

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC CẦN THƠ
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



LUẬN VĂN TỐT NGHIỆP
MẠNG MÁY TÍNH VÀ TRUYỀN THÔNG DỮ LIỆU

Đề tài
XÂY DỰNG HỆ THỐNG THỐNG KÊ
KẾT QUẢ XỔ SỐ KIẾN THIẾT
TRÊN NỀN TẢNG WEB

(Developing a Statistical System for Lottery Results on a Web Platform)

Sinh viên thực hiện:

HỒ ANH KIỆT

MSSV: B2004731

KHOÁ: 46

Cần Thơ, 12/2024

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC CẦN THƠ
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



LUẬN VĂN TỐT NGHIỆP
MẠNG MÁY TÍNH VÀ TRUYỀN THÔNG DỮ LIỆU

Đề tài
XÂY DỰNG HỆ THỐNG THỐNG KÊ
KẾT QUẢ XỔ SỐ KIẾN THIẾT
TRÊN NỀN TẢNG WEB

(Developing a Statistical System for Lottery Results on a Web Platform)

Giảng viên hướng dẫn
ThS. TRANG THANH TRÍ

Sinh viên thực hiện
HỒ ANH KIỆT
MSSV: B2004731
KHOÁ: 46

Cần Thơ, 12/2024

TRƯỜNG ĐẠI HỌC CẦN THƠ
TRƯỜNG CÔNG NGHỆ THÔNG TIN
VÀ TRUYỀN THÔNG

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

XÁC NHẬN CHỈNH SỬA LUẬN VĂN THEO YÊU CẦU CỦA HỘI ĐỒNG

Tên luận văn (tiếng Việt và tiếng Anh): **Xây dựng hệ thống thống kê kết quả xổ số
kiến thiết trên nền tảng web - Developing a Statistical System for Lottery Results
on a Web Platform**

Họ tên sinh viên: Hồ Anh Kiệt

MASV: B2004731

Mã lớp: DI20T9A1

Đã báo cáo tại hội đồng ngành: Mạng máy tính và truyền thông dữ liệu

Ngày báo cáo: 11/12/2024

Luận văn đã được chỉnh sửa theo góp ý của Hội đồng.

Cần Thơ, ngày 16. tháng 12. năm 2024

Ký thay Giáo viên hướng dẫn

(Ký và ghi họ tên)


TS. Trần Thị Tố Quyên

LỜI CẢM ƠN

Đầu tiên, em muốn bày tỏ lòng biết ơn sâu sắc đến Trường Công nghệ Thông tin & Truyền Thông vì đã tạo điều kiện học tập tốt nhất cho em. Đặc biệt, em xin gửi lời cảm ơn chân thành đến Thầy, Cô đã chia sẻ những kiến thức quý báu và truyền đạt những lời khuyên chân thành. Trong quãng thời gian học tập, em đã tích lũy được nhiều kiến thức bổ ích, là nguồn động viên quan trọng giúp em vững bước trên con đường học nghiên cứu sắp tới.

Em không quên bày tỏ lòng biết ơn sâu sắc đến Thầy Trang Thanh Trí, người đã tận tâm hỗ trợ và hướng dẫn em trong suốt quá trình thực hiện bài luận văn này. Mặc dù em đã nỗ lực hết mình, nhưng với vốn kiến thức còn hạn chế, cũng như năng lực phù hợp với các yêu cầu bài luận văn nên không tránh khỏi những khuyết điểm. Em kính mong Thầy xem xét và góp ý, để em có cơ hội rút kinh nghiệm và hoàn thiện bản thân hơn trong những công việc nghiên cứu sắp tới.

Cuối cùng, em xin chân thành cảm ơn tất cả sự hỗ trợ và đóng góp, mang đến cho em những trải nghiệm quý giá. Lời cảm ơn này không chỉ là sự biểu hiện của lòng biết ơn cá nhân mà còn là lời cam kết về việc áp dụng những kiến thức và kỹ năng đã học để phát triển bản thân và đóng góp cho xã hội.

Em xin chân thành cảm ơn!

MỤC LỤC

CHƯƠNG 1. GIỚI THIỆU	1
1.1. Đặt vấn đề.....	1
1.2. Mục tiêu đề tài	3
1.3. Đối tượng nghiên cứu.....	3
1.4. Mục tiêu nghiên cứu.....	4
1.4.1. Mục tiêu tổng quát.....	4
1.4.2. Mục tiêu cụ thể	4
1.5. Phạm vi nghiên cứu	4
1.5.1. Phạm vi dữ liệu	4
1.5.2. Phạm vi chức năng của hệ thống:	4
1.5.3. Phạm vi công nghệ.....	4
1.5.4. Phạm vi người dùng.....	5
1.5.5. Phạm vi thời gian.....	5
1.6. Phương pháp nghiên cứu.....	5
1.7. Ý nghĩa của đề tài nghiên cứu	6
1.8. Bố cục của quyển luận văn.....	6
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT	7
2.1. Frontend	7
2.1.1. HTML (Hypertext Markup Language)	7
2.1.2 CSS (Cascading Style Sheets)	8
2.1.3. JavaScript.....	8
2.2. Backend	9
2.3. Các công cụ và các thư viện	9
2.3.1. Playwright.....	9
2.3.2. Chart.js.....	10
2.3.3. html2canvas	11
2.3.4. jsPDF	11

2.3.5. Công cụ hỗ trợ lập trình: Visual Studio Code (VSCode)	11
2.3.6. Công cụ quản lý cơ sở dữ liệu: MySQL Workbench.....	12
2.3.7. Công cụ hỗ trợ vẽ sơ đồ: Draw.io	13
2.3.8. Phần mềm thiết kế đồ họa: Adobe Illustrator	14
CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG.....	16
3.1. Mô tả hệ thống.....	16
3.2. Kiến trúc hệ thống	16
3.3. Thiết kế thành phần xử lý	17
3.3.1. Sơ đồ chức năng.....	17
3.3.2. Sơ đồ usecase	18
3.3.3. Sơ đồ tuần tự “Thống kê cho người dùng”	24
3.3.4. Sơ đồ tuần tự “Dự đoán”	25
3.3.5. Lưu đồ hoạt động của hệ thống	26
3.4. Thiết kế thành phần dữ liệu	30
3.4.1. Mô hình dữ liệu.....	30
3.4.2. Thu thập dữ liệu	33
3.4.3. Tự động hóa trong việc thu thập dữ liệu	37
3.4.4. Các phương pháp thống kê	38
3.4.5. Phương pháp tạo vé và quay thử.....	43
3.5. Thiết kế giao diện	45
3.5.1. Giao diện trang chủ.....	45
3.5.2. Giao diện chọn miền.....	46
3.5.3. Giao diện thông báo, thông tin	46
3.5.4. Giao diện thống kê.....	47
3.5.5. Giao diện hiển thị kết quả thống kê	48
3.5.6. Giao diện dự đoán.....	48
CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU.....	51
4.1. Môi trường thực hiện.....	51

4.2. Đánh giá kiểm thử	51
CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....	54
5.1. Kết quả đạt được.....	54
5.2. Những hạn chế.....	54
5.3. Hướng phát triển.....	54
TÀI LIỆU THAM KHẢO	56
PHỤ LỤC	57

DANH MỤC BẢNG

Bảng 3.1. Mô tả Usecase toàn bộ hệ thống	22
Bảng 3.2. Bảng dữ liệu xổ số	31
Bảng 3.3. Tạo mô hình cho cơ sở dữ liệu	32
Bảng 3.4. Mô tả quy trình thu thập dữ liệu.	34
Bảng 3.5. Các thành phần có trong các tỉnh để thu thập dữ liệu.....	35
Bảng 3.6. Các phần tử được chọn cho Miền Nam và Miền Trung.	36
Bảng 3.7. Các phần tử được chọn cho Miền Bắc.....	37
Bảng 3.8. Bảng xử lý tác vụ hẹn giờ thu thập dữ liệu.....	37
Bảng 3.9. Xử lý con số đầu tiên.	39
Bảng 3.10. Xử lý con số cuối.....	39
Bảng 3.11. Xử lý 2 số cuối.....	40
Bảng 3.12. Xử lý 3 số cuối.....	40
Bảng 3.13. Xử lý 2 con số đầu.....	41
Bảng 3.14. Xử lý chọn ra 10 giá trị nhiều nhất và ít nhất.	41
Bảng 3.15. Xử lý 2 con số ở các vị trí khác nhau.	42
Bảng 3.16. Các thành phần xử lý tạo vé.	43
Bảng 3.17. Các thành phần xử lý quay số.....	44
Bảng 3.18. Bảng hiệu năng của các trang web khi thống kê.	52
Bảng 3.19. Bảng so sánh các website.	53

DANH MỤC HÌNH

Hình 1.1. Trang xổ số Minh Ngọc.	2
Hình 1.2. Trang web xoso.com.vn	2
Hình 3.1. Kiến trúc của hệ thống.	16
Hình 3.2. Sơ đồ chức năng của hệ thống.	17
Hình 3.3. Sơ đồ usecase tổng quát của hệ thống.....	19
Hình 3.4. Sơ đồ usecase của tác nhân người dùng.....	20
Hình 3.5. Sơ đồ usecase của tác nhân người quản trị trong hệ th.....	21
Hình 3.6. Sơ đồ tuần tự chức năng thống kê.....	24
Hình 3.7. Sơ đồ tuần tự chức năng dự đoán.....	25
Hình 3.8. Lưu đồ hoạt động của hệ thống.	27
Hình 3.9. Lưu đồ hoạt động cho chức năng thống kê.	28
Hình 3.10. Lưu đồ hoạt động cho chức năng dự đoán.	29
Hình 3.11. Lưu đồ hoạt động chức năng xem tin tức.....	30
Hình 3.12. Giao diện thống kê của trang web xoso.com.vn	33
Hình 3.13. Giao diện trang chủ Xổ Vàng.....	45
Hình 3.14. Giao diện chọn miền của Xổ Vàng	46
Hình 3.15. Giao diện xem tin tức trong Xổ Vàng.....	46
Hình 3.16. Giao diện thống kê	47
Hình 3.17. Giao diện kết quả thống kê kèm biểu đồ.....	48
Hình 3.18. Giao diện dự đoán.	49
Hình 3.19. Giao diện kết quả khi dự đoán.	49
Hình 3.20. Giao diện lịch sử.	50
Hình 3.21. Độ chính xác của kết quả.	52

TÓM TẮT

Trong bối cảnh hiện nay, việc theo dõi và phân tích kết quả xổ số ngày càng trở nên quan trọng, đặc biệt khi nhu cầu thống kê và dự đoán kết quả xổ số gia tăng. Tuy nhiên, việc tra cứu và thống kê thủ công không chỉ tốn nhiều thời gian mà còn dễ dẫn đến sai sót. Chính vì vậy, việc xây dựng một hệ thống tự động hóa và dễ sử dụng để thu thập, lưu trữ và phân tích kết quả xổ số trở nên cần thiết.

Vậy nên hệ thống thống kê kết quả xổ số kiến thiết trên nền tảng web hướng đến phát triển một hệ thống trực tuyến với công nghệ Framework Django và MySQL Workbench để quản lý và lưu trữ dữ liệu kết quả xổ số. Hệ thống cung cấp các chức năng chính bao gồm: hiển thị kết quả xổ số theo từng tỉnh thành, thống kê tần suất xuất hiện của các con số, các biểu đồ, chọn ra các con số xuất hiện nhiều nhất và tính năng quay thử dự đoán kết quả. Giao diện người dùng được thiết kế thân thiện, dễ sử dụng, giúp người dùng dễ dàng truy cập các thông tin cần thiết. Ngoài ra, hệ thống còn hỗ trợ tự động thu thập dữ liệu từ các nguồn trực tuyến và có khả năng mở rộng, nâng cấp trong tương lai.

ABSTRACT

In the current context, tracking and analyzing lottery results has become increasingly important, especially with the growing demand for statistical analysis and predictions. However, manual tracking and analysis are time-consuming and prone to errors. Therefore, the development of an automated and user-friendly system to collect, store, and analyze lottery results is essential.

The "Lottery Result Statistics System on a Web Platform" aims to create an online system using the Django framework and MySQL Workbench to manage and store lottery results data. The system provides key features, including displaying lottery results by province, analyzing the frequency of number occurrences, displaying charts, identifying the most frequently drawn numbers, and offering a trial prediction feature. The user interface is designed to be intuitive and easy to use, allowing users to quickly access the necessary information. Additionally, the system supports automatic data collection from online sources and offers scalability and the potential for future upgrades.

CHƯƠNG 1. GIỚI THIỆU

Chương này giới thiệu bối cảnh, mục tiêu và phạm vi nghiên cứu, định hướng xây dựng hệ thống thống kê kết quả xổ số kiến thiết trên web.

1.1. Đặt vấn đề

Trong thời đại số hóa hiện nay, việc sử dụng công nghệ thông tin tạo ra các hệ thống thông tin thông minh và hữu ích ngày càng trở nên quan trọng. Trong lĩnh vực thống kê dự đoán, xổ số kiến thiết là một trong những lĩnh vực thu hút sự quan tâm của nhiều người. Tuy nhiên, việc phân tích và thống kê kết quả xổ số một cách hiệu quả có thể gặp khó khăn.

Đa phần, người dùng trong việc tiếp cận và sử dụng công cụ hỗ trợ thống kê, phân tích và dự đoán kết quả xổ số có thể gặp một số thách thức sau: Thứ nhất, các thông tin liên quan đến kết quả xổ số đều được cung cấp trên rất nhiều nền tảng khác nhau, không phải lúc nào nó cũng được trình bày dưới dạng rõ ràng, đầy đủ lại dễ tìm như người dùng mong muốn. Điều này gây khó khăn cho người dùng trong việc tổng hợp, so sánh, và đưa ra những dự đoán có cơ sở. Không những thế, việc tự thống kê hoặc dự đoán bằng các phương pháp truyền thống tiêu tốn thời gian mà còn có sai sót do yếu tố con người. Người dùng không chỉ cần một hệ thống chính xác mà còn phải trực quan và thân thiện để dễ dàng thao tác, ngay cả với những người không rành về công nghệ. Cuối cùng, trong thời đại phát triển của công nghệ hiện nay, người dùng ngày càng mong chờ một công cụ có khả năng không chỉ lưu trữ dữ liệu mà còn hỗ trợ các chức năng tiên tiến như tự động phân tích xu hướng, dự đoán kết quả và cung cấp báo cáo thống kê chi tiết. Do đó, một hệ thống thống kê và phân tích kết quả xổ số tối ưu sẽ mang lại nhiều giá trị thiết thực, đáp ứng đúng nhu cầu ngày càng cao của người dùng.

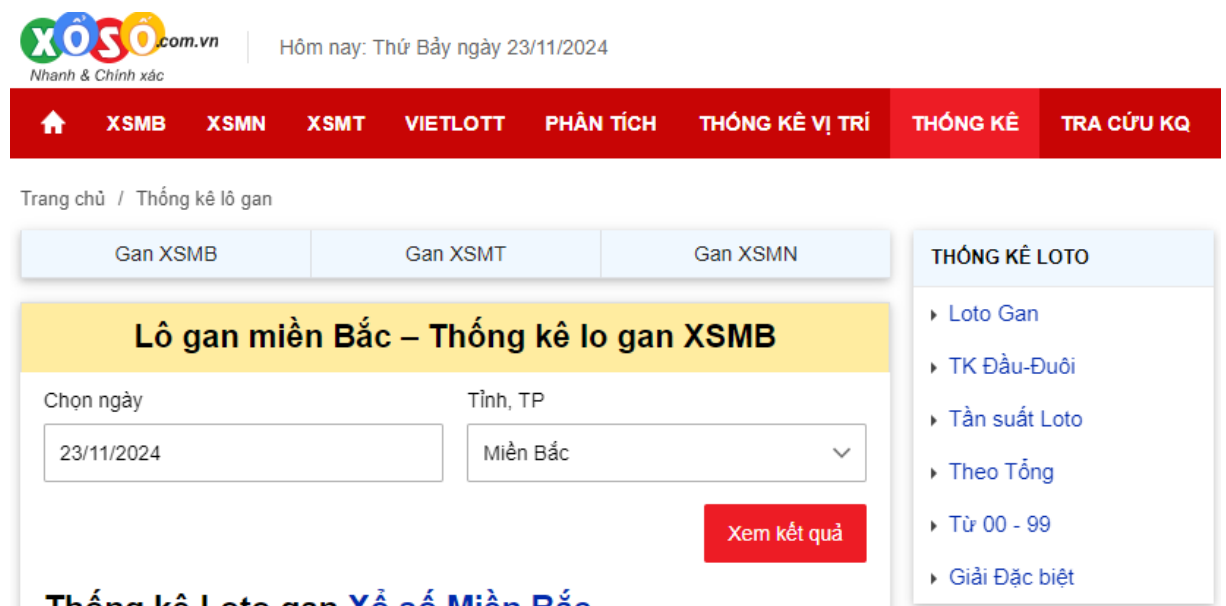
Hiện nay có một số trang web thống kê tiêu biểu như:

+ Trang web xổ số Minh Ngọc



Hình 1.1. Trang xổ số Minh Ngọc.

+ Trang web xoso.com.vn



Hình 1.2. Trang web xoso.com.vn

Các website này cung cấp các tính năng thống kê phổ biến như thống kê tần suất, lô gan, và đầu - đuôi số. Tuy nhiên, việc hiển thị các giải chưa thực sự rõ ràng, gây khó khăn cho người mới tiếp cận với thống kê. Đặc biệt, giao diện có phần phức tạp và thường xuyên xuất hiện quảng cáo, làm ảnh hưởng đến trải nghiệm người dùng.

1.2. Mục tiêu đề tài

Mục tiêu chính của đề tài là xây dựng một hệ thống thống kê kết quả xổ số kiến thiết trên nền tảng web với khả năng cập nhật, tra cứu, thống kê, cung cấp thông tin chính xác để hỗ trợ người dùng trong việc dự đoán và nắm bắt xu hướng của kết quả xổ số. Dưới đây là mô tả về các mục tiêu chính của đề tài:

- Truy cập thuận tiện và dễ dàng: Việc truy cập thông tin kết quả xổ số kiến thiết cần phải thuận tiện, nhanh chóng và dễ dàng đối với người dùng. Hệ thống cần phải cung cấp một giao diện web thân thiện, tối ưu hóa các kích thước hiển thị cho máy tính hay các thiết bị điện tử có thể truy cập với các kích thước khác nhau.
- Tính chính xác và đáng tin cậy: Thông tin về kết quả xổ số cần phải được cập nhật đúng lịch và chính xác. Người dùng cần tin tưởng vào tính đáng tin cậy của dữ liệu được hiển thị trên hệ thống.
- Phân tích và thống kê thông tin: Hệ thống cần cung cấp khả năng phân tích và thống kê các kết quả xổ số kiến thiết. Điều này giúp người dùng có cái nhìn rõ ràng về xu hướng và cơ hội trúng thưởng.

