DUYTAN UNIVERSITY



INTERNATIONAL SCHOOL

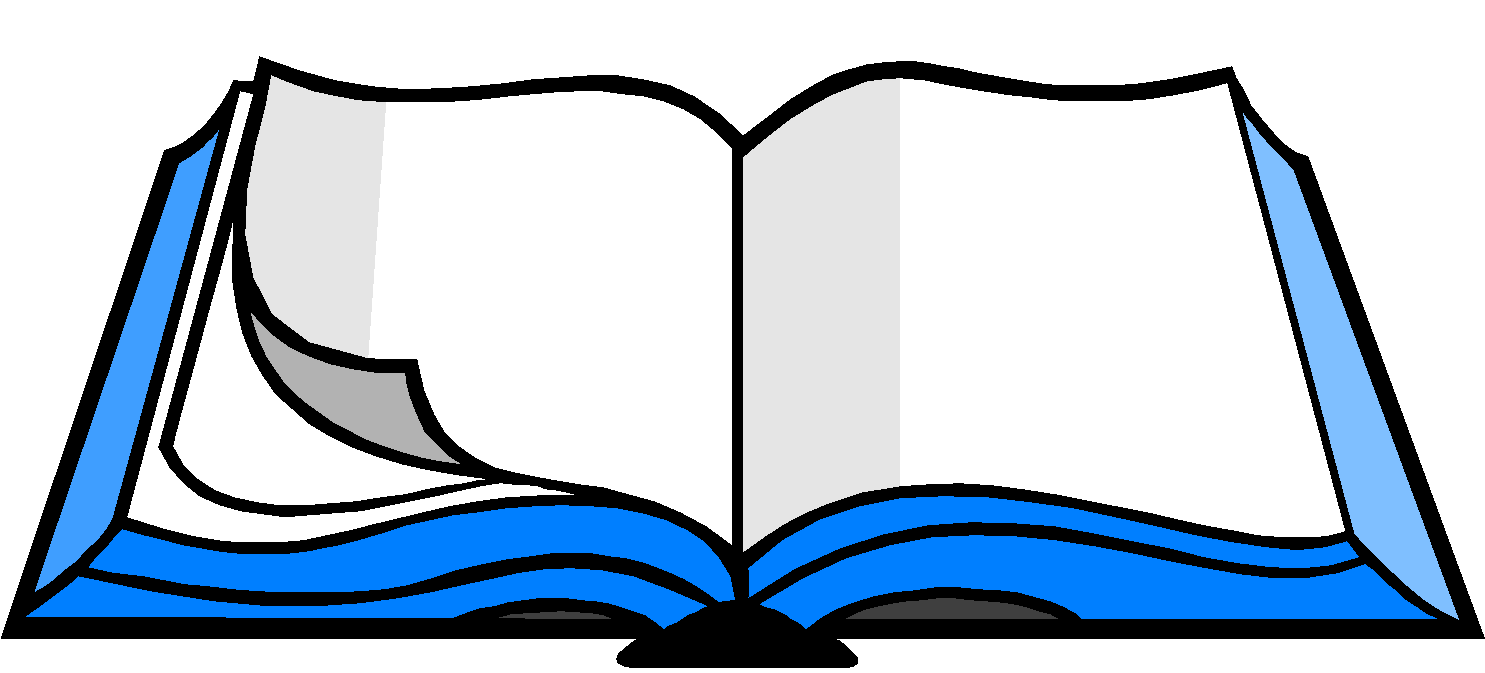
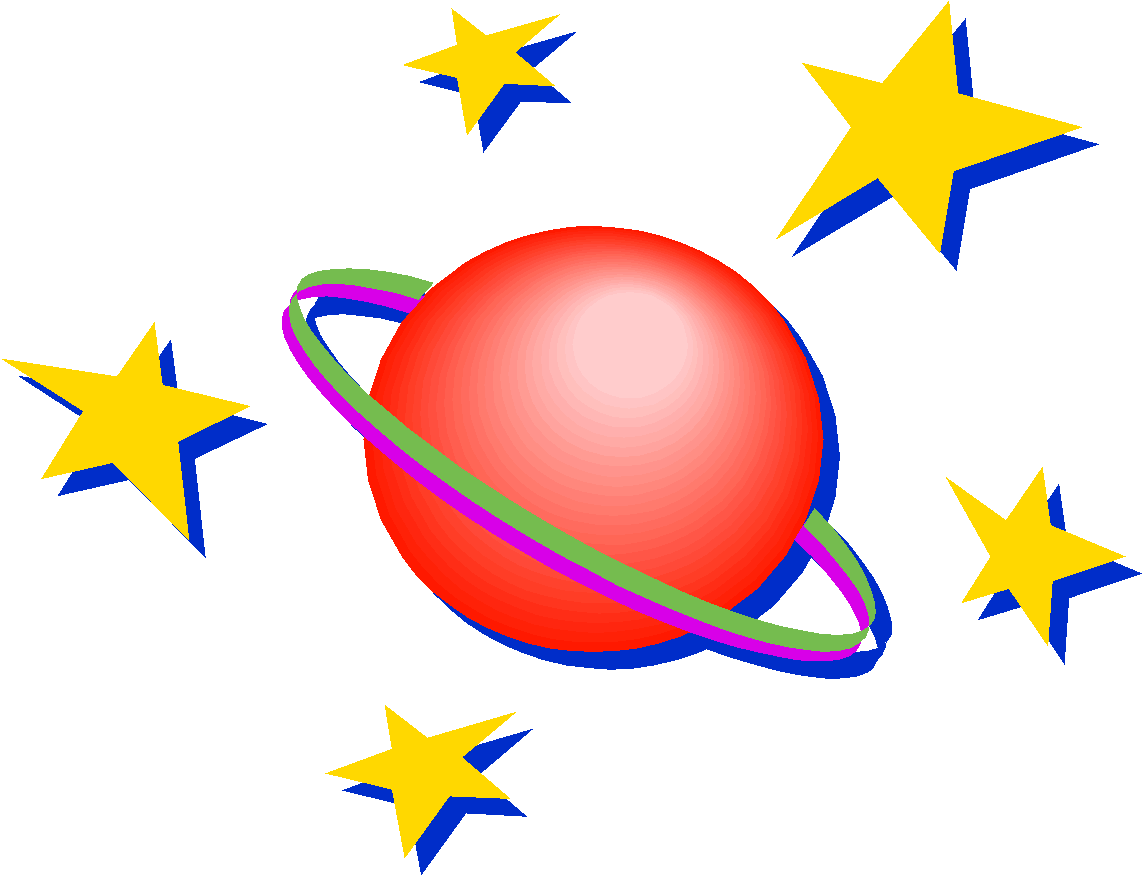
--------------🙖🙐✰🙖🙐-------------



