express my gratitude for the knowledge, experience, and connections that I will carry forward in my career. It has been an honor to be part of such a remarkable organization.

References

- https://learn.ineuron.ai/lesson/Course-Overview/62eaa6ba766d65f5ae3169f7/course/Full-Stack-Data-Science-BootCamp-2.0/62eaa6ba766d6539c53164bd
- https://zapcom.udemy.com/course/
- https://zapcom.udemy.com/course/postgresqlmasterclass/l
- https://docs.python.org/3/
- https://docs.streamlit.io/
- https://docs.aws.amazon.com/textract/latest/dg/API_Reference.html
- https://docs.aws.amazon.com/
- https://www.postgresql.org/docs/
- https://www.linkedin.com/learning/?u=108216466
- https://rasa.com/
- https://www.youtube.com/
- https://www.google.com/