1.3. Đối tượng nghiên cứu

Đối tượng nghiên cứu của đề tài bao gồm:

+ Dữ liệu kết quả xổ số kiến thiết:

- Kết quả xổ số từ các tỉnh, thành phố trên cả nước theo từng giải và từng ngày.
- Tần suất xuất hiện của các con số trong các giải thưởng.

+ Phương pháp thống kê và phân tích:

- Các kỹ thuật thống kê dữ liệu và xác suất để phân tích xu hướng.

+ Công nghệ xây dựng hệ thống:

- Nền tảng phát triển web (Django) và cơ sở dữ liệu (MySQL).
- Công cụ lập trình thu thập dữ liệu tự động (Playwright).

+ Người dùng hệ thống:

- Người dùng có nhu cầu tra cứu, thống kê, và sử dụng tính năng dự đoán ngẫu nhiên của kết quả xổ số.

1.4. Mục tiêu nghiên cứu

1.4.1. Mục tiêu tổng quát

Mục tiêu chính là nghiên cứu cách vận hành của xổ số từ đó xây dựng hệ thống thống kê và phân tích kết quả xổ số kiến thiết trên nền tảng web, tích hợp tính năng dự đoán để hỗ trợ người dùng.

1.4.2. Mục tiêu cụ thể

- Phát triển hệ thống tự động thu thập và lưu trữ kết quả xổ số từ nhiều tỉnh, thành phố.
- Thiết kế giao diện web thân thiện, cung cấp công cụ thống kê và phân tích dữ liệu.
- Đảm bảo hệ thống hoạt động ổn định và đáp ứng tốt nhu cầu của người dùng.

1.5. Phạm vi nghiên cứu

1.5.1. Phạm vi dữ liệu

- Đề tài tập trung vào kết quả xổ số kiến thiết của các tỉnh, thành phố ở ba miền tại Việt Nam.
- Dữ liệu được giới hạn trong một khoảng thời gian nhất định: Miền Nam và Miền Trung được thu thập trong khoảng 10 năm, Miền Bắc khoảng 8 năm.

1.5.2. Phạm vi chức năng của hệ thống:

- Thống kê: phân tích tần suất xuất hiện của các con số theo giải, theo tỉnh, và theo số lượng con số.
- Dự đoán: Cung cấp dự đoán các con số ngẫu nhiên và cho lịch sử để người dùng quan sát
- Giao diện: Thiết kế giao diện web hiển thị kết quả xổ số, công cụ thống kê và phân tích một cách trực quan, dễ sử dụng.

1.5.3. Phạm vi công nghệ

- Hệ thống được phát triển trên nền tảng web sử dụng framework Django và cơ sở dữ liệu MySQL.
- Công cụ Playwright được sử dụng để thu thập dữ liệu tự động từ các trang web xổ số chính thống.
- Các thuật toán thống kê sẽ được áp dụng để phân tích dữ liệu.

1.5.4. Phạm vi người dùng

- Người dùng có nhu cầu thống kê kết quả xổ số.
- Người sử dụng đòi hỏi tính phân tích dữ liệu để tìm kiếm số liệu tiềm năng.
- Người sử dụng không đòi hỏi chuyên môn cao về kỹ thuật để sử dụng hệ thống.

1.5.5. Phạm vi thời gian

- Đề tài này được thực hiện trong khoảng thời gian đầu tháng 8 đến cuối tháng 11
- Hệ thống sẽ ưu tiên xử lý dữ liệu hiện tại và quá khứ gần để đảm bảo tính thực tiễn của kết quả.

1.6. Phương pháp nghiên cứu

Sau khi chọn đề tài, đầu tiên là xác định các yêu cầu cụ thể của hệ thống cũng như phạm vi nghiên cứu để đảm bảo rằng hướng phát triển phù hợp với nhu cầu thực tế, tham khảo các trang web tìm các ưu và nhược điểm, các cơ chế của xổ số. Quá trình thu thập dữ liệu là bước nền tảng, trong đó các kết quả xổ số từ các trang web uy tín sẽ được tự động thu thập qua công cụ Playwright. Công cụ này cho phép quá trình thu thập diễn ra nhanh chóng và chính xác, loại bỏ những hạn chế của phương pháp thủ công. Sau đó, dữ liệu sẽ được lưu lại cơ sở dữ liệu tạo tiền đề cho các bước phân tích tiếp theo.

Dữ liệu thu thập sẽ được phân tích để tìm ra quy luật trong kết quả xổ số. Các thống kê sẽ được sử dụng để tính toán tần suất xuất hiện của từng con số theo thời gian, giúp người dùng có cái nhìn tổng quan. Hệ thống sẽ được phát triển trên nền tảng web sử dụng framework Django, mang lại sự linh hoạt và mạnh mẽ. Giao diện sẽ được thiết kế thân thiện và trực quan, giúp người dùng dễ dàng truy cập và sử dụng các tính năng thống kê, phân tích và dự đoán. Sẽ tích hợp các biểu đồ, bảng số liệu và công cụ tra cứu để hỗ trợ tối đa nhu cầu của người dùng.

Trong quá trình phát triển, hệ thống sẽ được kiểm thử liên tục để đảm bảo tính chính xác và ổn định. Những kiểm tra bao gồm việc xác minh dữ liệu thu thập, đánh giá hiệu suất của hệ thống và thử nghiệm về độ chính xác của các thuật toán dự đoán. Kết quả của những thử nghiệm này sẽ được sử dụng để thực hiện tối ưu hóa của hệ thống nhằm cải thiện trải nghiệm người dùng và làm cho sản phẩm cuối cùng tốt hơn.

1.7. Ý nghĩa của đề tài nghiên cứu

Đề tài "Xây dựng hệ thống thống kê và phân tích kết quả xổ số kiến thiết trên nền tảng web" mang lại ý nghĩa cả về mặt khoa học lẫn thực tiễn, góp phần giải quyết một bài toán trong lĩnh vực thống kê và dự đoán.

Về mặt khoa học, đề tài thể hiện các phương pháp thống kê truyền thống để phân tích và dự đoán dữ liệu một cách hiệu quả. Quá trình triển khai hệ thống là một ví dụ cụ thể về việc ứng dụng các nền tảng công nghệ như framework Django và cơ sở dữ liệu MySQL để xây dựng một hệ thống thông minh và tự động hóa. Điều này không chỉ làm phong phú nền tảng lý thuyết về khoa học dữ liệu và công nghệ thông tin mà còn mở ra những hướng nghiên cứu mới.

Về mặt thực tiễn, hệ thống này mang lại lợi ích cho người dùng trong việc phân tích và dự đoán kết quả xổ số kiến thiết, đặc biệt là những người mới tiếp xúc với xổ số kiến thiết. Xổ số không chỉ là một hoạt động giải trí phổ biến mà còn có tác động lớn đến nguồn thu ngân sách nhà nước. Việc cung cấp một hệ thống trực quan và dễ sử dụng giúp người dùng tiếp cận thông tin một cách nhanh chóng và chính xác, từ đó hỗ trợ họ đưa ra các quyết định hiệu quả hơn. Ngoài ra, giao diện thân thiện và khả năng hiển thị dữ liệu trực quan của hệ thống góp phần nâng cao trải nghiệm người dùng.

1.8. Bố cục của quyển luận văn

Bố cục của quyển luận văn này bao gồm 5 chương

Chương 1: Giới thiệu

Chương 2: Cơ sở lý thuyết

Chương 3: Phân tích và thiết kế hệ thống

Chương 4: Kết quả nghiên cứu

Chương 5: Kết luận và hướng phát triển

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

Chương này giới thiệu các phương pháp thống kê, công nghệ cần thiết các kỹ thuật phân tích dữ liệu để phát triển hệ thống thống kê kết quả xổ số.

2.1. Frontend

2.1.1. HTML (Hypertext Markup Language)

HTML (Ngôn ngữ đánh dấu siêu văn bản) là ngôn ngữ cơ bản nhất để xây dựng một trang web. Đây là nền tảng cấu trúc của mọi trang web hiện nay. HTML giúp mô tả cấu trúc và tổ chức của các tài liệu web thông qua các thẻ (tags), cho phép trình duyệt web hiển thị nội dung của các trang web đó. HTML không chịu trách nhiệm về phần giao diện hay các tính năng động của trang web, mà chỉ đơn giản là cấu trúc của thông tin.

+ Trang HTML có một cấu trúc rất cơ bản, được chia thành ba phần chính:

- Doctype: Thẻ `<!DOCTYPE html>` cho biết đây là một tài liệu HTML5.
- Thẻ `<html>`: Là thẻ bao bọc toàn bộ nội dung của tài liệu HTML.
- Thẻ `<head>`: Chứa các thông tin không hiển thị trên trang như liên kết đến các tệp CSS, thông tin meta (charset, tác giả, v.v.) và các kịch bản JavaScript.
- Thẻ `<body>`: Phần này chứa nội dung chính mà người dùng nhìn thấy khi truy cập trang web, bao gồm văn bản, hình ảnh, liên kết, bảng biểu, v.v.

+ Các phân tử HTML được định nghĩa bởi các thẻ mở và thẻ đóng

- Đoạn văn bản: `<p>` dùng để chèn văn bản vào trang.
- Tiêu đề: Các thẻ `<h1>` đến `<h6>` dùng để đánh dấu tiêu đề các cấp độ, với `<h1>` là tiêu đề chính và `<h6>` là tiêu đề phụ.
- Liên kết: Thẻ `<a>` giúp tạo liên kết đến các trang khác hoặc các phần khác trong cùng một trang, thông qua thuộc tính href.
- Hình ảnh: Thẻ `` được sử dụng để nhúng hình ảnh vào trang.

HTML là công cụ duy nhất giúp trình duyệt hiểu được cấu trúc của nội dung trang web. Tuy nhiên, HTML không có khả năng tạo kiểu dáng hay các tính năng tương tác động. Đó là lý do tại sao các công nghệ như CSS và JavaScript được sử dụng phối hợp với HTML để xây dựng giao diện người dùng đẹp mắt và tính năng động.

2.1.2 CSS (Cascading Style Sheets)

CSS là một ngôn ngữ thiết kế, sử dụng để điều chỉnh giao diện và cách hiển thị của các phần tử HTML trên trang web. CSS không thay đổi nội dung của trang web mà chỉ điều chỉnh cách những phần tử HTML đó được trình bày.

+ Lợi ích và đặc điểm của CSS:

- Tách biệt nội dung và giao diện: Trong CSS có rất nhiều thế mạnh mà ta có thể kể đến là nó giúp phân biệt nội dung của HTML với giao diện CSS, giúp cho việc đọc mã nguồn được dễ dàng và bảo quản.
- Responsiveness (Khả năng phản ứng): CSS hỗ trợ thiết kế đáp ứng, tức là trang web có thể tự động điều chỉnh giao diện để phù hợp với các kích thước màn hình khác nhau, từ desktop đến di động.
- Hiệu suất và mở rộng: Dựa vào các sử dụng hợp lý của CSS, trang web có thể load nhanh hơn, giảm tối đa dung lượng load trang mà dễ dàng mở rộng tính năng mới mà không cần thay đổi mã nguồn toàn bộ.
- Thiết kế linh hoạt: CSS có thể tạo ra các kiểu dáng cho trang web rất đa dạng, từ việc thay đổi màu sắc, kiểu chữ, định dạng văn bản đến các hiệu ứng chuyển động và các bố cục phức tạp như Grid và Flexbox

+ CSS cũng có thể được chia thành ba loại:

- CSS Inline: Được viết trực tiếp vào trong thẻ HTML.
- CSS Internal: Được viết trong phần `<style>` của thẻ `<head>` trong tài liệu HTML.
- CSS External: Được lưu trữ trong tệp riêng biệt và liên kết với trang HTML qua thẻ `<link>`. Đây là cách phổ biến nhất trong phát triển

2.1.3. JavaScript

JavaScript là một ngôn ngữ lập trình phía client (tại trình duyệt) cho phép các nhà phát triển thêm tính năng tương tác và động vào trang web. Không giống như HTML và CSS, JavaScript có thể thay đổi nội dung trang web một cách linh động mà không cần tải lại trang.

Các điểm nổi bật của JavaScript:

- Tính tương tác (User Interaction): JavaScript giúp xây dựng các ứng dụng web tương tác, như việc kiểm tra dữ liệu đầu vào trong biểu mẫu, tạo thông báo, và

thay đổi nội dung trang dựa trên hành vi người dùng (nhấp chuột, di chuột, nhập liệu, v.v).

- Thao tác với DOM (Document Object Model): JavaScript có thể thay đổi nội dung HTML bằng cách thao tác với DOM. DOM là một mô hình đối tượng được tạo ra từ mã HTML của trang web, cho phép JavaScript đọc và thay đổi cấu trúc tài liệu.
- Xử lý sự kiện (Event Handling): JavaScript hỗ trợ các sự kiện như click, keyup, scroll, giúp trang web đáp ứng nhanh với các hành động của người dùng.

2.2. Backend

Django là một framework phát triển ứng dụng web mạnh mẽ sử dụng Python. Django hỗ trợ lập trình web nhanh chóng, giúp bạn xây dựng các ứng dụng web phức tạp với ít dòng mã hơn và bảo mật cao. Dưới đây là các tính năng nổi bật của Django:

- Django sử dụng mô hình MVC (Model-View-Controller), giúp phân chia rõ ràng các phần trong ứng dụng: Model quản lý dữ liệu, View xử lý logic, và Template xử lý giao diện người dùng.
- Hệ thống quản trị admin: Django cung cấp hệ thống quản trị tự động rất mạnh mẽ cho phép quản lý và thao tác dữ liệu dễ dàng mà không cần viết mã.
- Bảo mật: Django có các cơ chế bảo mật tự động, bao gồm chống lại các cuộc tấn công như SQL Injection, Cross-Site Scripting (XSS), và Cross-Site Request Forgery (CSRF).
- Khả năng mở rộng và bảo trì: Django khuyến khích sử dụng các ứng dụng độc lập và có thể tái sử dụng, giúp dễ dàng mở rộng hệ thống khi cần.

2.3. Các công cụ và các thư viện

2.3.1. Playwright¹

Playwright là một công cụ tự động hóa trình duyệt do Microsoft phát triển, giúp kiểm thử và tự động hóa các thao tác trên trình duyệt. Nó hỗ trợ các trình duyệt phổ biến như Chromium, Firefox, và WebKit, và có thể tích hợp với các framework kiểm thử như Jest. Dưới đây là các điểm mạnh của Playwright:

¹ <https://github.com/microsoft/playwright-examples>

- Đa trình duyệt: Playwright hỗ trợ kiểm thử trên nhiều trình duyệt, giúp đảm bảo rằng ứng dụng web hoạt động ổn định trên các trình duyệt khác nhau.
- Tự động hóa tác vụ phức tạp: Playwright có thể thực hiện các tác vụ như đăng nhập, tải trang, xử lý popup, hoặc điền vào các biểu mẫu mà không cần can thiệp thủ công.
- Chế độ headless: Playwright có thể chạy trong chế độ không có giao diện người dùng (headless), giúp tăng tốc quá trình kiểm thử và tự động hóa.

2.3.2. Chart.js²

Chart.js là một thư viện mã nguồn mở của JavaScript được thiết kế để tạo các biểu đồ trực quan trên trang web. Thư viện này cung cấp nhiều loại biểu đồ khác nhau như biểu đồ đường, cột, tròn và nhiều loại khác. Dưới đây là các điểm nổi bật của Chart.js:

- Dễ sử dụng: Chart.js có cú pháp đơn giản và dễ hiểu, giúp người phát triển tạo ra các biểu đồ một cách nhanh chóng mà không cần phải viết mã phức tạp.
- Hỗ trợ nhiều loại biểu đồ: Thư viện hỗ trợ nhiều loại biểu đồ đa dạng, bao gồm biểu đồ đường, cột, tròn, bánh, radar, scatter và nhiều loại khác, giúp người dùng linh hoạt trong việc lựa chọn loại biểu đồ phù hợp với nhu cầu.
- Tương tác và thanh cuộn: Chart.js cung cấp khả năng tương tác, như hiển thị thông tin chi tiết khi người dùng rê chuột qua các điểm dữ liệu trên biểu đồ. Thư viện cũng hỗ trợ thanh cuộn cho các biểu đồ lớn, giúp dễ dàng duyệt qua các dữ liệu chi tiết.
- Tùy chỉnh linh hoạt: Người phát triển có thể tùy chỉnh nhiều yếu tố của biểu đồ, từ màu sắc, độ dày của nét vẽ cho đến cách thức hiển thị các nhãn, mang lại sự linh hoạt cao trong việc thiết kế giao diện người dùng.
- Responsive (Phản hồi linh hoạt): Chart.js hỗ trợ tính năng phản hồi, tức là biểu đồ có thể tự động thay đổi kích thước và điều chỉnh giao diện để phù hợp với các kích thước màn hình khác nhau, mang đến trải nghiệm người dùng tốt trên mọi thiết bị.
- Thư viện nhẹ: Chart.js có kích thước nhẹ, giúp tối ưu hóa hiệu suất và giảm thời gian tải trang, điều này đặc biệt quan trọng trong các ứng dụng web cần tải nhanh.

² <https://github.com/chartjs>

- Dữ liệu động: Chart.js có khả năng cập nhật và hiển thị biểu đồ với dữ liệu thay đổi theo thời gian mà không cần phải tải lại toàn bộ trang, mang đến trải nghiệm người dùng mượt mà và hiệu quả.
- Tích hợp dễ dàng: Thư viện có thể dễ dàng tích hợp vào các dự án web sử dụng HTML, CSS và JavaScript. Nó cũng tương thích tốt với các framework như Bootstrap, giúp xây dựng các ứng dụng web nhanh chóng.
- Tài liệu chi tiết: Thư viện cung cấp tài liệu hướng dẫn chi tiết và nhiều ví dụ thực tế, giúp các nhà phát triển dễ dàng tiếp cận và triển khai vào dự án của mình.

2.3.3. html2canvas³

Thư viện html2canvas là công cụ mã nguồn mở dùng để chụp màn hình các phần tử HTML và chuyển đổi chúng thành hình ảnh. Nó dựng lại cây DOM và style của trang web, sau đó render lên một đối tượng canvas. Dưới đây là các tính năng chính:

- Chụp toàn bộ hoặc một phần giao diện HTML.
- Hỗ trợ hầu hết các thuộc tính CSS phổ biến.
- Xuất dữ liệu dưới dạng hình ảnh (PNG, JPEG).

2.3.4. jsPDF⁴

Thư viện jsPDF là công cụ mã nguồn mở cho phép tạo và quản lý tệp PDF trực tiếp bằng JavaScript, thường dùng để xuất báo cáo hoặc lưu trữ nội dung trang web dưới dạng PDF. Dưới đây là các tính năng chính:

- Tạo file PDF từ nội dung HTML, hình ảnh hoặc dữ liệu JSON.
- Hỗ trợ chèn văn bản, hình ảnh với tùy chỉnh kích thước, màu sắc, căn lề.
- Tự động phân trang khi nội dung vượt quá kích thước trang.
- Tùy chỉnh kích thước PDF (A4, Letter, hoặc kích thước tự chọn).