CS 466

PERL PYTHON

Hệ Thống Chấm Điểm Bài Tập Tự Động



Giảng Viên: Trần Hữu Minh Đăng

*Thành Viên*:

Đặng Nguyễn Nhật Quang

Lê Thị Thanh Nguyệt

Trần Phước Nhân

Đỗ Quốc Tú

Nguyễn Sơn Quân

Group Project - Mentor:

Name Signature Date

Da nang, 09/2025

TÓM TẮT / ABSTRACT (TIẾNG VIỆT & TIẾNG ANH)

# MỤC LỤC

1. [CHƯƠNG 1: GIỚI THIỆU](#_heading=h.x244otpyhcne) 4
   1. [Lý do chọn đề tài](#_heading=h.ukvborqusmmz) 4
   2. [Vấn đề đặt ra](#_heading=h.wii0c4gjasfs) 4
   3. [Mục tiêu và phạm vi](#_heading=h.pd0vdclvot3o) 4
   4. [Phương pháp nghiên cứu](#_heading=h.gu7uqpeoe3pu) 4
   5. [Bố cục báo cáo](#_heading=h.68uwdjpr6z0m) 4
2. [CHƯƠNG 2: CƠ SỞ LÝ THUYẾT & NGHIÊN CỨU LIÊN QUAN](#_heading=h.olvrz9qnfh1d) 4
   1. [Các khái niệm, công nghệ nền tảng](#_heading=h.5g6n085x62ry) 4
   2. [Các nghiên cứu/ứng dụng trước đây](#_heading=h.oso6tqtbggs1) 4
3. [CHƯƠNG 3: PHÂN TÍCH & THIẾT KẾ HỆ THỐNG/GIẢI PHÁP](#_heading=h.cx2pn1xxuli7) 4
   1. [Yêu cầu hệ thống](#_heading=h.4mr7bau6yfw3) 4
   2. [Kiến trúc / mô hình ERD, UML, sơ đồ mạng …](#_heading=h.zdi2kwwpll9h) 4
4. [CHƯƠNG 4: TRIỂN KHAI & KẾT QUẢ](#_heading=h.heg5rullktbe) 4
   1. [Công cụ, môi trường](#_heading=h.cz3bijgeqz63) 4
   2. [Cấu hình/thực nghiệm](#_heading=h.df4sv1sa1df) 4
   3. [Kết quả đạt được](#_heading=h.l4ansag6uf4q) 4
5. [CHƯƠNG 5: KẾT LUẬN & HƯỚNG PHÁT TRIỂN](#_heading=h.sgvepkpkwrr) 4
   1. [Tài liệu tham khảo (chuẩn APA/IEEE, không lấy từ Wikipedia)](#_heading=h.dy5z01lm1oiq) 4
   2. [Phụ lục (code, script, dataset, sơ đồ chi tiết …)](#_heading=h.nv65pn84gtft) 4

