Dynamic Mess Billing System using Face Recognition

by

Ravi Chopra

2019BCS-046

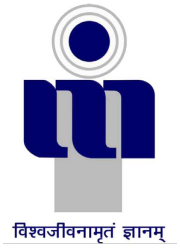
A thesis submitted in partial fulfilment of the requirements for the minor project of

Bachelor of Technology (Semester IV)

in

Computer Science Engineering

2019-23



ATAL BIHARI VAJPAYEE-

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY AND MANAGEMENT

GWALIOR - 474015, MADHYA PRADESH, INDIA

Thesis Certificate

I hereby certify that the work, which is being presented in the report/thesis, entitled Dynamic Mess Billing System using Face Recognition, in fulfillment of the requirement for minor project of **Bachelor of Technology (Semester IV)**
in **Computer Science Engineering** and submitted to the institution is an authentic record of my/our own work carried out during the period *May-2021* to *June-2021* under the supervision of Prof. Joydip Dhar Supervisor and Dr. Vinal Patel Co-Supervisor. I also cited the reference about the text(s)/figure(s)/table(s) from where they have been taken.

Prof. Joydip Dhar Supervisor and Dr. Vinal Patel Co-Supervisor

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Candidate's Declaration

I hereby certify that I have properly checked and verified all the items as prescribed in the check-list and ensure that my thesis is in the proper format as specified in the guideline for thesis preparation.
I declare that the work containing in this report is my own work. I understand that plagiarism is defined as any one or combination of the following:

(1) To steal and pass off (the ideas or words of another) as one's own

(2) To use (another's production) without crediting the source

(3) To commit literary theft

(4) To present as new and original idea or product derived from an existing source.

I understand that plagiarism involves an intentional act by the plagiarist of using someone else's work/ideas completely/partially and claiming authorship/originality of the work/ideas. Verbatim copy as well as close resemblance to some else's work constitute plagiarism.
I have given due credit to the original authors/sources for all the words, ideas, diagrams, graphics, computer programmes, experiments, results, websites, that are not my original contribution. I have used quotation marks to identify verbatim sentences and given credit to the original authors/sources.
I affirm that no portion of my work is plagiarized, and the experiments and results reported in the report/dissertation/thesis are not manipulated. In the event of a complaint of plagiarism and the manipulation of the experiments and results, I shall be fully responsible and answerable. My faculty supervisor(s) will not be responsible for the same.

Signature:

Name: Ravi Chopra

Roll. No: 2019BCS-046

Date: 25 June 2021

Abstract

This thesis represents fulfilled partial requirements for the Minor Project (Bachelor of Technology in Computer Science Engineering 2019-2023). **The idea proposed is a dynamic billing system for the student mess that internally utilises face recognition**. In this thesis, we discuss the work done until now, and the literature read for the project's development. The thesis also throws a shadow on potential research areas and the sub-modules of the whole project. Finally, the deadlines for completing the various tasks in the development process have been carefully drawn using a PERT chart.

Keywords: Face recognition, MERN stack, Insight Face, Computer Vision

Dedication

This project draws inspiration from existing systems available for billing systems in the student mess at different institutions in India. The other reason that dedicates us to develop this project is the current food wastage in the mess areas of our college. Mess workers have no idea of how much food will be utilised on the subsequent day, which leads to heavy wastage of food. This project will generate data that mess workers can use to get an idea of what quantity of food to prepare to minimise wastage. Many ideas have already been proposed for billing systems, but all lack at some points. Harnessing the power of deep learning, we can use face recognition for billing students on a per meal basis. However, this approach has some challenges that have been tried to be resolved in this work.

Acknowledgments

I want to extend my heartful thanks to **Prof. Joydip Dhar(Supervisor)** and **Dr Vinal Patel(Co-Supervisor)** for constantly guiding me through the project. They helped me to develop an excellent practice of reading the recent literature before pursuing the work. This practice increased my thirst for knowledge without any doubt. Their expertise in the field of Computer Vision helped me to think better and innovative.