2.3.5. Công cụ hỗ trợ lập trình: Visual Studio Code (VSCode)

Visual Studio Code (VSCode) là một trình soạn thảo mã nguồn phổ biến, được nhiều lập trình viên sử dụng nhờ các tính năng mạnh mẽ và dễ sử dụng. Dưới đây là một số điểm mạnh của Visual Studio Code:

³ <https://github.com/retifrav/html2canvas-example?tab=readme-ov-file>

⁴ <https://github.com/parallax/jsPDF>

- Mã nguồn mở và miễn phí: VSCode là phần mềm mã nguồn mở và miễn phí, giúp người dùng dễ dàng cài đặt và sử dụng mà không phải trả bất kỳ chi phí nào.
- Hỗ trợ nhiều ngôn ngữ: VSCode hỗ trợ nhiều ngôn ngữ lập trình như JavaScript, Python, Java, PHP, và nhiều ngôn ngữ khác, giúp đáp ứng nhu cầu phát triển đa dạng.
- Tiện ích mở rộng (Extensions): VSCode có hệ thống tiện ích mở rộng phong phú, cho phép người dùng mở rộng và tùy chỉnh các tính năng của phần mềm để phù hợp với nhu cầu phát triển cụ thể.
- IntelliSense: Cung cấp tính năng tự động hoàn thành mã và gợi ý thông minh, giúp tăng tốc quá trình lập trình và giảm thiểu lỗi.
- Debugging nâng cao: VSCode tích hợp khả năng gỡ lỗi mạnh mẽ với bảng điều khiển gỡ lỗi, cho phép xem giá trị biến và theo dõi các bước thực thi mã, giúp phát hiện lỗi hiệu quả.
- Hệ thống Quản lý Phiên bản: Tích hợp với các hệ thống quản lý phiên bản như Git, giúp theo dõi lịch sử thay đổi và quản lý mã nguồn một cách dễ dàng.
- Tích hợp Terminal: VSCode tích hợp terminal trong môi trường làm việc, giúp người dùng thực hiện các lệnh hệ thống mà không cần thoát khỏi phần mềm.
- Hỗ trợ đa nền tảng: VSCode hỗ trợ cả Windows, macOS và Linux, mang lại sự linh hoạt khi sử dụng trên các hệ điều hành khác nhau.
- Cộng đồng phát triển lớn: VSCode có cộng đồng người dùng lớn, cùng với nhiều tài nguyên học tập và tiện ích mở rộng được chia sẻ, giúp người dùng dễ dàng tìm kiếm sự hỗ trợ.

2.3.6. Công cụ quản lý cơ sở dữ liệu: MySQL Workbench

MySQL Workbench là công cụ mạnh mẽ giúp quản lý và tối ưu hóa cơ sở dữ liệu MySQL. Dưới đây là những điểm mạnh của MySQL Workbench:

- Giao diện đồ họa (GUI): MySQL Workbench cung cấp giao diện đồ họa để sử dụng, giúp người dùng dễ dàng tương tác với cơ sở dữ liệu MySQL mà không cần phải viết mã phức tạp.
- Tạo và quản lý cơ sở dữ liệu: Công cụ cho phép người dùng tạo mới cơ sở dữ liệu, bảng và chỉnh sửa cấu trúc của chúng thông qua giao diện đồ họa.

- Trình soạn thảo SQL: MySQL Workbench cung cấp trình soạn thảo SQL, cho phép người dùng viết và thực thi các truy vấn SQL một cách thuận tiện.
- Thực hiện sao lưu và khôi phục: MySQL Workbench hỗ trợ sao lưu (backup) và khôi phục (restore) cơ sở dữ liệu, giúp bảo vệ dữ liệu quan trọng của bạn.
- Quản lý người dùng và quyền truy cập: Công cụ này giúp người quản trị cơ sở dữ liệu quản lý người dùng và quyền truy cập vào các cơ sở dữ liệu, đảm bảo an toàn và bảo mật.
- Tối ưu hóa cơ sở dữ liệu: MySQL Workbench cung cấp các công cụ để tối ưu hóa cơ sở dữ liệu, cải thiện hiệu suất và đảm bảo tính ổn định của các bảng.
- Mô hình hóa cơ sở dữ liệu: MySQL Workbench cho phép vẽ mô hình cơ sở dữ liệu, giúp người dùng thiết kế cấu trúc cơ sở dữ liệu trước khi triển khai.
- Đồng bộ hóa cơ sở dữ liệu: Hỗ trợ đồng bộ hóa cơ sở dữ liệu giữa máy chủ và máy tính cá nhân, giúp duy trì tính đồng nhất dữ liệu.
- Visual Explain: Công cụ giúp người dùng phân tích và tối ưu hóa các câu truy vấn SQL thông qua việc xem xét các kế hoạch thực thi truy vấn.
- Cross-Platform: MySQL Workbench hỗ trợ nhiều nền tảng, bao gồm Windows, macOS và Linux, giúp người dùng sử dụng trên nhiều hệ điều hành khác nhau.

2.3.7. Công cụ hỗ trợ vẽ sơ đồ: Draw.io

Draw.io là một công cụ trực tuyến mạnh mẽ giúp người dùng tạo sơ đồ một cách nhanh chóng và dễ dàng. Dưới đây là những điểm nổi bật của Draw.io:

- Trực tuyến và miễn phí: Draw.io là một công cụ web miễn phí, giúp người dùng dễ dàng truy cập và tạo sơ đồ mà không cần cài đặt phần mềm.
- Đa dạng loại sơ đồ: Công cụ hỗ trợ nhiều loại sơ đồ như sơ đồ UML, sơ đồ luồng công việc, biểu đồ tư duy, sơ đồ mô hình dữ liệu và các loại sơ đồ khác.
- Giao diện dễ sử dụng: Giao diện của Draw.io được thiết kế đơn giản, giúp người dùng nhanh chóng làm quen và bắt đầu tạo sơ đồ mà không gặp khó khăn.
- Tích hợp với các dịch vụ lưu trữ: Draw.io cho phép tích hợp với các dịch vụ lưu trữ đám mây như Google Drive, OneDrive, Dropbox và GitHub, giúp dễ dàng lưu trữ và chia sẻ các sơ đồ.

- Hỗ trợ xuất và nhúng: Người dùng có thể xuất sơ đồ dưới nhiều định dạng như PNG, JPEG, SVG, PDF và có thể nhúng sơ đồ vào trang web hoặc ứng dụng của mình.
- Bảo mật: Draw.io cung cấp các tính năng bảo mật, bao gồm khả năng tạo sơ đồ offline và hỗ trợ các dịch vụ lưu trữ an toàn.
- Hỗ trợ vẽ tự do: Cho phép vẽ tự do và tùy chỉnh hình dạng của các đối tượng trong sơ đồ, mang lại sự linh hoạt trong thiết kế.
- Tính năng Undo và Redo: Draw.io hỗ trợ các tính năng "Undo" và "Redo", giúp người dùng quay lại hoặc lặp lại các thao tác một cách linh hoạt.
- Cộng đồng hỗ trợ lớn: Công cụ này có một cộng đồng người sử dụng lớn và năng động, cung cấp nhiều tài nguyên học liệu và hỗ trợ trực tuyến cho người dùng.

2.3.8. Phần mềm thiết kế đồ họa: Adobe Illustrator

Adobe Illustrator là phần mềm đồ họa vector mạnh mẽ được sử dụng phổ biến để tạo ra các thiết kế đồ họa, logo, minh họa, biểu đồ, đồ họa web và nhiều sản phẩm nghệ thuật khác. Dưới đây là điểm nổi bật của Illustrator.

- Khả năng tạo ra hình ảnh vector, giúp duy trì chất lượng hình ảnh khi thay đổi kích thước, rất hữu ích trong việc thiết kế logo hay các hình ảnh cần phóng to hoặc thu nhỏ nhiều lần mà không bị vỡ nét.
- Illustrator cung cấp một loạt công cụ vẽ mạnh mẽ như Pen Tool, Brush Tool, Shape Tools và Pencil Tool, cho phép người dùng tạo ra các đường nét và hình dạng chính xác. Các hiệu ứng mạnh mẽ, như hiệu ứng 3D, đổ bóng, làm mờ và biến dạng, giúp làm nổi bật các yếu tố trong thiết kế.
- Phần mềm này còn có hệ thống màu sắc linh hoạt với các công cụ Color Picker, Swatches Panel và Gradient Tool, cho phép tạo ra các màu sắc và gradient phong phú. Illustrator hỗ trợ tạo văn bản với nhiều kiểu chữ, kích thước và hiệu ứng đặc biệt, như Type on a Path và Text Effects, giúp sáng tạo và tùy chỉnh kiểu chữ một cách linh hoạt.
- Illustrator tích hợp tốt với các phần mềm khác của Adobe như Photoshop, InDesign và After Effects, tạo điều kiện thuận lợi cho các dự án phối hợp. Công cụ như Pathfinder, Live Paint và Clipping Masks giúp tạo ra các mẫu đồ họa phức tạp từ biểu đồ đến thiết kế tự do.

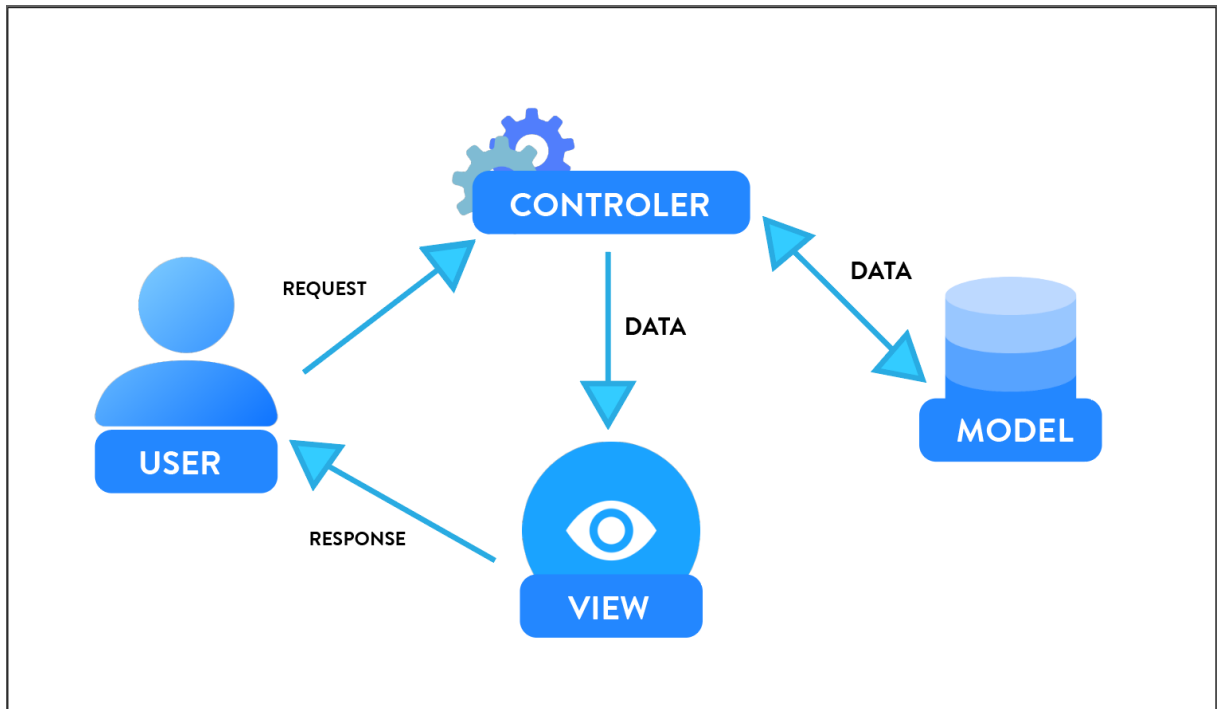
- Phần mềm này hỗ trợ nhiều định dạng tệp đồ họa vector như AI, EPS, SVG và PDF, dễ dàng chia sẻ và xuất bản các thiết kế. Illustrator còn hỗ trợ các tiện ích mở rộng và plug-ins, mở rộng tính năng của phần mềm, và cộng đồng người dùng tích cực cung cấp nhiều tài nguyên hỗ trợ.

CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1. Mô tả hệ thống

Hệ thống thống kê kết quả xổ số kiến thiết trên nền tảng web với các đối tượng chính là người chơi xổ số, quản trị viên đồng thời hệ thống trên sẽ bao gồm các chức năng như xem thống kê, thử dự đoán, xem tin tức hay các thông báo.

3.2. Kiến trúc hệ thống



Hình 3.1. Kiến trúc của hệ thống.

Đề tài này sử dụng mô hình kiến trúc MVC (Model-View-Controller) và đối với Django thì mặc định đã áp dụng mô hình này. Mô hình này đối với đề tài này sẽ bao gồm:

- Model: Lưu trữ dữ liệu kết quả xổ số trong cơ sở dữ liệu MySQL, bao gồm thông tin về các con số ứng với các giải thưởng.
- Controller: Tiếp nhận yêu cầu từ người dùng, truy vấn dữ liệu từ Model và xử lý các logic nghiệp vụ như thống kê tần suất các con số hoặc thực hiện dự đoán.
- View: Hiển thị kết quả đã được xử lý từ Controller dưới dạng giao diện người dùng, bao gồm bảng kết quả xổ số, biểu đồ thống kê, và các phân tích.

3.3. Thiết kế thành phần xử lý

3.3.1. Sơ đồ chức năng



Hình 3.2. Sơ đồ chức năng của hệ thống.

Sơ đồ chức năng của hệ thống "Website Thống Kê Xổ Số" được thiết kế để hiển thị cấu trúc các phần chính của trang web và các chức năng liên quan. Giao diện chính của trang web được chia thành ba phần lớn:

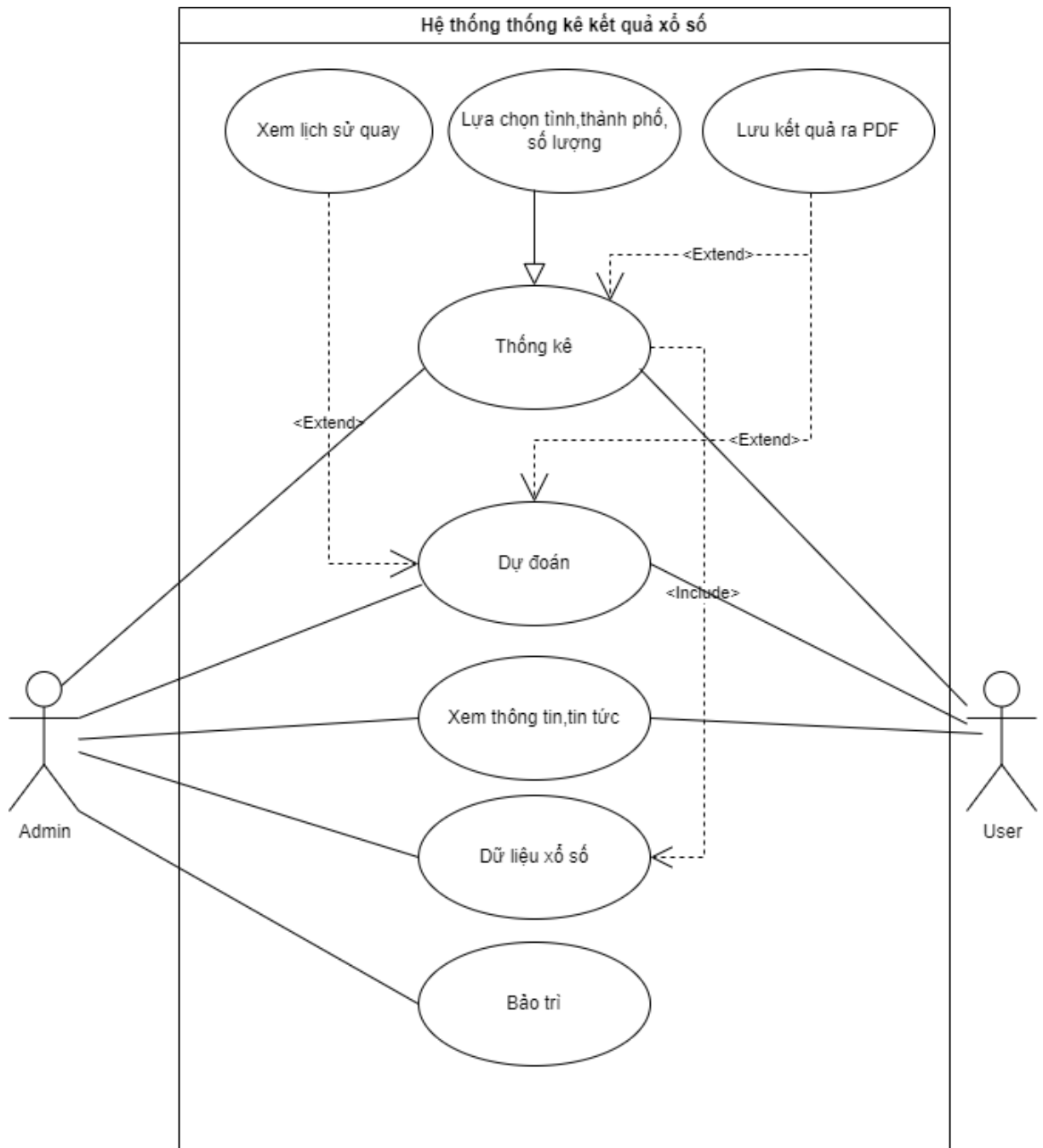
- Trang chủ: Là khu vực bắt đầu của trang web, dẫn đến các mục khác.
- Thống kê: Phần này cho phép người dùng thực hiện các thao tác thống kê. Gồm hai chức năng con:
 - + Lựa chọn để thống kê: Cho phép người dùng chọn các dữ liệu cần thống kê.
 - + Xem thống kê: Hiển thị kết quả thống kê cho người dùng.
- Dự đoán: Cung cấp các công cụ để dự đoán kết quả xổ số, gồm:
 - + Rút vé: Tạo và hiển thị vé xổ số ngẫu nhiên.

- + Quay số: Thực hiện quá trình quay số để tạo ra các kết quả dự đoán.
- + Lịch sử quay: Hiển thị các kết quả quay trước đó.
- Thông tin: Cung cấp thông tin liên quan đến hệ thống và cách sử dụng.