## 

## 

## 

## 

## CHƯƠNG 1: GIỚI THIỆU

### Lý do chọn đề tài

### 

### Vấn đề đặt ra

### 

### Mục tiêu và phạm vi

### 

### Phương pháp nghiên cứu

### 

### Bố cục báo cáo

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT & NGHIÊN CỨU LIÊN QUAN

### Các khái niệm, công nghệ nền tảng

### 

### Các nghiên cứu/ứng dụng trước đây

## 

## CHƯƠNG 3: PHÂN TÍCH & THIẾT KẾ HỆ THỐNG/GIẢI PHÁP

### Yêu cầu hệ thống

### 

### Kiến trúc / mô hình ERD, UML, sơ đồ mạng …

## 

## CHƯƠNG 4: TRIỂN KHAI & KẾT QUẢ

### Công cụ, môi trường

### 

### Cấu hình/thực nghiệm

### 

### Kết quả đạt được

## 

## CHƯƠNG 5: KẾT LUẬN & HƯỚNG PHÁT TRIỂN

### Tài liệu tham khảo (chuẩn APA/IEEE, không lấy từ Wikipedia)

### 

### Phụ lục (code, script, dataset, sơ đồ chi tiết …)

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT & NGHIÊN CỨU LIÊN QUAN

**2.1 Các khái niệm và công nghệ nền tảng.**

Trong chương này, nhóm sẽ trình bày các kiến thức nền tảng và công nghệ chính được sử dụng để xây dựng hệ thống chấm bài tự động. Đây là cơ sở khoa học và thực tiễn để hiểu rõ hơn về cách hệ thống hoạt động, cũng như những nghiên cứu và ứng dụng liên quan đã có trước đây.

#### **Ngôn ngữ Python:** Python là một ngôn ngữ lập trình bậc cao, thông dịch, được ưa chuộng trong nhiều lĩnh vực như khoa học dữ liệu, trí tuệ nhân tạo, phát triển web và tự động hóa. Python có cú pháp đơn giản, dễ đọc, cộng đồng lớn và hệ sinh thái thư viện phong phú. Trong dự án này, Python được dùng làm ngôn ngữ chính để phát triển ứng dụng web chấm điểm tự động.

* **Flask Framework:** Flask là một micro-framework của Python hỗ trợ xây dựng ứng dụng web. Với ưu điểm nhẹ, dễ mở rộng và tích hợp, Flask cho phép lập trình viên phát triển nhanh các ứng dụng web từ nhỏ đến trung bình. Các thành phần chính gồm:
* Routing: định nghĩa URL và hàm xử lý.
* Template Engine (Jinja2): hỗ trợ kết xuất giao diện HTML.
* Request & Response Handling: xử lý dữ liệu từ client gửi lên và trả về kết quả.

Trong hệ thống, Flask được dùng để: Quản lý **upload file phiếu trả lời**, gọi hàm từ grader.py để chấm điểm, trả kết quả về cho người dùng qua **HTML template**, cung cấp giao diện nhập đáp án cho giảng viên.

#### **JSON (JavaScript Object Notation):** JSON là định dạng dữ liệu phổ biến, nhẹ, dễ đọc và dễ xử lý trong nhiều ngôn ngữ. File answer\_key.json được dùng để lưu trữ đáp án chuẩn của các câu hỏi, phục vụ quá trình so sánh và chấm điểm.