I am indebeted to all the professors for allowing us to develop an industry-grade project in these times of peril. They gave their valuable time for evaluating and give much-needed insights about the project.

Thanks to my colleague **Pranav Kotgire** and my family for their constant support.

I would also like to thank **Dr Andrew NG** for starting a MOOC specialisation in deep learning. Specialisation was one of the primary sources of my learning.

(Ravi Chopra)

Contents

Chapter

Figures

Figure

Chapter 1

Introduction

This chapter presents an overview about the context as a part of project being developed in section 1.1. In section 1.2 the problems and motivations are presented. Next, in section 1.3, due to the presented motivations, a research work flow is introduced step by step. Finally, in section 1.4, the research objectives are presented.

1.1 Context

This project is a part of B.Tech. Computer Science Engineering curriculum for the third semester. The objective is to develop and experiment with solutions to automate and dynamise the billing system in the mess areas of the institute. This project will replace the current billing system (semester-based) with a dynamic system that bills students on a per meal basis. The aim is to use a face recognition module to achieve the above-stated objectives.

1.2 Problem/Motivation

Project Motivation : The problem with the existing billing system is that it bills all the students the same even if they skip some meals due to some reasons. Moreover, there is no way for our mess workers to predict how much food to prepare, which leads to a lot of food wastage. These reasons motivate me to develop a system that bills students dynamically and generates data for mess workers. This data will help them to estimate the amount of food to prepare to minimise food wastage.

Face recognition problems: Most existing face recognition systems suffer from problems like face orientation in the image, varying brightness, noise, etc. Some models also suffer from the problem of overfitting due to the non-linearity of the cosine function. (More details are discussed in further sections).

Web interface features : This project includes a web interface for students to view and pay their bills on monthly bills. Every student will be notified as a text message when he/she enters the mess area. A proper database for each student will be maintained where students' personal information will be stored in encrypted form.

Design and UX : Various aspects of design and User Experience have been kept in mind during the development of the project. Till now, a design has been made using Canva, a design software.

1.3 Objectives

Face recognition module : The project's objective is to develop a robust face recognition system that can work in a mess area condition. The model must be robust to changes in brightness, orientation and other factors.

Web application : The other aim is to provide a web interface for students to interact with the system. It will include a payment gateway developed using one of the APIs available in the market.

Deployment : Optionally, we are thinking to deploy the face recognition model on Raspberry Pi. This objective is just a speculation that we try to achieve.

1.4 Work flow

According to the objectives, the report will describe the work flow as below:

Step 1 Developing/Improving an existing model for face recognition that is not sensitive to changes to brightness and orientation.

Step 2 Developing a web interface for students to interact with system.

Step 3 Using requests module in python to help the model interact with web application's server side.

Step 4 Using Twilio API to send notification messages to students.

Step 5 Using Payments API to allow students clear their dues.

Step 6(Optional) Deployment on Raspberry Pi.

Chapter 2

Literature review

2.1 Background

This section of the thesis focuses on the literature review conducted before starting the development process. It includes a short write-up and a critical analysis of the previous works on the objective proposed in the thesis. There is a comparison drawn between different works based on standard benchmarks used in the industry.

2.2 Key related research

Various works were reviewed and analysed related to the objectives stated in the theses. These are listed below:

- ArcFace: Additive Angular Margin Loss for Deep Face Recognition (9 Feb 2019)
- Mis-classified Vector Guided Softmax Loss for Face Recognition (26 Nov 2019)
- LinCos-Softmax: Learning Angle-Discriminative Face Representations With Linearity-Enhanced Cosine Logits (15 June 2020)

2.3 Analysis

The above works were carefully analysed before moving forward. All the papers to be discussed propose different loss functions for the classification problem of faces into classes. They

tend to minimize intra-class variations and maximize inter-class distances. All the loss functions are just small variations of softmax loss. Accordingly, a short analysis for each paper is stated below:

- **ArcFace: Additive Angular Margin Loss for Deep Face Recognition** : The idea is to add an additive angular margin penalty m between X_i and W_{yi} to simultaneously enhance the intra-class compactness and inter-class discrepancy. This additive angular margin penalty is equal to the geodesic distance margin penalty in the normalised hypersphere. Loss:

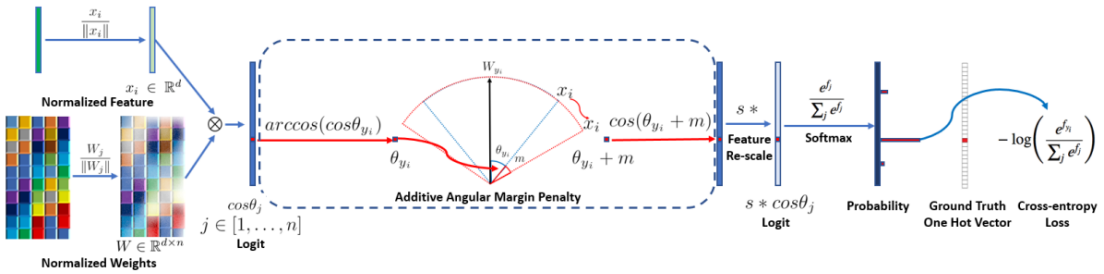


Figure 2. Training a DCNN for face recognition supervised by the ArcFace loss. Based on the feature x_i and weight W normalisation, we get the $\cos\theta_j$ (logit) for each class as $W_j^T x_i$. We calculate the $arccos\theta_{y_i}$ and get the angle between the feature x_i and the ground truth weight W_{y_i}. In fact, W_j provides a kind of centre for each class. Then, we add an angular margin penalty m on the target (ground truth) angle θ_{y_i}. After that, we calculate $\cos(\theta_{y_i} + m)$ and multiply all logits by the feature scale s. The logits then go through the softmax function and contribute to the cross entropy loss.

Figure 2.1: Source: ArcFace: Additive Angular Margin Loss for Deep Face Recognition

$$L_3 = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{n}e^{s\cos\theta_j}} \tag{2.1}$$

- **Mis-classified Vector Guided Softmax Loss for Face Recognition** : This paper tries to design a new loss function, which explicitly indicates the hard examples as mis-classified vectors and adaptively emphasizes on them to guide the discriminative feature learning. As a consequence, our new loss also absorbs the discrimiantibility from other non-ground truth classes as well as is with adaptive margins for different classes.This paper defines a binary indicator I_k to adaptively indicate whether a sample (feature) is mis-classified by a

specific classifier w_k (where k != y) in the current stage:

$$I_k = \begin{cases} 0, & f\left(m, \theta_{w_y,x}\right) - \cos\left(\theta_{w_k,x}\right) \geq 0 \\ 1, & f\left(m, \theta_{w_y,x}\right) - \cos\left(\theta_{w_k,x}\right) < 0 \end{cases} \tag{2.2}$$

Loss:

$$\mathcal{L}_5 = -\log \frac{e^{sf\left(m,\theta_{w_y,x}\right)}}{e^{sf\left(m,\theta_{w_y,x}\right)} + \sum_{k\neq y}^{K} h\left(t, \theta_{w_k,x}, I_k\right) e^{s\cos\left(\theta_{w_k,x}\right)}} \tag{2.3}$$

- **LinCos-Softmax: Learning Angle-Discriminative Face Representations With Linearity-Enhanced Cosine Logits** : Despite the excellent performance achieved by the angle-based softmax loss variants, one weakness is that the angle is nonlinearly mapped by a cosine function. The nonlinearity of the cosine function may lead to insufficient angular optimization between features and corresponding class weights. As a result, the angular discriminability of the features may be compromised, resulting in a reduced generalization ability. To tackle this issue, we propose a Linear-Cosine Softmax Loss to learn angle-discriminative face features more effectively. The main idea is the use of a linear-cosine logit, which is designed by performing Taylor expansion on a linear logit. This paper represents θ_j as the arccosine of $\cos\theta_j$, then perform a Taylor expansion over the arccosine function, and approximate the angle using the first K terms:

  Loss:

$$\theta_j = \arccos\left(\cos\theta_j\right) \approx \hat{\theta}_j$$

$$\text{where } \hat{\theta}_j = \frac{\pi}{2} - \sum_{n=0}^{K-1} c_n \left(\cos\theta_j\right)^{2n+1}$$

$$\text{and } c_n = \frac{(2n)!}{2^{2n}(n!)^2(2n+1)}$$

$$f_j^{linear} = -\theta_j + \frac{\pi}{2} = -\arccos\left(\cos\theta_j\right) + \frac{\pi}{2}$$

$$f_j^{LinCos} = -\hat{\theta}_j + \frac{\pi}{2} = \sum_{n=0}^{K-1} c_n \left(\cos\theta_j\right)^{2n+1}$$

$$L_i^{LinCos} = -\log\left(P_{y_i}^{LinCos}\right)$$

$$P_{yi}^{LinCos} = \frac{e^{sf_{yi}\,Lin\cos}}{\sum_{j=1}^{C} e^{sf_j^{LinCos}}}$$

2.4 Research gaps

In this section we compare different loss functions which were reviewed in the previous section. The comparison is done on basis of recognition loss(in) on the basis of standard benchmarks LFW and AgeDB. It is observed that all the algorithms are on their saturation, i.e. it is quite not possible to decide the best of them.

Results: See figure 2.6

Loss	LFW	AgeDB
ArcFace	99.75%	98.2%
MV-Softmax	99.76%	98.01%
LinCos	99.2%	98.06%

Figure 2.2: Results

2.5 Problem formulation

As we can see that all of the above papers already give a pretty good accuracy, it is pretty impossible to choose the best one. However, due to ease of implementation and familiarity with Pytorch, I chose to proceed with ArcFace. Various experiments can be done with the above methods. One of them is to combine the ideas of LinCos and ArcFace. We can approximate $\cos(\theta + m)$ (where m = Marginal Penalty) using the method proposed in LinCos paper and reduce the problem of over-fitting. Another possible work-around is to replace cosine with the tangent.

2.6 Conclusion

To conclude, there has been extensive research in the field of face recognition. Due to this, many proposed models have already achieved an accuracy that is possibly less only by the Bayes error. However, there is still room for experimentation, and my focus will remain on trying the above-proposed experiments.

Chapter 3

Software Design and Requirements

This section gives description about the requirements and the design of the project. It utilises DFDs(Data Flow Diagrams) to represent the working and inter-dependency of different modules of the system.

3.1 System Requirements

The major requirements for the system are listed below:

- Students are recognised by their faces and mapped to their IDs.
- Students can view their attended meals.
- Students get notified on entering the mess.
- Students can pay and view bills on monthly basis.

3.2 System Design

System design shown in figure 3.1 is a representation of how different modules and users will interact with each other.