3.3.2. Sơ đồ usecase

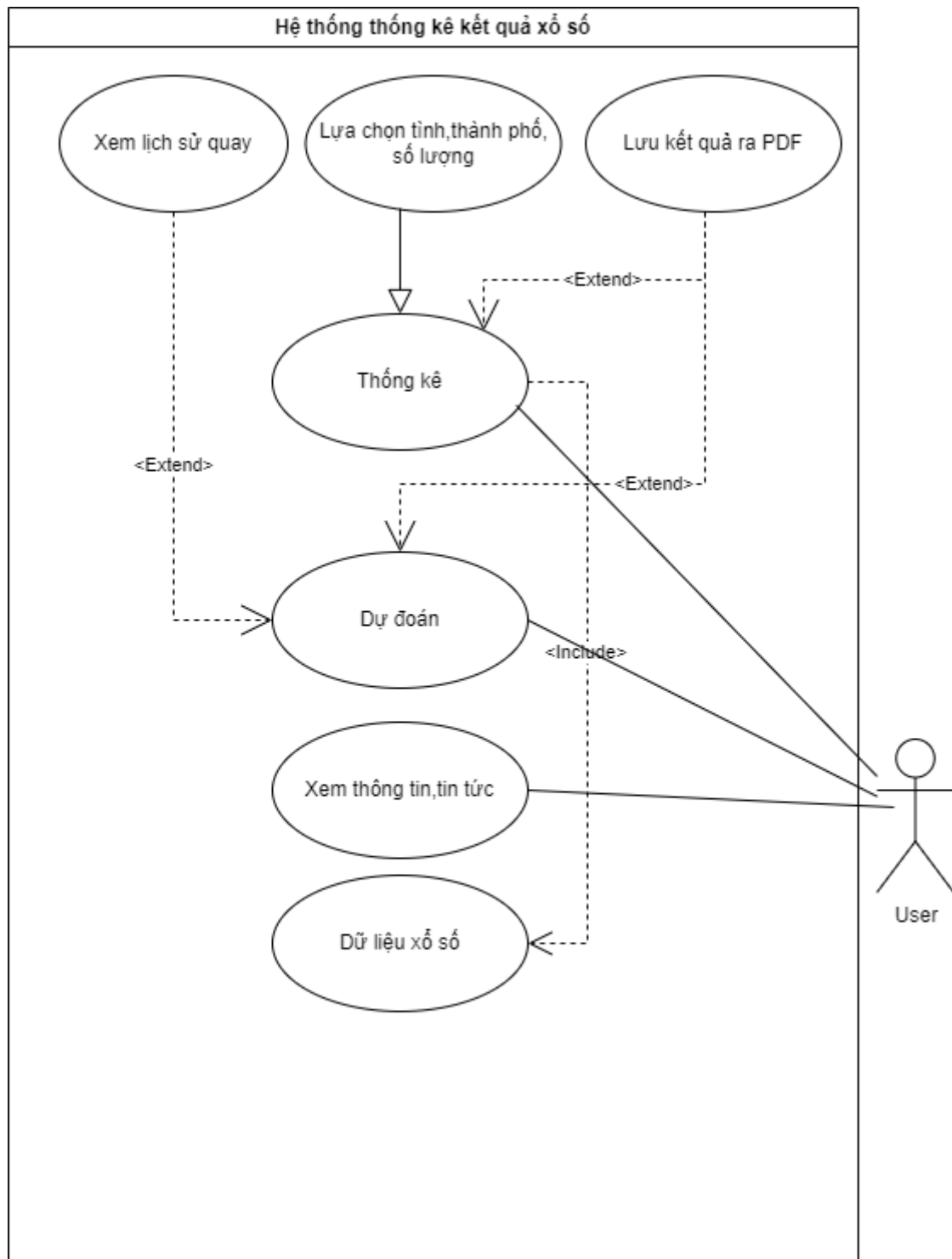
Sơ đồ usecase mô tả các chức năng chính của hệ thống "Website Thống Kê Xổ Số" và mối quan hệ giữa các tác nhân với hệ thống. Cấu trúc này giúp xác định các hành động mà Admin và User có thể thực hiện. Sau đây là các sơ đồ usecase và mô tả chi tiết.

- Sơ đồ usecase tổng quát.



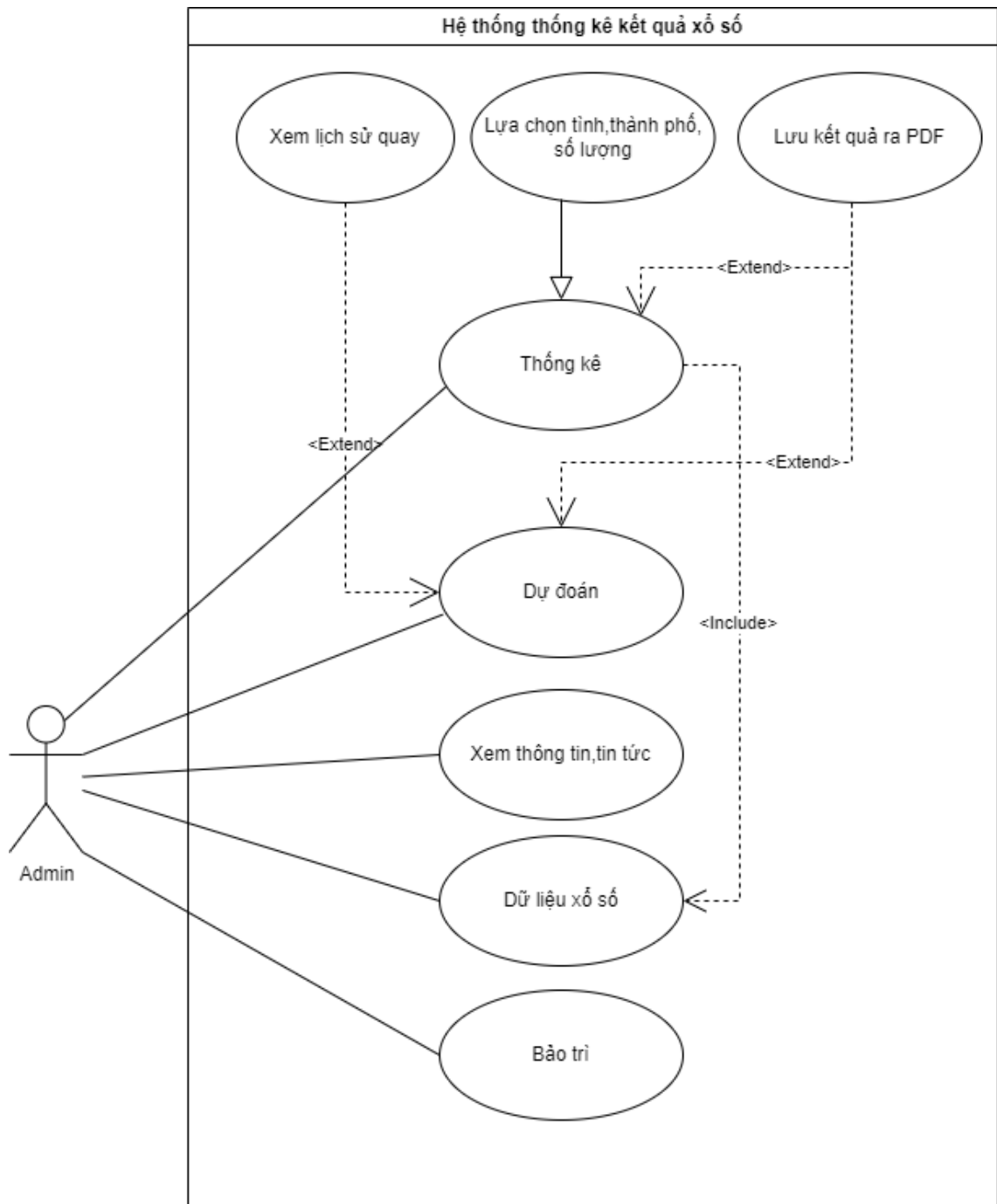
Hình 3.3. Sơ đồ usecase tổng quát của hệ thống

- Sơ đồ usecase cho người dùng.



Hình 3.4. Sơ đồ usecase của tác nhân người dùng.

- Sơ đồ usecase đối với người quản trị.



Hình 3.5. Sơ đồ usecase của tác nhân người quản trị trong hệ th

Dưới đây là bảng mô tả chi tiết usecase của hệ thống:

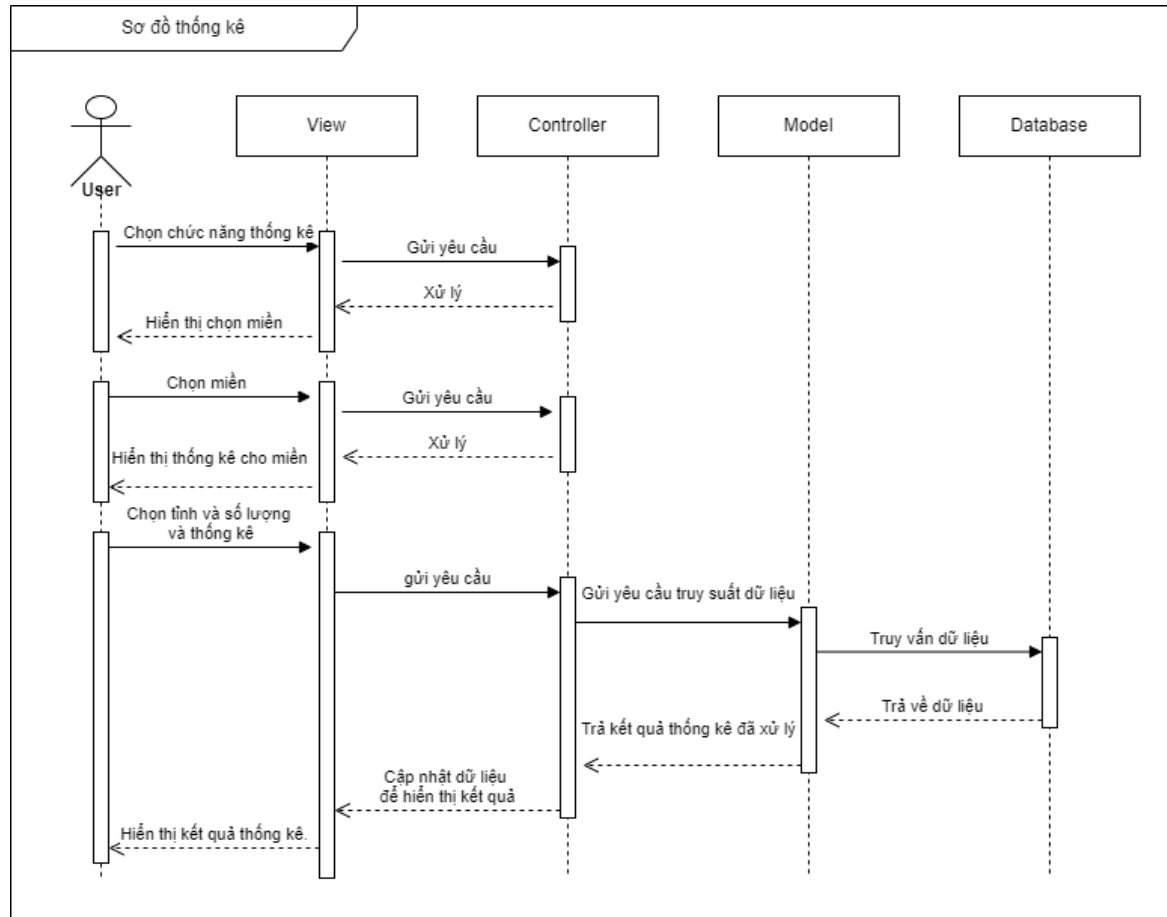
Bảng 3.1. Mô tả Usecase toàn bộ hệ thống

Tên Usecase	Mô tả	Tác nhân	Kịch bản chính	Kịch bản thay thế	Mối quan hệ
Thống kê	Người dùng có thể dùng thống kê kết quả xổ số theo các tiêu chí như tỉnh, thành phố, và số lượng.	Admin, User	1. Người dùng chọn các tiêu chí thống kê. 2. Hệ thống hiển thị kết quả thống kê dựa trên dữ liệu hiện có.	Nếu người dùng không chọn tiêu chí, hệ thống sẽ yêu cầu nhập lại để bổ sung.	
Dự đoán	Hệ thống cung cấp chức năng dự đoán kết quả xổ số dựa trên dữ liệu trước đó.	Admin, User	1. Người dùng chọn loại dự đoán. 2. Hệ thống thực hiện dự đoán và hiển thị kết quả.		
Xem thông tin, tin tức	Cung cấp thông tin và tin tức liên quan đến xổ số cho người dùng.	Admin, User	1. Người dùng chọn mục tin tức. 2. Hệ thống hiển thị các bài viết và tin tức. 3. Nếu người dùng là admin, hệ thống cho phép cập nhật và hiển thị các tin tức mới hoặc chỉnh sửa các bài viết có sẵn.		
Dữ liệu xổ số	Quản lý và truy xuất	Admin	1. Admin cập nhật dữ	Nếu dữ liệu không	<include>

	dữ liệu kết quả xổ số.		liệu hoặc hệ thống tự động cập nhật 2. Dữ liệu được lưu trữ và có thể truy xuất khi cần.	hợp lệ, hệ thống yêu cầu nhập lại hoặc nếu hệ thống có sự cố trong việc tự động thông báo lỗi hoặc không hiện kết quả	
Bảo trì	Admin thực hiện bảo trì hệ thống để đảm bảo hoạt động ổn định.	Admin	1. Admin kiểm tra và cập nhật hệ thống. 2. Hệ thống hoạt động bình thường sau khi bảo trì.	Nếu phát hiện lỗi, hệ thống cần phải sửa chữa.	
Xem lịch sử quay	Người dùng có thể xem lại các kết quả quay xổ số đã thực hiện trước đó.	Admin, User	1. Người dùng chọn mục xem lịch sử quay. Hoặc lướt xuống cuối trang. 2. Hệ thống dẫn hiển thị các kết quả đã quay trước đó.		<extend>
Lưu kết quả ra PDF.	Người dùng có xuất các kết quả ra PDF.	Admin, User	1. Người dùng chọn mục xem lịch sử quay. 2. Hệ thống hiển thị các kết quả đã quay trước đó.		<extend>

3.3.3. Sơ đồ tuần tự “Thống kê cho người dùng”

Sơ đồ tuần tự “Thống kê cho người dùng” mô tả chi tiết quá trình tương tác giữa người dùng và các thành phần của hệ thống để thực hiện một yêu cầu thống kê. Cấu trúc của hệ thống được chia thành các lớp theo mô hình MVC (Model-View-Controller), với từng thành phần hoạt động như sau:



Hình 3.6. Sơ đồ tuần tự chức năng thống kê

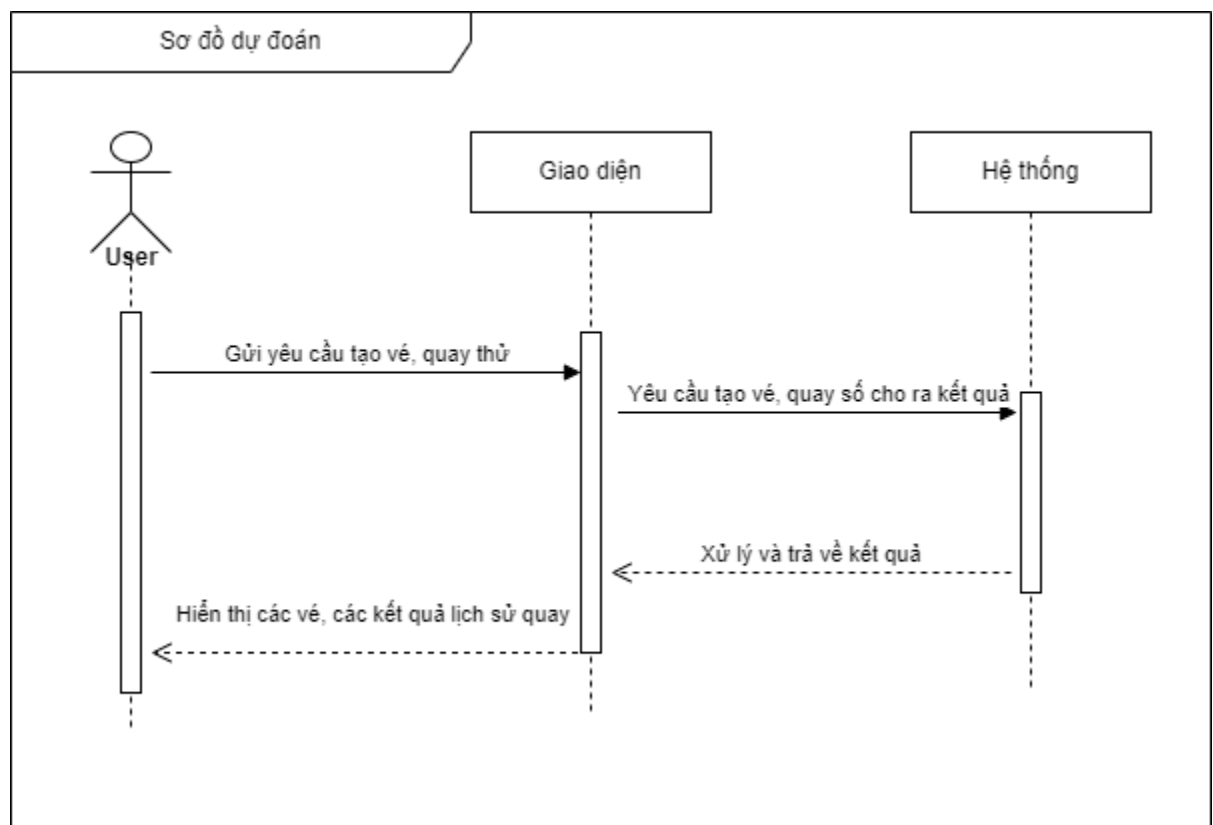
Mô tả quy trình:

1. Người dùng chọn chức năng trên giao diện hệ thống.
2. Gửi yêu cầu từ view.
3. Nhận yêu cầu và xử lý.
4. Hiển thị chọn miền cho người sử dụng.
5. Người dùng chọn miền mong muốn trên màn hình hiển thị
6. Gửi yêu cầu từ view.
7. Nhận yêu cầu và xử lý.

8. Hiện thị thống kê của miền.
9. Người dùng chọn tỉnh và số lượng cần thống kê trên giao diện hệ thống.
10. Gửi yêu cầu từ view.
11. Gửi yêu cầu truy suất dữ liệu từ cơ sở dữ liệu.
12. Truy vấn dữ liệu.
13. Trả về dữ liệu.
14. Trả về kết quả thống kê đã xử lý.
15. Cập nhật dữ liệu để hiển thị kết quả.
16. hiển thị kết quả thống kê.

3.3.4. Sơ đồ tuần tự “Dự đoán”

Sơ đồ tuần tự “Thống kê cho người dùng” mô tả chi tiết quá trình tương tác giữa người dùng và các thành phần của hệ thống để thực hiện một yêu cầu dự đoán bao gồm việc quay thử và rút vé.