#### **OpenCV:** OpenCV là thư viện mã nguồn mở hỗ trợ xử lý ảnh và thị giác máy tính. Nó có khả năng đọc, ghi, tiền xử lý ảnh (lọc nhiễu, chuyển đổi màu, phát hiện cạnh) để nâng cao chất lượng dữ liệu đầu vào cho các bước nhận dạng. Trong hệ thống, dùng để phát hiện contour, phân tích vùng tô của phiếu trả lời grader

* **NumPy**: NumPy là thư viện Python mạnh mẽ cho tính toán khoa học, cung cấp cấu trúc mảng đa chiều và nhiều hàm toán học tối ưu. Trong dự án, NumPy được sử dụng để hỗ trợ xử lý dữ liệu số, đồng thời kết hợp với OpenCV và Pillow trong các phép tính ảnh.
* **Pillow (PIL Fork):** Pillow là thư viện xử lý ảnh trong Python, hỗ trợ mở, chỉnh sửa, và lưu trữ ảnh ở nhiều định dạng khác nhau. Nó thường được dùng để chuẩn hóa ảnh trước khi áp dụng OCR.
* **Tesseract OCR**: Tesseract là một công cụ nhận dạng ký tự quang học (OCR) mã nguồn mở do Google phát triển. Nó cho phép trích xuất văn bản từ ảnh scan hoặc ảnh chụp, giúp hệ thống có thể chấm điểm trực tiếp từ file ảnh chứa đáp án. Trong đề tài, Tesseract OCR được sử dụng để đọc thông tin cá nhân từ header của phiếu (Họ tên, MSSV), xác định **Mã đề** để biết bộ đáp án chuẩn cần sử dụng.
* **OpenPyXL**: OpenPyXL là thư viện Python dùng để đọc và ghi file Excel (.xlsx). Trong ứng dụng, OpenPyXL cho phép xuất kết quả chấm điểm ra bảng điểm Excel, giúp việc lưu trữ và tổng hợp dữ liệu thuận tiện hơn.

### **2.2 Các nghiên cứu và ứng dụng liên quan**

#### Google Forms / Microsoft Forms: là những công cụ trực tuyến miễn phí, dễ sử dụng và hỗ trợ chấm điểm tự động. Ưu điểm nổi bật của các hệ thống này là giao diện thân thiện và khả năng triển khai nhanh chóng. Tuy nhiên, hạn chế lớn nhất là chúng chỉ phù hợp cho các kỳ thi trực tuyến và hoàn toàn không thể xử lý các phiếu trả lời trên giấy.

#### Hệ thống LMS (Moodle, Canvas): là các hệ thống quản lý học tập (LMS) toàn diện, cung cấp đầy đủ chức năng quản lý lớp học, bài kiểm tra, chấm điểm và xuất báo cáo. Ưu điểm của các hệ thống này là tính năng đa dạng và khả năng quản lý trên quy mô lớn. Tuy vậy, nhược điểm là việc cài đặt và vận hành khá phức tạp, đòi hỏi server mạnh, nên khó áp dụng cho các lớp học nhỏ hoặc nhu cầu thử nghiệm.

#### Phần mềm thương mại (ZipGrade, Remark OMR): có ưu điểm là độ chính xác cao, tốc độ xử lý nhanh và hỗ trợ tốt cho các kỳ thi với quy mô lớn. Nhưng hạn chế của chúng nằm ở chi phí bản quyền cao và yêu cầu thiết bị phần cứng chuyên dụng, điều này gây khó khăn cho việc triển khai trong môi trường học đường phổ thông hoặc các lớp học quy mô nhỏ.

#### Các dự án mã nguồn mở: Trên các nền tảng như GitHub, có nhiều dự án mã nguồn mở sử dụng OpenCV để nhận diện và chấm phiếu trắc nghiệm. Ưu điểm của những dự án này là tính linh hoạt, dễ tiếp cận và phù hợp cho mục đích học tập hoặc nghiên cứu. Tuy nhiên, hầu hết các dự án chỉ dừng lại ở mức độ minh họa (demo), chưa phát triển thành hệ thống hoàn chỉnh, thiếu giao diện web thân thiện và không có chức năng quản lý đáp án.

So sánh với các giải pháp trên, đề tài của nhóm có đặc điểm:

* **Chi phí thấp**: chỉ cần Python và thư viện mã nguồn mở.
* **Gọn nhẹ**: dễ triển khai trong phạm vi lớp học.
* **Có giao diện web trực quan**: giảng viên chỉ cần upload ảnh và nhận kết quả ngay.
* **Có khả năng mở rộng:** có thể thêm nhiều mã đề, xuất báo cáo Excel.