3.3 Data Flow Diagrams

- Level 0 DFD: See figure 3.2

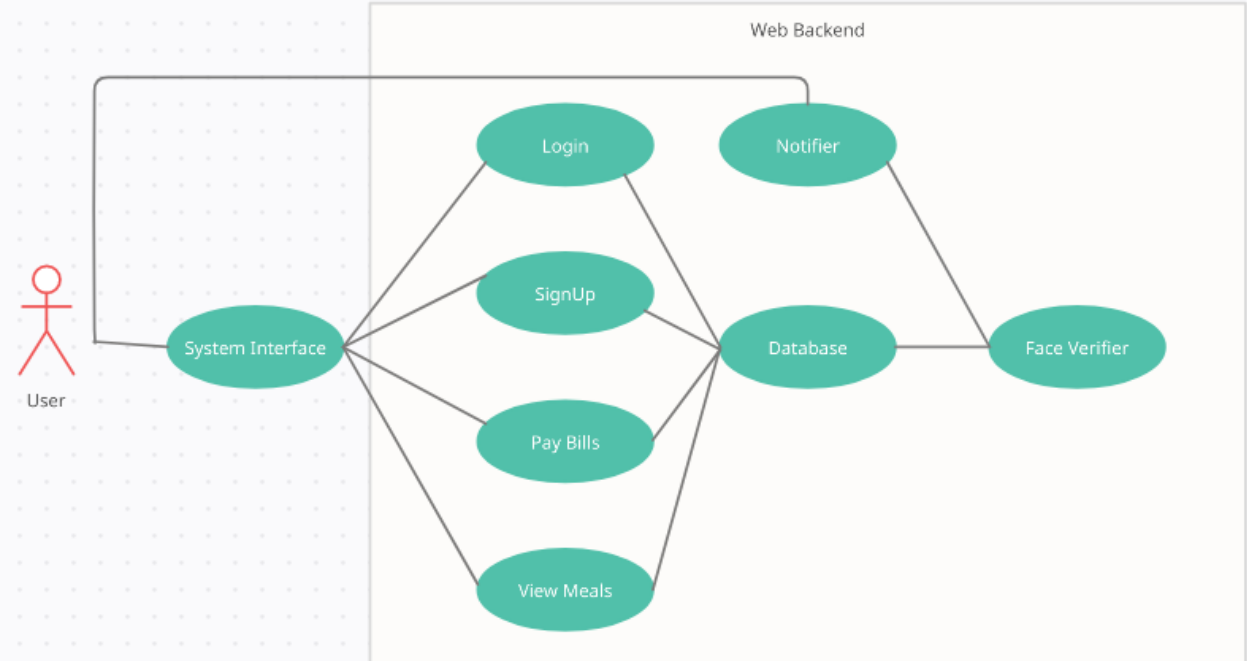


Figure 3.1: System Design

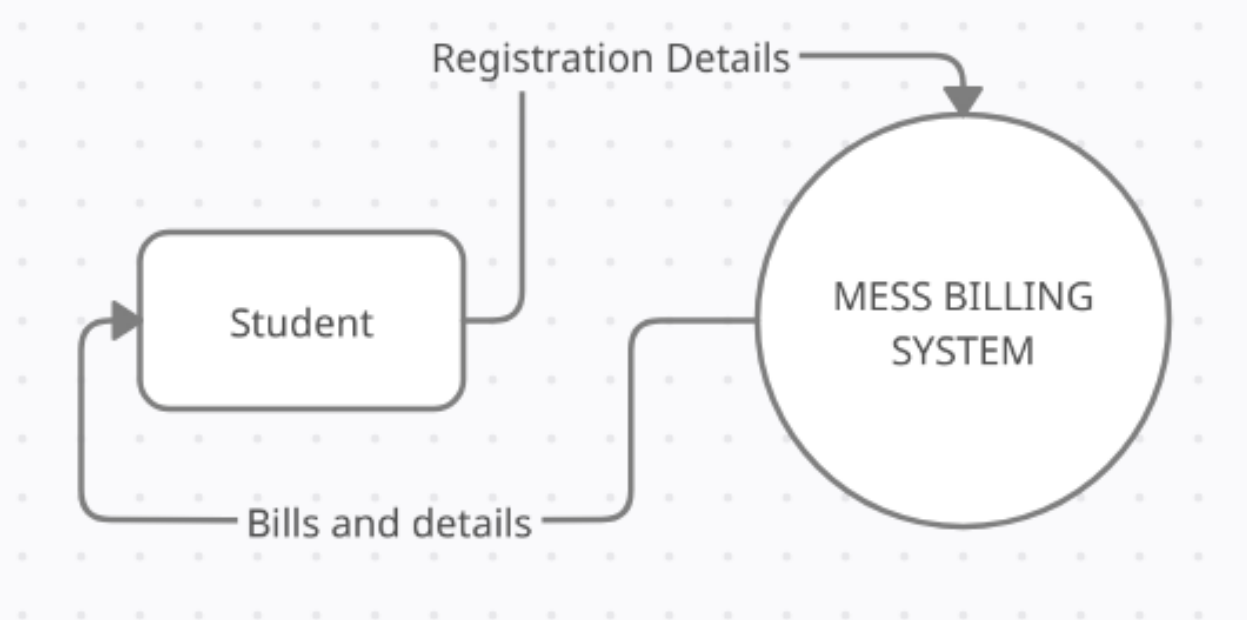


Figure 3.2: Level 0 DFD

- Level 1 DFD: See figure 3.3

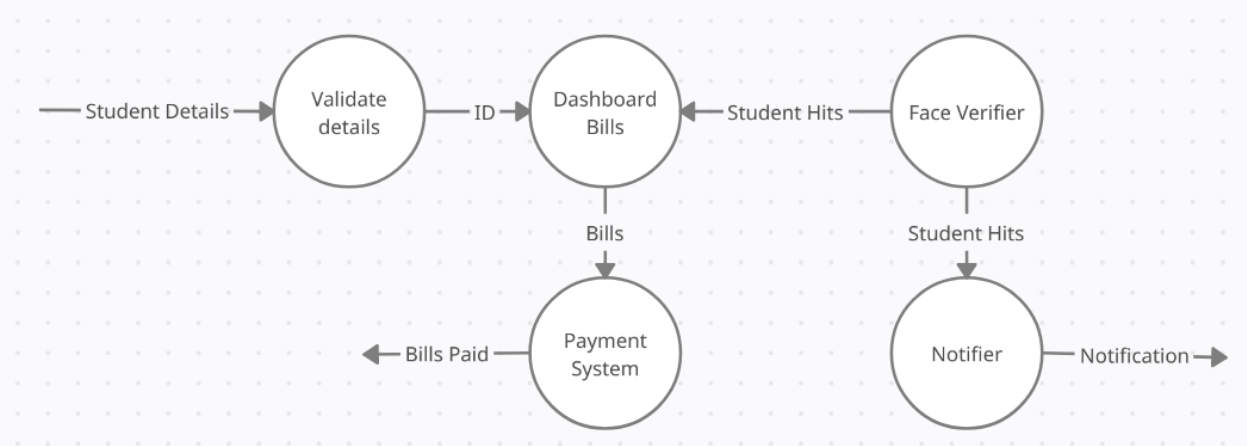


Figure 3.3: Level 1 DFD

- Level 2 DFD: See figure 3.4

3.4 Implementation specifics

This sub-section proposes the tech-stack used in the project.

- **Pytorch** : It is used as the framework for development of face recognition model.
- **MERN stack** : MERN stands for MongoDB database, Express.JS, React.JS, Node.JS. This is the used stack for web interface.
- **requests** : It is a module in python used for making calls to web server from face recognition system.
- **Twilio API** : This is the Application Programming Interface(API) used to send notification messages to students.
- **Payment gateway** : It will be an API used to allow students pay their bills online. We are yet to choose which API to use. The deciding factors for the chosen API will be latency, cost effectiveness, and, ease of use.