Hình 3.7. Sơ đồ tuần tự chức năng dự đoán

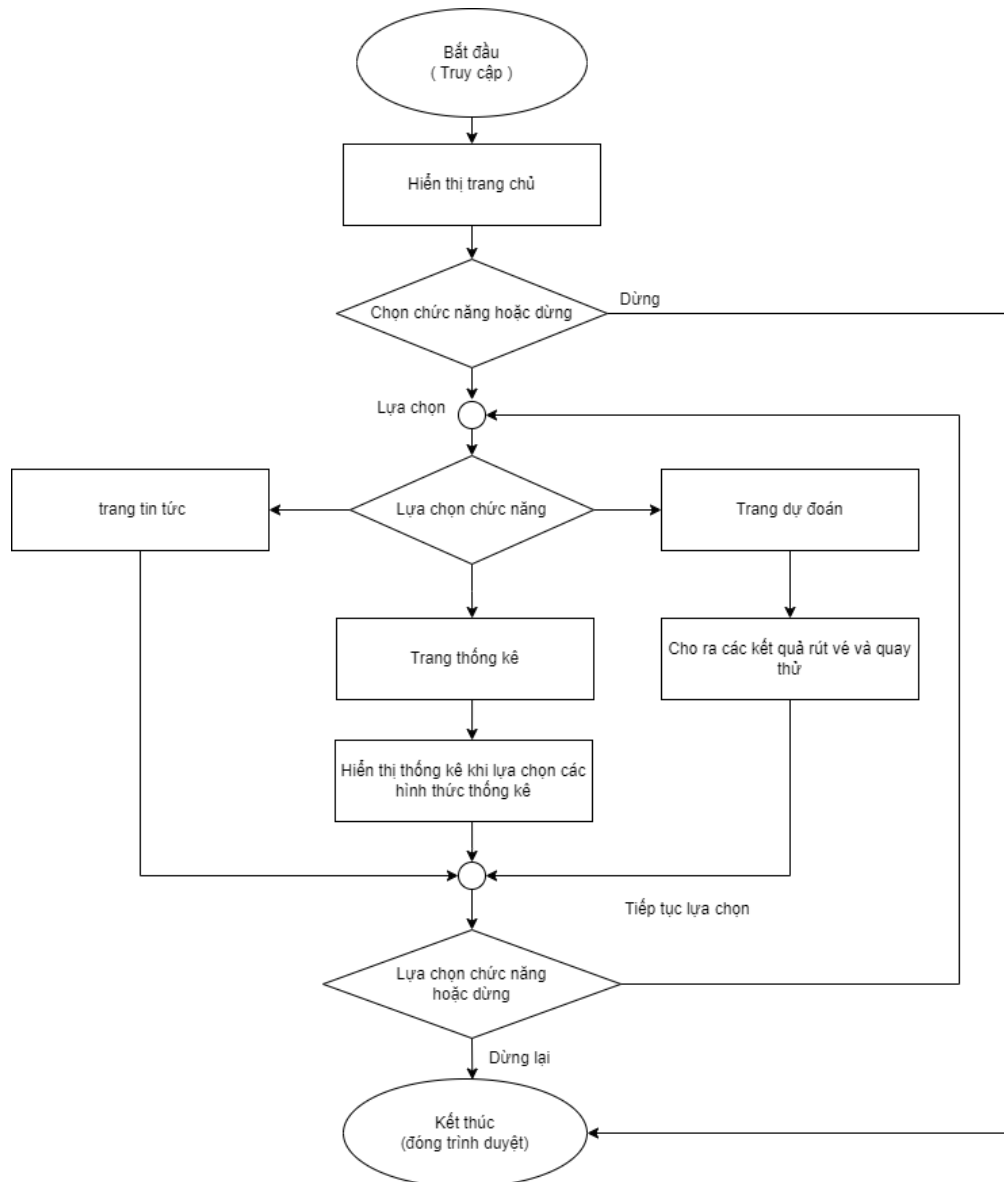
Mô tả quy trình:

1. Người dùng chọn chức năng tạo vé hoặc quay thử hay sử dụng cả hai trên giao diện hệ thống.
2. Gửi yêu cầu đến hệ thống
3. Nhận yêu cầu và xử lý cho ra kết quả
4. Hiển thị các kết quả cho người sử dụng.

3.3.5. Lưu đồ hoạt động của hệ thống

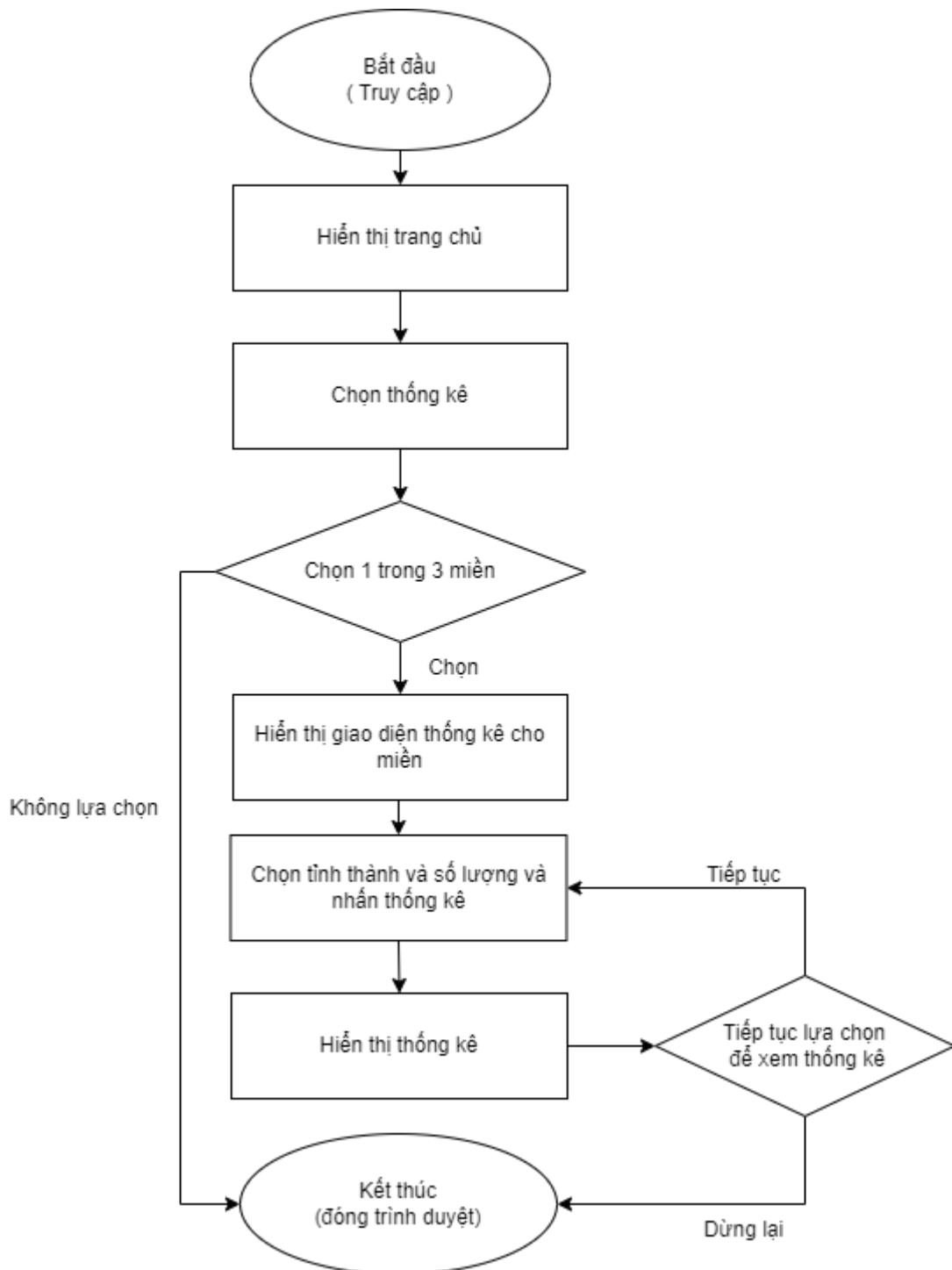
Lưu đồ hoạt động là sơ đồ trực quan hữu ích trong việc mô tả các quá trình, quy trình hoạt động của hệ thống từ đầu đến cuối. Trong trường hợp này, lưu đồ hoạt động của hệ thống "Website Thống kê Xổ số" được thiết kế nhằm minh họa các bước xử lý khi người dùng thực hiện các yêu cầu thống kê, dự đoán hay xem các tin tức trên trang web. Lưu đồ này giúp xác định các trạng thái, hành động và điều kiện của hệ thống, từ đó đảm bảo sự chính xác và hiệu quả của quy trình xử lý và dựa theo sơ đồ tuần tự thì dưới đây là các lưu đồ hoạt động gồm : Lưu đồ tổng quát và các lưu đồ theo chức năng.

- Lưu đồ hoạt động tổng quát của hệ thống.



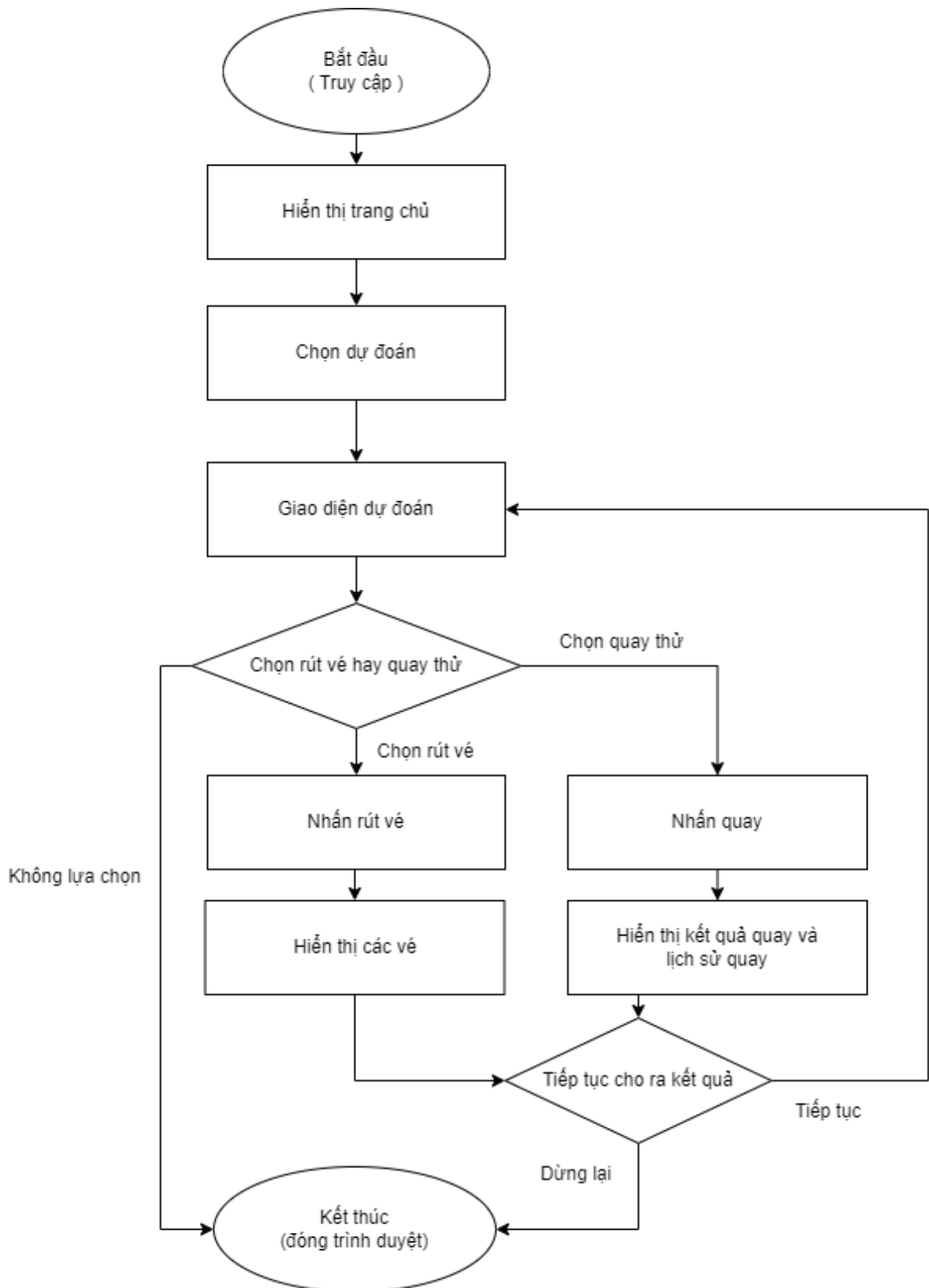
Hình 3.8. Lưu đồ hoạt động của hệ thống.

- Lưu đồ tổng quát của chức năng thống kê



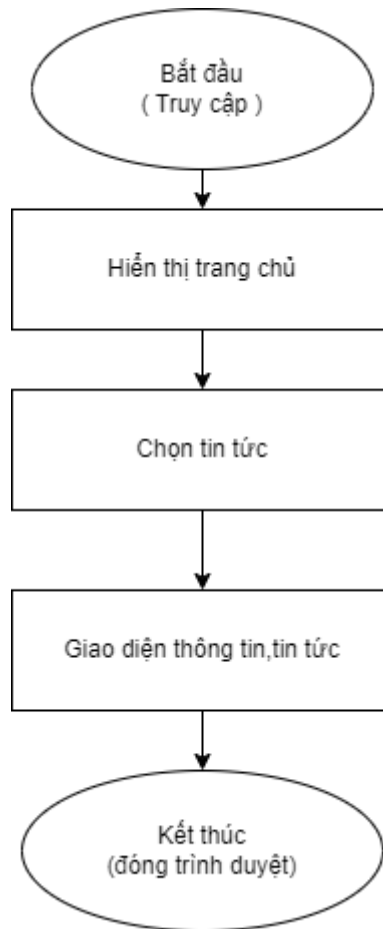
Hình 3.9. Lưu đồ hoạt động cho chức năng thống kê.

- Lưu đồ tổng quát của chức năng dự đoán



Hình 3.10. Lưu đồ hoạt động cho chức năng dự đoán.

- Lưu đồ hoạt động của chức năng xem tin tức



Hình 3.11. Lưu đồ hoạt động chức năng xem tin tức.

3.4. Thiết kế thành phần dữ liệu

3.4.1. Mô hình dữ liệu

Trước tiên xổ số kiến thiết là một hình thức giải trí phổ biến và hợp pháp tại Việt Nam, được tổ chức bởi các công ty xổ số kiến thiết ở từng tỉnh thành.

❖ Cơ chế của xổ số:

- Các vé số thường có dãy số từ 000000 đến 999999 (6 con số đối với Miền Nam và Miền Trung 5 con số đối với Miền Bắc).
- Miền Nam và Miền Trung sẽ có 9 giải từ giải 8 đến giải đặc biệt riêng Miền Bắc không có giải 8. Cả 2 sẽ có các kết quả khác nhau về bao nhiêu số ở trên một giải.
- Quy trình mở thưởng được diễn ra công khai mỗi tỉnh thành (có 41 tỉnh thành) sẽ có thời gian và ngày trực tuyến khác nhau

Tiếp theo là thiết kế mô hình sử dụng một bảng cơ sở dữ liệu áp dụng cho tất cả các tỉnh (41 tỉnh) dưới đây là bảng cơ sở dữ liệu được sử dụng chung.

Bảng 3.2. Bảng dữ liệu xổ số

STT	Tên trường	Kiểu dữ liệu	Khóa chính	Mô tả
1	id	Int Auto-Increment	X	Khóa chính (Primary Key), tự động tăng
2	prize_8	Varchar (255)		Số liệu giải 8, mặc định là chuỗi rỗng
3	prize_7	Varchar (255)		Số liệu giải 7, mặc định là chuỗi rỗng
4	prize_6	Varchar (255)		Số liệu giải 6, mặc định là chuỗi rỗng
5	prize_5	Varchar (255)		Số liệu giải 5, mặc định là chuỗi rỗng
6	prize_4	Varchar (255)		Số liệu giải 4, mặc định là chuỗi rỗng
7	prize_3	Varchar (255)		Số liệu giải 3, mặc định là chuỗi rỗng
8	prize_2	Varchar (255)		Số liệu giải 2, mặc định là chuỗi rỗng
9	prize_1	Varchar (255)		Số liệu giải nhất, mặc định là chuỗi rỗng
10	prize_db	Varchar (255)		Số liệu giải đặc biệt, mặc định là chuỗi rỗng

Trong Django ở file models.py đầu tiên tạo mô hình lớp cha sao đó tạo mô hình con tương ứng cho từng tỉnh thành.

Bảng 3.3. Tạo mô hình cho cơ sở dữ liệu

```
# Mô hình cha
class BaseLotteryNumber(models.Model):
    id = models.AutoField(primary_key=True)
    prize_8 = models.CharField(max_length=255, blank=True)
    prize_7 = models.CharField(max_length=255, blank=True)
    prize_6 = models.CharField(max_length=255, blank=True)
    prize_5 = models.CharField(max_length=255, blank=True)
    prize_4 = models.CharField(max_length=255, blank=True)
    prize_3 = models.CharField(max_length=255, blank=True)
    prize_2 = models.CharField(max_length=255, blank=True)
    prize_1 = models.CharField(max_length=255, blank=True)
    prize_db = models.CharField(max_length=255, blank=True)

    class Meta:
        abstract = True

# Mô hình con cho từng tỉnh
class AnGiangLotteryNumber(BaseLotteryNumber):
    class Meta:
        db_table = 'angiang'

#Thêm các tỉnh khác
```

Trong đó:

- Lớp cha BaseLotteryNumber chứa các trường chung và kiểu dữ liệu cho các bảng dữ liệu liên quan đến kết quả xổ số.

- Lớp con TestLotteryNumber kế thừa từ lớp cha, và thêm thông tin chi tiết về bảng dữ liệu mà nó đại diện.
- class Meta: abstract = True: Câu lệnh này đánh dấu lớp BaseLotteryNumber là mô hình trừu tượng. Điều này có nghĩa là Django sẽ không tạo ra một bảng cơ sở dữ liệu cho mô hình này. Thay vào đó, các mô hình con kế thừa từ nó sẽ được sử dụng để tạo bảng. Lớp trừu tượng chỉ được sử dụng như một khuôn mẫu chung để chia sẻ các trường dữ liệu cho các lớp con.
- class Meta: db_table = 'test': Câu lệnh này định nghĩa tên bảng cơ sở dữ liệu sẽ được tạo ra cho mô hình AnGiangLotteryNumber. Trong trường hợp này, bảng sẽ có tên các tỉnh thành. Bằng cách này, bạn có thể tạo ra nhiều lớp con từ BaseLotteryNumber và mỗi lớp con có thể có một bảng cơ sở dữ liệu riêng với tên khác nhau.

3.4.2. Thu thập dữ liệu

Trong việc thu thập dữ liệu web, Playwright là công cụ mạnh mẽ được sử dụng để tự động hóa trình duyệt, giúp thu thập dữ liệu một cách chính xác và hiệu quả. Đầu tiên là chọn trang web <https://xoso.com.vn> để thu thập.

XSAG - Kết quả Xổ số An Giang - SXAG hôm nay			
XSMN / XSMN Thứ 5 / XSMN 21/11/2024			
XSAG Thứ 5 / XSAG 21/11/2024			
8	10		
7	264		
6	3483	4264	5469
5	0163		
4	40963	14081	69414
	77218	17849	12793
		20331	
3	04373	53832	
2	89119		
1	85274		
ĐB	524341		

Hình 3.12. Giao diện thống kê của trang web xoso.com.vn

Quy trình thực hiện:

Bảng 3.4. Mô tả quy trình thu thập dữ liệu.

Bước	Mô tả
Bước 1: Khởi tạo Django	Mã khởi tạo môi trường Django, thiết lập các cài đặt và kết nối với cơ sở dữ liệu thông qua <code>os.environ</code> và <code>django.setup()</code> .
Bước 2: Truy cập trang Web	Sử dụng Playwright để mở trang <code>Xoso.com.vn</code> và truy cập trang kết quả xổ số của tỉnh An Giang (<code>xsag-pl.html</code>) hoặc các tỉnh khác.
Bước 3: Xóa Dữ Liệu Cũ	Trước khi thu thập dữ liệu mới, tất cả bản ghi trong bảng <code>angiang</code> sẽ được xóa để tránh trùng lặp. Thực hiện bằng câu lệnh SQL <code>DELETE FROM angiang</code> và reset ID trường bằng <code>ALTER TABLE angiang AUTO_INCREMENT = 1;</code> .
Bước 4: Nhấn Nút "Xem Thêm"	Mã tự động nhấn nút "Xem thêm" để tải thêm kết quả xổ số cho đến khi tìm thấy ngày cần tìm
Bước 5: Thu Thập Kết Quả	Mã sử dụng các selector của Playwright để trích xuất kết quả từ các giải thưởng: Giải 8, Giải 7, Giải 6, Giải 5, Giải 4, Giải 3, Giải 2, Giải 1 và Giải Đặc Biệt. Mỗi giải sẽ được lưu vào danh sách tương ứng.
Bước 6: Lưu Dữ Liệu Vào CSDL	Sau khi thu thập xong kết quả xổ số, mã sẽ lưu dữ liệu vào bảng <code>angiang</code> trong cơ sở dữ liệu MySQL thông qua phương thức <code>save angiang lottery_number()</code> .
Bước 7: Đóng Trình Duyệt	Sau khi hoàn thành thu thập và lưu trữ, trình duyệt được đóng lại bằng <code>await browser.close()</code> .

Các thành phần:

Bảng 3.5. Các thành phần có trong các tình để thu thập dữ liệu

Thành phần	Giải thích
@sync_to_async	Trang bị cho các phương thức Django để có thể chạy trong môi trường bất đồng bộ (async), giúp kết hợp Playwright với Django một cách hiệu quả.
save_angiang_lottery_number()	Phương thức này lưu dữ liệu xổ số (giải thưởng từ giải 8 đến giải Đặc Biệt) vào cơ sở dữ liệu MySQL, sử dụng mô hình AnGiangLotteryNumber. Các tình khác tương tự
delete_all_angiang_lottery_numbers()	Phương thức này xóa tất cả các bản ghi hiện có trong bảng angiang và đặt lại ID tự động (AUTO_INCREMENT) về 1, đảm bảo dữ liệu mới được lưu vào từ đầu.
async with async_playwright()	Khởi tạo Playwright trong chế độ bất đồng bộ (async) để tự động hóa việc truy cập và thao tác với trình duyệt.
await page.goto('https://xoso.com.vn/xo-so-can-tho/xsag-p1.html')	Mã truy cập trang web kết quả xổ số An Giang từ địa chỉ URL đã cho.
while not target_date_found: hoặc for i in range(): await page.click('div.load-more a#loadmore')	Mã vòng lặp để nhấn nút "Xem thêm" và tải thêm kết quả xổ số cho đến khi tìm thấy kết quả với ngày 19/09/2024 hoặc nhấn xem thêm bao nhiêu lần.
numbers_8 = await page.query_selector_all('tr > td.name-prize:has-text("8") + td span.number_DB')	Trích xuất các số trúng thưởng của giải 8, sử dụng selector Playwright để tìm các phần tử có chứa "8".

<code>prize_8_list.append(number_text)</code>	Lưu các số trúng giải 8 vào danh sách <code>prize_8_list</code> . Tương tự với các giải thưởng khác (Giải 7, Giải 6, ...).
<code>max_length = max(len(prize_8_list), len(prize_7_list), ...)</code>	Xác định chiều dài tối đa của các danh sách giải thưởng để vòng lặp sau có thể lặp qua tất cả các giải thưởng (dù số lượng có thể khác nhau).
<code>await save_angiang_lottery_number(...)</code>	Lưu mỗi dãy số của từng giải vào cơ sở dữ liệu với phương thức đã định nghĩa ở trên.
<code>await browser.close()</code>	Đóng trình duyệt sau khi thu thập dữ liệu xong.

Các Selector được sử dụng cho các tỉnh thuộc miền Nam và Miền Trung

Bảng 3.6. Các phần tử được chọn cho Miền Nam và Miền Trung.

Giải	Selector
8	<code>'tr > td.name-prize:has-text("8") + td span.number_DB'</code>
7	<code>'tr > td.name-prize:has-text("7") + td span.xs_prize1'</code>
6	<code>'tr > td.name-prize:has-text("6") + td span.xs_prize3'</code>
5	<code>'tr > td.name-prize:has-text("5") + td span.xs_prize1'</code>
4	<code>'tr > td.name-prize:has-text("4") + td span.xs_prize3, tr > td.name-prize:has-text("4") + td span.xs_prize1'</code>
3	<code>'tr > td.name-prize:has-text("3") + td span.xs_prize2'</code>
2	<code>'tr > td.name-prize:has-text("2") + td span.xs_prize1'</code>
1	<code>'tr > td.name-prize:has-text("1") + td span.xs_prize1'</code>
Đặc Biệt	<code>'tr > td.name-prize:has-text("ĐB") + td span.number_DB'</code>

Các Selector được sử dụng cho các tỉnh thuộc miền Bắc

Bảng 3.7. Các phần tử được chọn cho Miền Bắc.

Giải	Selector
7	'tr > td:has-text("7") + td span.prize7'
6	'tr > td:has-text("6") + td span.prize6'
5	'tr > td:has-text("5") + td span.prize5'
4	'tr > td:has-text("4") + td span.prize4'
3	'tr > td:has-text("3") + td span.prize3'
2	'tr > td:has-text("2") + td span.prize2'
1	'tr > td:has-text("1") + td span.prize1'
Đặc Biệt	'tr > td:has-text("ĐB") + td span.special-prize'

3.4.3. Tự động hóa trong việc thu thập dữ liệu

Dựa theo cơ chế mỗi lần xổ số các file thu thập dữ liệu sẽ được thu thập khoảng 10 phút sau khi kết thúc xổ số trong ngày và được thực hiện trong 1 file schedule.py khi khởi động server thì đồng thời chạy file này để hoạt động liên tục.

Trong python có một thư viện có sẵn là thư viện schedule giúp lập lịch trình thực hiện các tác vụ tự động trong Python. Thư viện này dễ sử dụng và phù hợp cho các tác vụ định kỳ, như chạy các tập lệnh theo ngày, giờ.

Bảng 3.8. Bảng xử lý tác vụ hẹn giờ thu thập dữ liệu.

<pre>import schedule import time import subprocess</pre>
<pre>def run_script(file_path): try: subprocess.run(["python", file_path], check=True) except subprocess.CalledProcessError as e: print(f"Lỗi khi chạy {file_path}: {e}")</pre>


```
def schedule_jobs():
    schedule.every().thursday.at("18:00").do(run_script,"D:/TAI_LIEU_HOC_TAP/
    MON_HOC/THONGKEXOSO/TKXS/angiang.py")
    schedule_jobs()
    while True:
        schedule.run_pending()
        time.sleep(1)
```

Trong đó:

- `schedule`: Thư viện để định nghĩa các tác vụ theo lịch.
- `time`: Sử dụng để tạo vòng lặp kiểm tra và chạy các lịch trình.
- `subprocess`: Thực hiện các tập lệnh Python khác bằng cách gọi python file_path.
- `file_path`: Đường dẫn của file Python cần chạy.
- `subprocess.run`: Chạy file Python được chỉ định.
- `check=True`: Nếu file xảy ra lỗi khi chạy, sẽ báo lỗi.
- Xử lý lỗi: In ra thông báo nếu có lỗi khi chạy file.
- `schedule.every().thursday.at("18:00")`: Định nghĩa lịch chạy vào lúc 18:00 mỗi thứ 5.
- `.do(run_script, "path_to_script")`: Thực hiện hàm `run_script` với đường dẫn file chỉ định.
- Gọi hàm `schedule_jobs` để thêm các lịch trình vào hệ thống
- `schedule.run_pending()`: Kiểm tra nếu có tác vụ nào đến thời gian cần thực hiện, nó sẽ chạy.
- `time.sleep(1)`: Tạm dừng 1 giây mỗi lần lặp để tránh sử dụng quá nhiều tài nguyên CPU.

3.4.4. Các phương pháp thống kê

Trong quá trình thống kê và phân tích kết quả xổ số, việc xử lý và phân nhóm dữ liệu đóng vai trò quan trọng trong việc tìm ra các mẫu và xu hướng. Các phương

pháp thống kê áp dụng cho dữ liệu xổ số có thể bao gồm các bước xử lý chữ số đầu tiên, các chữ số cuối, và việc tìm ra các nhóm phổ biến nhất hoặc ít phổ biến nhất.

+ Đầu tiên xử lý con số đầu tiên.

Bảng 3.9. Xử lý con số đầu tiên.

```
grouped_by_first_digit_5 = {}
for prize_5, count in prize_5_count.items():
    first_digit = str(prize_5)[0] # Lấy chữ số đầu tiên
    grouped_by_first_digit_5[first_digit] =
grouped_by_first_digit_5.get(first_digit, 0) + c
```

Trong đó

- Khởi tạo: Tạo từ điển `grouped_by_first_digit_5` để lưu kết quả. (số 5 tương ứng giải 5)
- Lấy chữ số đầu tiên: Dùng `str(prize_5)[0]` để lấy chữ số đầu tiên của `prize_5`.
- Cộng dồn: Dùng `get()` để cộng dồn `count` vào nhóm theo chữ số đầu tiên.

+ Lấy con số cuối.

Bảng 3.10. Xử lý con số cuối.

```
grouped_by_units_5 = {}
for prize_5, count in prize_5_count.items():
    unit = prize_5[-1] # Lấy chữ số đơn vị (cuối)
    grouped_by_units_5[unit] = grouped_by_units_5.get(unit, 0) + count
grouped_by_units_5 = dict(sorted(grouped_by_units_5.items(), key=lambda item:
int(item[0])))
```

Trong đó:

- Khởi tạo: Tạo từ điển `grouped_by_units_5` để lưu kết quả nhóm theo chữ số đơn vị (cuối cùng) của `prize_5`.
- Lấy chữ số đơn vị: Dùng `prize_5[-1]` để lấy chữ số cuối cùng của `prize_5`.
- Cộng dồn: Dùng `get()` để cộng dồn `count` vào nhóm theo chữ số đơn vị.

- sorted(): Sắp xếp các phần tử của từ điển theo chữ số đơn vị.
 - key=lambda item: int(item[0]): Sắp xếp theo giá trị số thực sự của chữ số đơn vị, không phải theo thứ tự chữ cái.
 - dict(): Chuyển kết quả đã sắp xếp thành từ điển
- + Lấy 2 số cuối

Bảng 3.11. Xử lý 2 số cuối.

```
grouped_by_last_two_digits_5 = {}
for prize_5, count in prize_5_count.items():
    last_two_digits = prize_5[-2:] # Lấy 2 chữ số cuối
    grouped_by_last_two_digits_5[last_two_digits] =
grouped_by_last_two_digits_5.get(last_two_digits, 0) + count
```

Trong đó:

- Khởi tạo: Tạo từ điển grouped_by_last_two_digits_5 để lưu kết quả nhóm theo 2 chữ số cuối của prize_5.
 - Lấy 2 chữ số cuối: Dùng prize_5[-2:] để lấy 2 chữ số cuối của prize_5.
 - Cộng dồn: Dùng get() để cộng dồn count vào nhóm theo 2 chữ số cuối.
- + Lấy 3 số cuối: Tương tự lấy 2 số chỉ thay đổi

Bảng 3.12. Xử lý 3 số cuối.

```
grouped_by_last_three_digits_5 = {}
for prize_5, count in prize_5_count.items():
    last_three_digits = prize_5[-3:] # Lấy 3 chữ số cuối
    grouped_by_last_three_digits_5[last_three_digits] =
grouped_by_last_three_digits_5.get(last_three_digits, 0) + count
```

Trong đó:

- Khởi tạo từ điển grouped_by_last_three_digits_5 để lưu kết quả nhóm theo 3 chữ số cuối của prize_5.

- Duyệt qua từng phần tử trong từ điển prize_5_count.
 - Lấy 3 chữ số cuối từ prize_5 bằng prize_5[-3:].
 - Cộng dồn count vào nhóm tương ứng trong grouped_by_last_three_digits_5 với việc sử dụng get() để kiểm tra và cập nhật giá trị.
- + Lấy 2 giá trị đầu

Bảng 3.13. Xử lý 2 con số đầu.

```
grouped_by_first_two_digits_4 = {}
for prize_4, count in prize_4_count.items():
    first_two_digits = prize_4[:2] # Lấy 2 chữ số đầu tiên
    grouped_by_first_two_digits_4[first_two_digits] =
    grouped_by_first_two_digits_4.get(first_two_digits, 0) + count
```

Trong đó:

- Khởi tạo từ điển grouped_by_first_two_digits_4 để lưu kết quả nhóm theo 2 chữ số đầu tiên của prize_4.
 - Duyệt qua từng phần tử trong từ điển prize_4_count.
 - Lấy 2 chữ số đầu tiên của prize_4 bằng cách sử dụng prize_4[:2].
 - Cộng dồn giá trị count vào nhóm tương ứng trong grouped_by_first_two_digits_4 với việc dùng get() để kiểm tra và cập nhật giá trị.
- + Top 10 các giá trị của những giá trị đầu áp dụng cho cả 2 số cuối 3 số cuối

Bảng 3.14. Xử lý chọn ra 10 giá trị nhiều nhất và ít nhất.

```
top_10_most_common_f2digits_4 =
sorted(grouped_by_first_two_digits_4.items(), key=lambda item: item[1],
reverse=True)[:10]

top_10_least_common_f2digits_4 =
sorted(grouped_by_first_two_digits_4.items(), key=lambda item: item[1]):[:10]
```

Trong đó:

- Liệt kê ra 10 số liệu nhiều nhất: Sắp xếp theo tần suất giảm dần bằng `sorted(..., reverse=True)` và lấy 10 nhóm đầu, lưu vào `top_10_most_common_f2digits_4`.
 - Liệt kê ra 10 số liệu ít nhất: Sắp xếp theo tần suất tăng dần bằng `sorted(...)` và lấy 10 nhóm đầu, lưu vào `top_10_least_common_f2digits_4`.
- + Lấy 2 giá trị sau 2 số đầu hoặc các vị trí

Bảng 3.15. Xử lý 2 con số ở các vị trí khác nhau.

```
grouped_by_second_two_digits_db = {}
for prize_db, count in prize_db_count.items():
    second_two_digits = prize_db[2:4] # Lấy 2 chữ số tiếp theo sau 2 chữ số đầu
    grouped_by_second_two_digits_db[second_two_digits] =
grouped_by_second_two_digits_db.get(second_two_digits, 0) + count
```

Trong đó:

- Khởi tạo từ điển `grouped_by_second_two_digits_db` để lưu kết quả nhóm theo 2 chữ số thứ hai (vị trí 3 và 4) của `prize_db`.
- Duyệt qua từng phần tử trong từ điển `prize_db_count`.
- Lấy 2 chữ số thứ hai của `prize_db` bằng cách sử dụng `prize_db[2:4]`.
- Cộng dồn giá trị `count` vào nhóm tương ứng trong `grouped_by_second_two_digits_db` bằng cách sử dụng `get()`.

Ngoài ra có thể sử dụng các vị trí khác như

- `prize_db[3:5]` : Lấy ký tự tại vị trí thứ 4 và 5. nếu `prize_db = "123456"`, kết quả sẽ là `'45'`.
- `prize_db[:3]` : Lấy từ đầu chuỗi đến trước vị trí thứ 3. Nếu `prize_db = "123456"`, kết quả sẽ là `'123'`. Hoặc ngược lại với `prize_db[3:]` tức là lấy từ vị trí thứ 4 đến hết chuỗi.
- `prize_db[::2]` #Lấy từng ký tự cách 2 vị trí. Nếu `prize_db = "123456"`, kết quả sẽ là `'135'`.

3.4.5. Phương pháp tạo vé và quay thử

Mục tiêu của phương pháp tạo vé và quay thử là cung cấp một cách thức ngẫu nhiên và mô phỏng quá trình rút vé xổ số, giúp người dùng có thể trải nghiệm việc quay số mà không cần tham gia vào quá trình rút vé thực tế. Phương pháp này được thiết kế để tạo ra các kết quả ngẫu nhiên cho các dãy số xổ số, phục vụ cho việc thống kê, phân tích, hoặc chỉ đơn giản là cho người dùng thử vận may. Dưới đây là các thành phần được sử dụng.

Bảng 3.16. Các thành phần xử lý tạo vé.

Thành phần	Ý nghĩa
<code>let isSpinning = false;</code>	Biến cờ dùng để kiểm tra trạng thái quay số, mặc định là chưa quay (false).
<code>let spinTimeout;</code>	Biến lưu trữ timeout nếu cần thêm hiệu ứng thời gian hoặc dừng quay số sau khoảng thời gian.
<code>const results = [];</code>	Mảng lưu kết quả các lần quay trước đó để sử dụng cho lịch sử hoặc các mục đích khác.
<code>document.getElementById("generateButton")</code>	Lấy phần tử HTML có id là "generateButton" (nút "Rút vé").
<code>.addEventListener("click", function() {</code>	Gắn sự kiện click cho nút, khi bấm sẽ chạy hàm bên trong.
<code>function generateRandomNumber()</code>	Hàm tạo số ngẫu nhiên gồm 6 chữ số từ 100000 đến 999999.
<code>Math.floor(100000 + Math.random() * 900000)</code>	Tính toán giá trị ngẫu nhiên: <code>Math.random()</code> tạo số thập phân từ 0 đến 1, nhân với 900000 và cộng 100000.