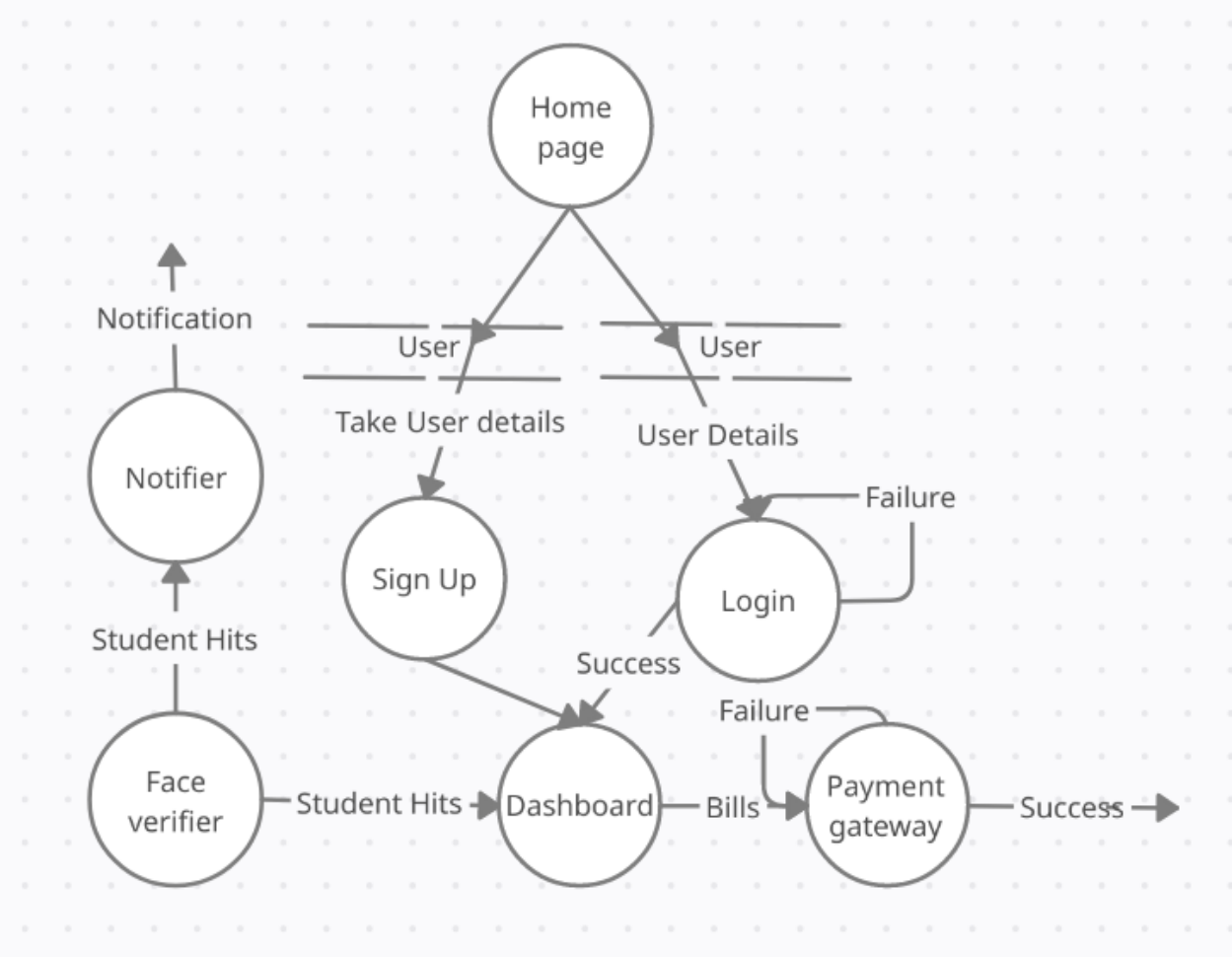


Figure 3.4: Level 2 DFD

3.5 Conclusion

This section was a representation of system design and interface using Data Flow Diagrams. This will help me to develop things more precisely. Also, the addition of new features would be more straightforward.

Chapter 4

Discussions and conclusion

In this chapter, the thesis is concluded with a plan of action. As discussed in the thesis, Data Flow Diagrams(DFDs) would be duly followed during development. Also, the focus will be primarily on the potential research areas pointed out in the thesis. This thesis will stand out as a base for the project's requirements and is a proof of study done about the project before starting the development process.

Bibliography

[1] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[2] Wei-Feng Ou, Lai-Man Po, Chang Zhou, Yu-Jia Zhang, Li-Tong Feng, Yasar Abbas Ur Rehman, and Yu-Zhi Zhao. Lincos-softmax: Learning angle-discriminative face representations with linearity-enhanced cosine logits. IEEE Access, 8:109758–109769, 2020.

[3] Xiaobo Wang, Shifeng Zhang, Shuo Wang, Tianyu Fu, Hailin Shi, and Tao Mei. Mis-classified vector guided softmax loss for face recognition, 2019.