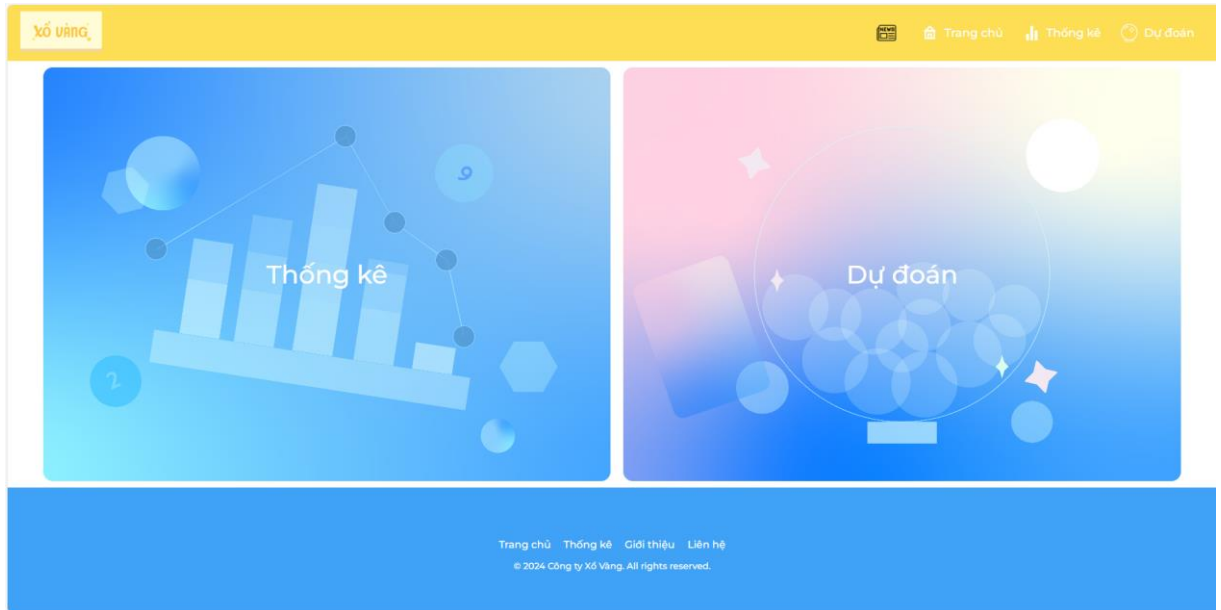
<code>let ticket1 = generateRandomNumber();</code>	Tạo số ngẫu nhiên đầu tiên và gán vào biến ticket1. Tương tự với vé thứ 2 và vé thứ 3
<code>document.getElementById("number1").innerText = ticket1;</code>	Gán số ticket1 vào phần tử HTML có id là "number1". Tương tự với vé thứ 2 và vé thứ 3

Bảng 3.17. Các thành phần xử lý quay số.

Thành phần	Ý nghĩa
<code>document.querySelector(".nutdudoan[onclick*='startContinuousSpin()']").addEventListener("click", startContinuousSpin);</code>	Lắng nghe sự kiện click trên nút "Quay" để gọi hàm startContinuousSpin().
Hàm startContinuousSpin()	Khởi động quá trình quay liên tục, kiểm tra nếu đang quay thì không thực hiện quay mới.
<code>spinCircles();</code>	Gọi hàm spinCircles() để bắt đầu quay các vòng.
<code>setTimeout(() => { stopContinuousSpin(); }, 2000);</code>	Dừng quay tự động sau 2 giây.
Hàm spinCircles()	Quay các vòng số, tạo số ngẫu nhiên và màu sắc cho từng vòng. Lặp lại quay mỗi 100ms cho đến khi dừng.
Hàm getCurrentSpinResults()	Thu thập kết quả hiện tại từ các vòng quay và trả về mảng kết quả.
Hàm displayResults()	Hiển thị kết quả quay trên giao diện, định dạng theo các giải (G8, G7, G6, v.v.) và thêm vào phần tử container kết quả.

3.5. Thiết kế giao diện

3.5.1. Giao diện trang chủ



Hình 3.13. Giao diện trang chủ Xổ Vàng.

- Với giao diện trang chủ sẽ có các chức năng mà mục đích sau:

- Thống kê: hiển thị tổng số lượt xổ số, tỷ lệ trúng thưởng và các thông tin khác về xổ số.
- Quay thử - Dự đoán: cho phép người dùng quay thử xổ số hoặc dự đoán các con số xổ số.
- Tin tức (biểu tượng báo màu đen): cho phép người dùng xem tin tức ,thông tin về các bảng cập nhật mới nhất.

3.5.2. Giao diện chọn miền

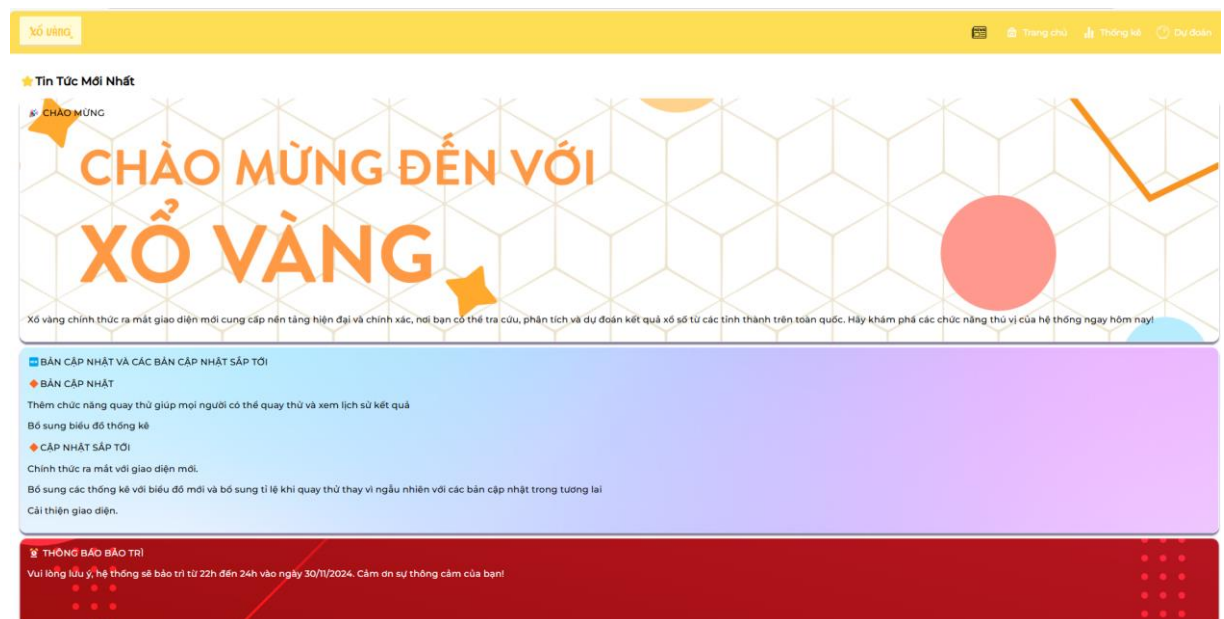
- Với giao diện chọn miền sẽ cho ra các lựa chọn để xem thống kê



Hình 3.14. Giao diện chọn miền của Xổ Vàng

3.5.3. Giao diện thông báo, thông tin

Giao diện thông báo và thông tin là phần không thể thiếu trong thiết kế của hệ thống "Website Thống kê Xổ số", giúp người dùng nhận được các cập nhật và thông báo quan trọng từ hệ thống. Được thiết kế để đảm bảo sự rõ ràng và dễ hiểu, giao diện này cung cấp thông tin cần thiết giúp người dùng có cái nhìn tổng quan về các sự kiện và tình trạng của hệ thống.



Hình 3.15. Giao diện xem tin tức trong Xổ Vàng

3.5.4. Giao diện thống kê

Thống kê Miền Nam & Miền Trung

Chọn tỉnh: Chọn số lượng: **Thống kê**

Thống kê Miền Bắc

Chú ý khi thống kê

★ **XEM GIẢI:**

Giải 7 **Giải 6** **Giải 5** **Giải 4** **Giải 3** **Giải 2** **Giải 1** **Giải DB** **Chụp màn hình và lưu PDF**

★ **thống kê giải 8:**

Không có dữ liệu thống kê.

Các số không xuất hiện:

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59

3 Con số xuất hiện nhiều nhất:

Con số	Số lần xuất hiện
00	0 lần
01	0 lần
02	0 lần

3 Con số xuất hiện ít nhất:

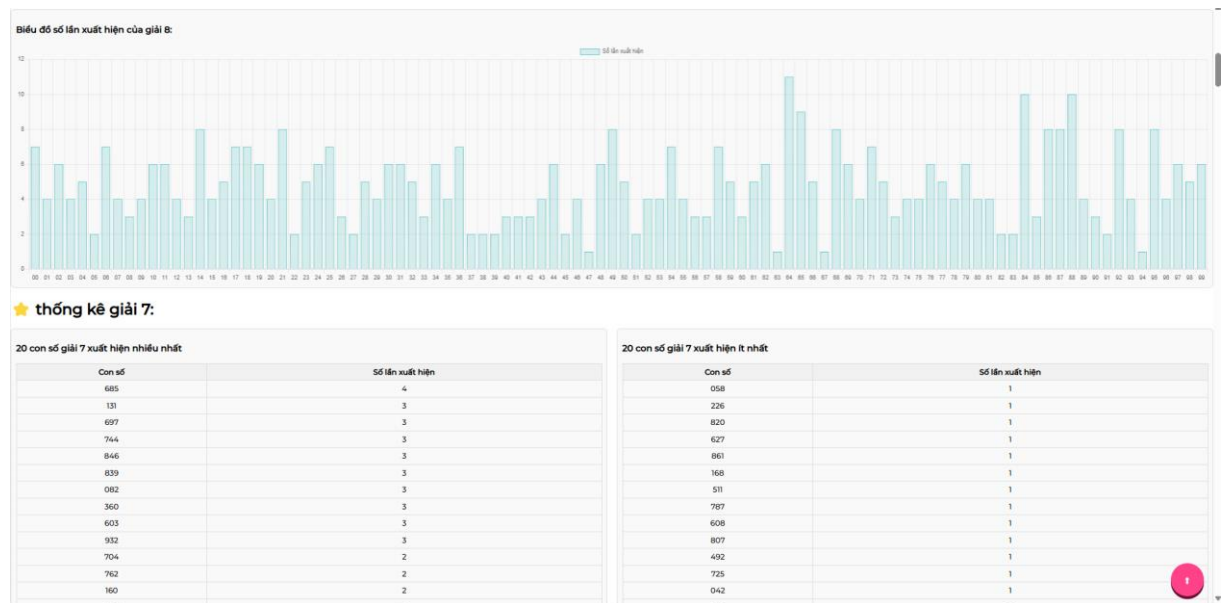
Con số	Số lần xuất hiện
97	0 lần

Hình 3.16. Giao diện thống kê

Với giao diện trang chủ sẽ có các chức năng mà mục đích sau:

- Chọn tỉnh: Đây là nút để chọn tỉnh cần thống kê
- Số lượng thống kê: Là nút chọn số lượng con số cần thống kê
- Nút thông kê các miền khác: Đây là nút lổi tắt để chuyển trang của miền khác
- Nút xuất kết quả ra PDF: xuất ra kết quả dưới dạng PDF

3.5.5. Giao diện hiển thị kết quả thống kê



Hình 3.17. Giao diện kết quả thống kê kèm biểu đồ.

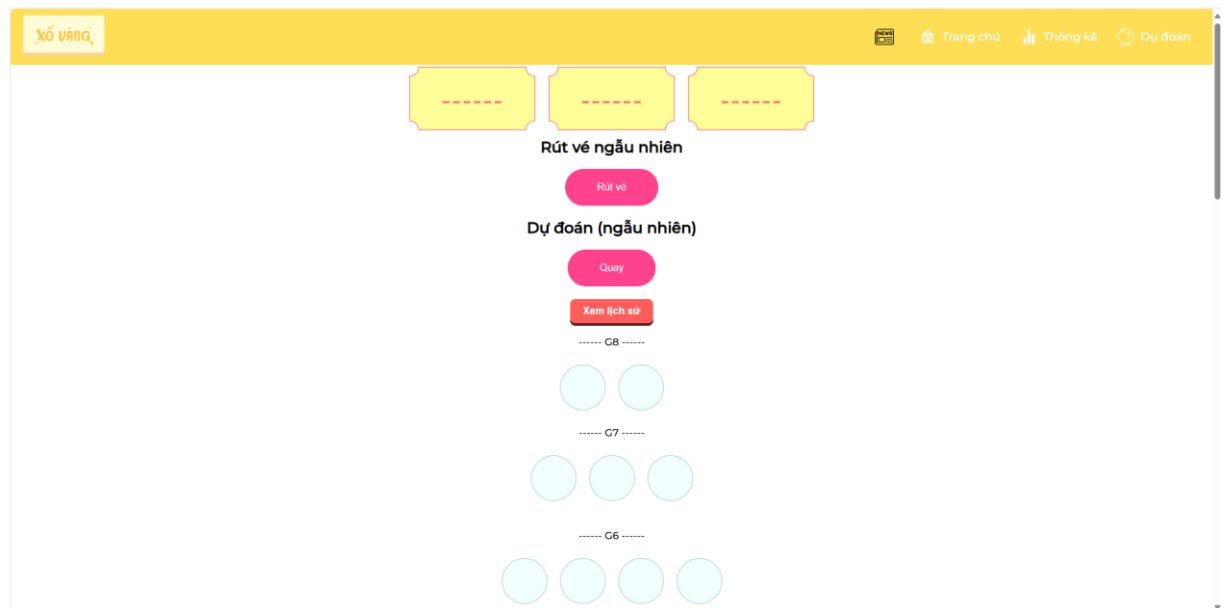
Với giao diện trên sẽ cho ra kết quả các loại thống kê gồm:

- thời gian cụ thể.
- Thống kê chu kỳ: Cho biết các số nào thường xuất hiện theo chu kỳ đều đặn.
- Thống kê Thống kê tần suất: Hiển thị số lần xuất hiện của các con số trong một khoảng
- số lâu chưa ra: Xác định những con số chưa xuất hiện trong thời gian dài, từ đó gợi ý khả năng xuất hiện trong tương lai.
- Hiển thị thêm thống kê theo các thứ tự các số

3.5.6. Giao diện dự đoán

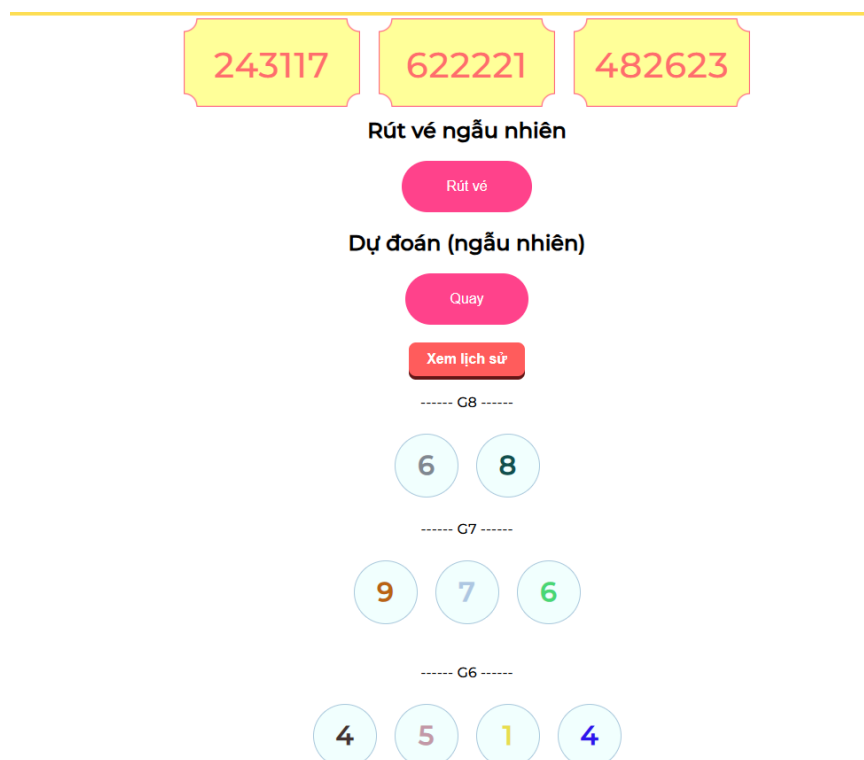
Với giao diện dự đoán sẽ bao gồm:

- Nút rút vé: Tạo ba vé số ngẫu nhiên và hiển thị
- Nút quay: Giúp tạo các con số cho từng giải
- Lịch sử: Hiện kết quả lịch sử sau khi ra kết quả quay



Hình 3.18. Giao diện dự đoán.

Khi thực hiện rút vé hay quay hệ thống sẽ cho ra các con số để có thể xem và lựa chọn cho mình con số phù hợp.



Hình 3.19. Giao diện kết quả khi dự đoán.

Ngoài ra các vé sẽ được giữ cố định trên màn hình để có thể dò thay vì phải cuộn lên xuống để xem lại khi bỏ lỡ.

không chỉ cung cấp kết quả quay thử số ngẫu nhiên mà còn lưu trữ lịch sử các lần quay trước đó. Sau mỗi lần quay, kết quả sẽ được lưu lại và người dùng có thể dễ dàng truy cập vào một danh sách lịch sử các lần quay, cho phép họ xem lại các kết quả đã thực hiện. Lịch sử quay này sẽ được hiển thị một cách rõ ràng, bao gồm thông tin về các con số được quay ở từng giải thưởng và số lần lượt quay xuất hiện của chúng. Người dùng có thể dễ dàng theo dõi sự thay đổi của các dãy số theo thời gian, cũng như phân tích và tìm ra các xu hướng trong kết quả xổ số.

Lịch sử quay

Lượt quay 1:
---G8---
6, 8
---G7---
9, 7, 6
---G6---
4, 5, 1, 4
0, 7, 7, 3
4, 7, 9, 8
---G5---
0, 2, 0, 1
---G4---
5, 4, 5, 2, 9
7, 7, 0, 0, 4
3, 9, 7, 6, 0
7, 0, 6, 8, 5
4, 5, 7, 4, 8
8, 9, 6, 3, 2
9, 9, 5, 8, 6
---G3---
3, 6, 0, 4, 2
2, 9, 0, 1, 5
---G2---
1, 3, 1, 2, 8
---G1---
9, 5, 5, 3, 9
ĐẶC BIỆT
5, 0, 8, 3, 0, 9

Lượt quay 2:
---G8---
8, 9
---G7---
5, 7, 1
---G6---
4, 7, 9, 7
2, 0, 0, 2
0, 7, 9, 8
---G5---
4, 5, 6, 9
---G4---
9, 5, 9, 8, 3
4, 1, 4, 8, 2
6, 3, 5, 2, 8
5, 6, 7, 4, 7
3, 8, 8, 3, 4
0, 1, 7, 9, 5
9, 6, 9, 8, 0
---G3---
1, 5, 3, 6, 5
5, 0, 8, 4, 3
---G2---
3, 6, 9, 1, 6
---G1---
7, 4, 6, 1, 9
ĐẶC BIỆT
4, 0, 6, 7, 2, 3

Lượt quay 3:
---G8---
2, 5
---G7---
6, 0, 5
---G6---
2, 3, 5, 1
4, 7, 7, 7
2, 0, 6, 6
---G5---
1, 6, 5, 1
---G4---
6, 6, 2, 2, 0
5, 6, 2, 0, 4
3, 8, 9, 7, 3
8, 5, 0, 1, 1
6, 6, 8, 6, 8
5, 7, 4, 9, 2
9, 4, 3, 7, 7
---G3---
9, 9, 6, 1, 7
8, 0, 6, 8, 8
---G2---
4, 3, 4, 4, 4
---G1---
9, 9, 3, 0, 8
ĐẶC BIỆT
0, 3, 7, 6, 8, 7

Lượt quay 4:
---G8---
8, 6
---G7---
0, 0, 2

Lượt quay 5:
---G8---
2, 4
---G7---
1, 0, 9

Hình 3.20. Giao diện lịch sử.

CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU

4.1. Môi trường thực hiện

Môi trường thực hiện là yếu tố quan trọng để đảm bảo quá trình phát triển và kiểm thử hệ thống được diễn ra trong điều kiện ổn định và hiệu quả. Trong chương này sẽ mô tả các yếu tố cơ bản về phần cứng, phần mềm và cấu hình hệ thống đã được sử dụng trong quá trình thực hiện nghiên cứu.

❖ Cấu hình phần cứng

- Vi xử lý (CPU): Intel Core i5-1035G1.
- RAM: 20GB.
- Card đồ họa (VGA): NVIDIA GeForce MX330.

❖ Cấu hình phần mềm

- Hệ điều hành: Windows 11
- Python: v3.12.6
- Visual Studio Code (VSCode): v1.95.3
- Django: v5.1.1
- MySQL Workbench: v8.0
- Playwright: v1.47.0

4.2. Đánh giá kiểm thử

❖ Một số nội dung kiểm thử

- Kiểm tra xem trang web có tải và hiển thị đúng các thành phần giao diện như trang chủ, trang thống kê, trang dự đoán, và trang tin tức.
- Xác minh các nút chức năng trên các trang có hoạt động và thực hiện đúng nhiệm vụ như chuyển trang, hiển thị dữ liệu hoặc thực hiện thao tác.
- Kiểm tra phần thống kê để đảm bảo các con số được hiển thị đúng cách và khớp với dữ liệu trong cơ sở dữ liệu.
- Kiểm tra các nút điều chỉnh số lượng ở trang thống kê, đảm bảo chúng hiển thị và cập nhật kết quả thống kê đúng theo yêu cầu.
- Kiểm tra biểu đồ thống kê, đảm bảo dữ liệu hiển thị chính xác và không bị lỗi hiển thị.

- Kiểm tra thao tác tạo vé, nhấn nút quay và lưu kết quả vào lịch sử.
 - Xác minh việc hệ thống có tạo ra các con số ngẫu nhiên khi quay thử và hiển thị đúng kết quả.
 - Đảm bảo các thông tin, tin tức hiển thị đầy đủ trên các trang liên quan.
 - Kiểm tra chức năng chụp màn hình và tải xuống hoạt động đúng cách.
- ❖ Kiểm tra hiệu năng.

Bảng 3.18. Bảng hiệu năng của các trang web khi thống kê.

Các hệ thống hoạt động trong 5 phút khi thực hiện thống kê	CPU (%)	RAM (MB)
Xổ vàng	8.8%	97.7 MB
Xoso.com.vn	5.5%	107.8 MB
Minh Ngọc	8.6%	124.3 MB

- ❖ Kiểm tra độ chính xác.

Dựa theo kết quả xổ số đã thu thập. Thực hiện hiển thị thống kê 4 số gần nhất ở giải 8 cho thấy các con số thống kê đã được thực thi và hiển thị chính xác. Với các giải khác đã hiển thị đúng với số liệu và tần suất hiển thị.

	id	prize_8
▶	1	69
	2	10
	3	72
	4	38
	5	55
	6	75

Kết quả thống kê giải 8:	
Giải 8	Số lần xuất hiện
10	1
38	1
69	1
72	1

Hình 3.21. Độ chính xác của kết quả.

Đối với tính năng quay thử và rút vé trong dự đoán, nó được thiết kế nhằm tạo ra các dãy số ngẫu nhiên để người dùng tham khảo, đưa ra các suy luận và giải trí. Vì vậy, không cần kiểm tra độ chính xác cho tính năng này, vì kết quả quay thử không nhằm mục đích khớp với kết quả thực tế.

❖ So sánh các website phổ biến.

Bảng 3.19. Bảng so sánh các website.

Tiêu chí	Xổ Vàng	Minh Ngọc	Xoso.com.vn
Giao diện hiện đại trực quan	x	-	x
Thống kê tần suất theo (2 số cuối, đầu cuối,)	x	x	x
Thống kê tần suất theo (3 số cuối, vị trí giữa hay các vị trí khác, 2 số đầu, hàng chục)	x	-	-
Hiển thị các kết quả thống kê các giải	x	-	-
Tính năng quay thử	x	-	x
Tính năng rút vé	x	-	-
Lịch sử quay thử	x	-	-
Biểu đồ trực quan	x	-	-
Chức năng tin tức	x	-	-
Quảng cáo trong website	Không quảng cáo	Có quảng cáo	Có quảng cáo

❖ Kết quả kiểm thử

Sau khi thực hiện các kiểm thử trên, kết quả cho thấy trang web hoạt động ổn định, dữ liệu hiển thị chính xác và đã đáp ứng đầy đủ các yêu cầu đề ra.

CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1. Kết quả đạt được

Trong quá trình nghiên cứu và thực hiện đề tài, hệ thống thống kê kết quả xổ số kiến thiết trên nền tảng web đã được xây dựng thành công, hoạt động ổn định và không gặp phải lỗi nghiêm trọng. Hệ thống đã đảm bảo lưu trữ, quản lý và cập nhật dữ liệu kết quả xổ số của nhiều tỉnh thành một cách tự động, duy trì độ chính xác và tính kịp thời cho dữ liệu.

Giao diện web được thiết kế thân thiện, giúp người dùng dễ dàng chọn tỉnh thành và xác định số lượng con số cần thống kê. Kết quả thống kê không chỉ được thể hiện qua bảng số liệu mà còn được trực quan hóa bằng các biểu đồ như biểu đồ cột, giúp người dùng dễ dàng theo dõi và phân tích. Trang tin tức cũng được hiển thị các thông tin mới và thông báo các bản cập nhật cho người dùng kịp thời theo dõi.

Bên cạnh đó, hệ thống còn tích hợp trang dự đoán, cho phép người dùng quay thử và rút vé nhằm kiểm tra khả năng trúng thưởng. Sau mỗi lượt quay, kết quả được lưu trữ và hiển thị lại dưới dạng lịch sử, tăng tính tiện ích cho người sử dụng. Hơn nữa, tính năng xuất kết quả ra PDF cũng được bổ sung, hỗ trợ người dùng lưu trữ thông tin một cách thuận tiện.

Giao diện được thiết kế đơn giản, dễ sử dụng và tương thích với nhiều loại thiết bị như máy tính, điện thoại di động và máy tính bảng, đáp ứng nhu cầu sử dụng đa dạng của người dùng.

5.2. Những hạn chế

Mặc dù hệ thống hoạt động tốt và đạt được những kết quả nhất định, vẫn tồn tại một số hạn chế cần khắc phục. Giao diện hiển thị các thống kê đôi khi chưa được tối ưu, có thể gây rối mắt cho một số người dùng khi phải đối mặt với lượng dữ liệu và thông tin lớn được hiển thị cùng lúc. Ngoài ra, các phân tích thống kê hiện tại còn khá ít và chưa thể cung cấp những thông tin chuyên sâu cho những người có nhu cầu nghiên cứu kỹ lưỡng hơn.

Trang dự đoán hiện vẫn chưa được tối ưu hóa về giao diện và chức năng phân tích, chưa đáp ứng được kỳ vọng của người dùng trong việc đưa ra các gợi ý hoặc dự đoán có cơ sở.

5.3. Hướng phát triển

Để cải thiện và mở rộng chức năng của hệ thống, hướng phát triển trong tương lai sẽ tập trung vào việc nâng cao tính năng và hiệu quả hoạt động của ứng dụng. Một trong những cải tiến quan trọng là việc nghiên cứu và áp dụng các loại hình thống kê

khác, cho phép người dùng tự do lựa chọn phương pháp phân tích phù hợp với nhu cầu cá nhân. Đồng thời, hệ thống sẽ bổ sung tính năng lọc và tùy chỉnh cách hiển thị thống kê để giảm sự phức tạp trong giao diện và nâng cao trải nghiệm người dùng.

Về tính năng dự đoán, hệ thống cần áp dụng các phương pháp học máy (machine learning) để cải thiện khả năng phân tích dữ liệu lịch sử và đưa ra dự đoán chính xác hơn. Việc sử dụng các thuật toán như Decision Tree (Cây Quyết Định) sẽ giúp phân tích các mẫu dữ liệu và tạo ra các gợi ý mang tính xác suất cao hơn cho người dùng khi thực hiện quay thử.

Bên cạnh đó, việc tích hợp các mốc thời gian và tùy chỉnh thông tin thống kê theo từng khoảng thời gian cụ thể sẽ thay thế phương pháp chọn số lượng con số thủ công, mang lại sự linh hoạt và chính xác hơn trong việc đáp ứng nhu cầu người dùng. Những cải tiến này không chỉ giúp hệ thống hoàn thiện hơn mà còn nâng cao khả năng cạnh tranh so với các ứng dụng hiện có trên thị trường.

TÀI LIỆU THAM KHẢO

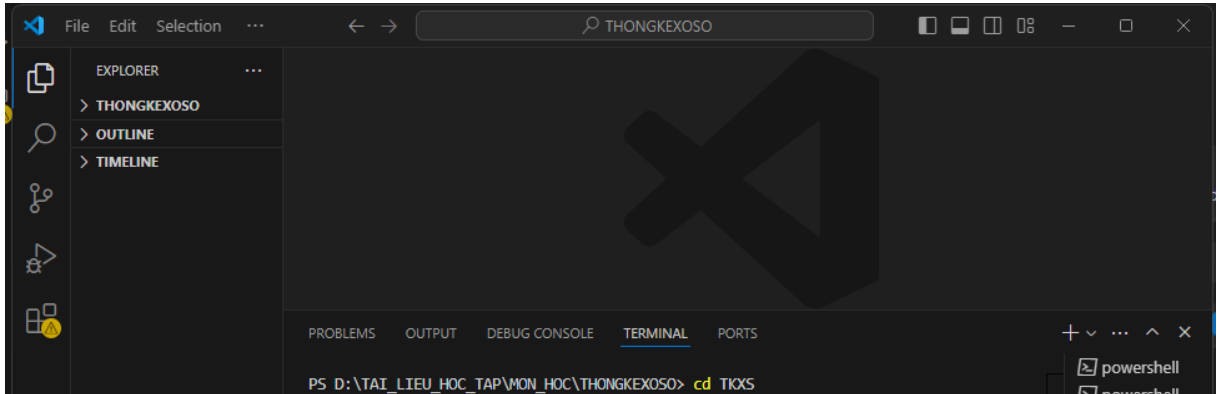
- [1] Gà Python. (2022). Lập trình web bằng Python và Django.
<https://www.youtube.com/c/GàPython>
- [2] Bader, D. (2021). Schedule Library. <https://github.com/dbader/schedule>
- [3] Microsoft. (2020). Playwright Python Documentation.
<https://playwright.dev/python/docs/intro>
- [4] Hege, Ståle, and Jan Egil Refsnes (1998). W3Schools documentation.
<https://www.w3schools.com>
- [5] Chart.js Contributors. (2022). Chart.js Documentation.
<https://www.chartjs.org/docs/latest/>
- [6] Mozilla Foundation. (2015). Math.random Documentation.
https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Math/random
- [7] Django Software Foundation. (2024). Django Documentation Version 5.1.
<https://docs.djangoproject.com/en/5.1/>
- [8] Mozilla Foundation. (2005). Web Technologies Documentation.
<https://developer.mozilla.org/en-US/docs/Web>

PHỤ LỤC

1. Hướng dẫn cài đặt

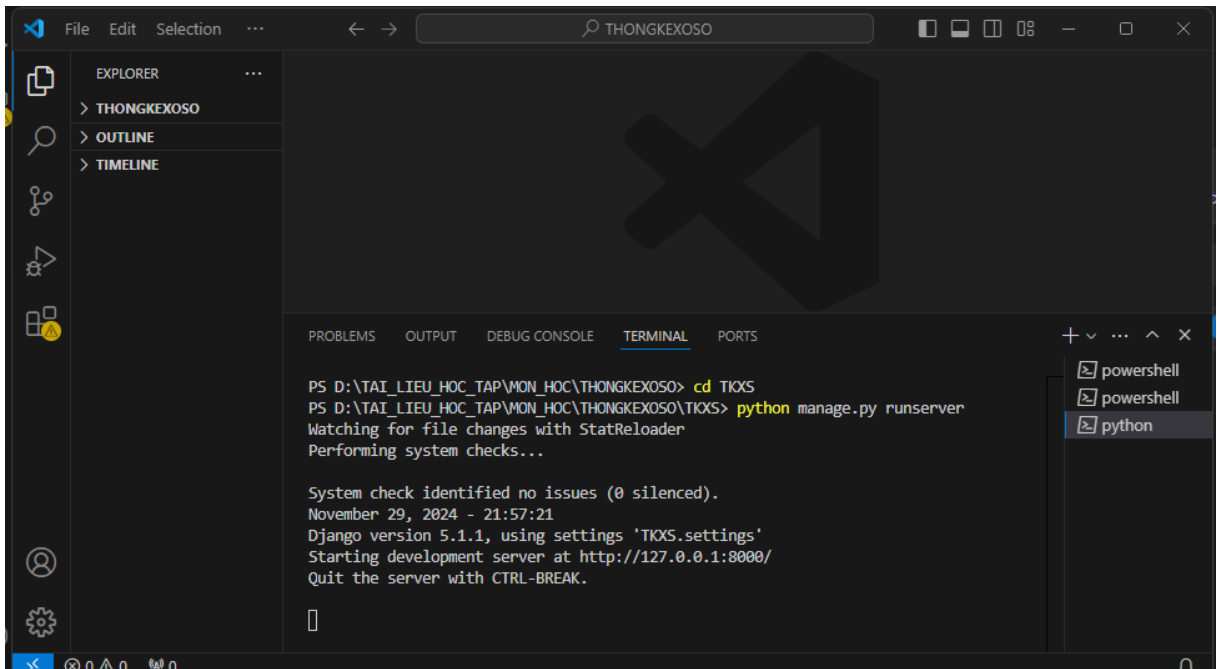
1.1. Cài đặt VS Code

- Tải và cài đặt phần mềm thông qua trang chủ <https://code.visualstudio.com/Download>
- Mở Visual Studio Code chọn thư mục của dự án.



Giao diện Visual Studio Code

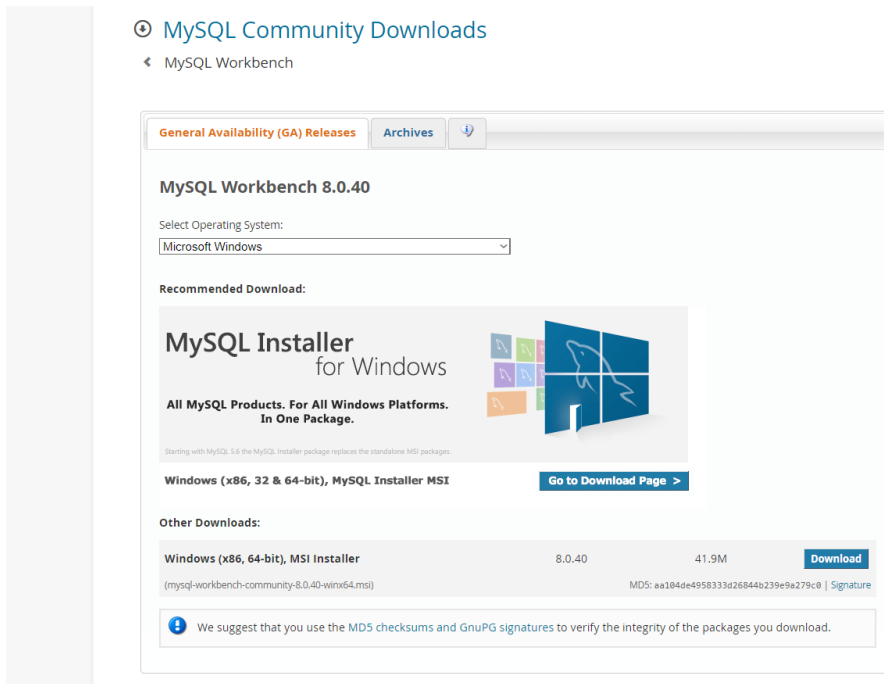
- Khởi chạy server để truy cập trang web.



Giao diện Visual Studio Code

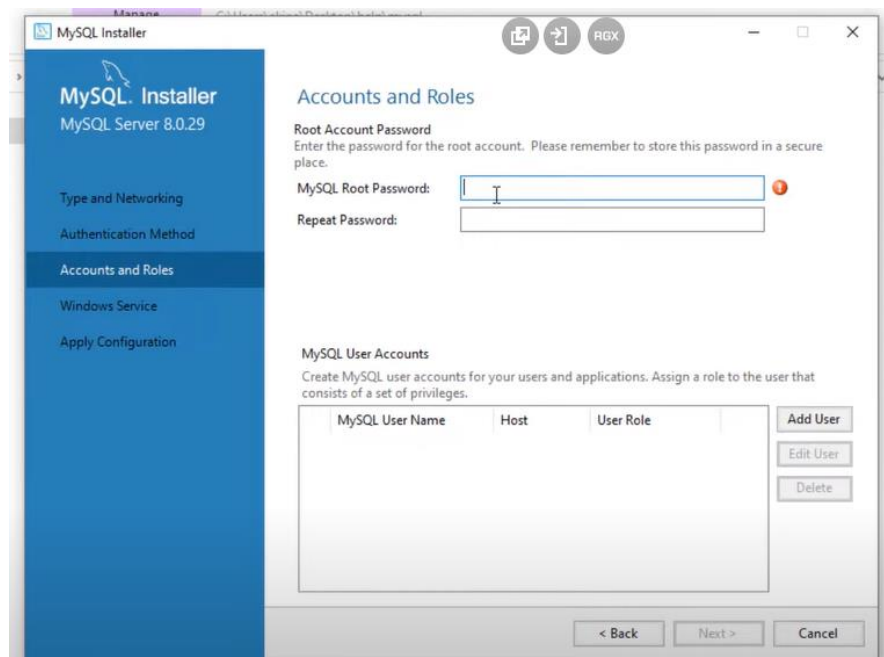
1.2. Cài đặt MySql Workbench

- Tải và cài đặt phần mềm thông qua trang chủ hoặc các kênh youtube.



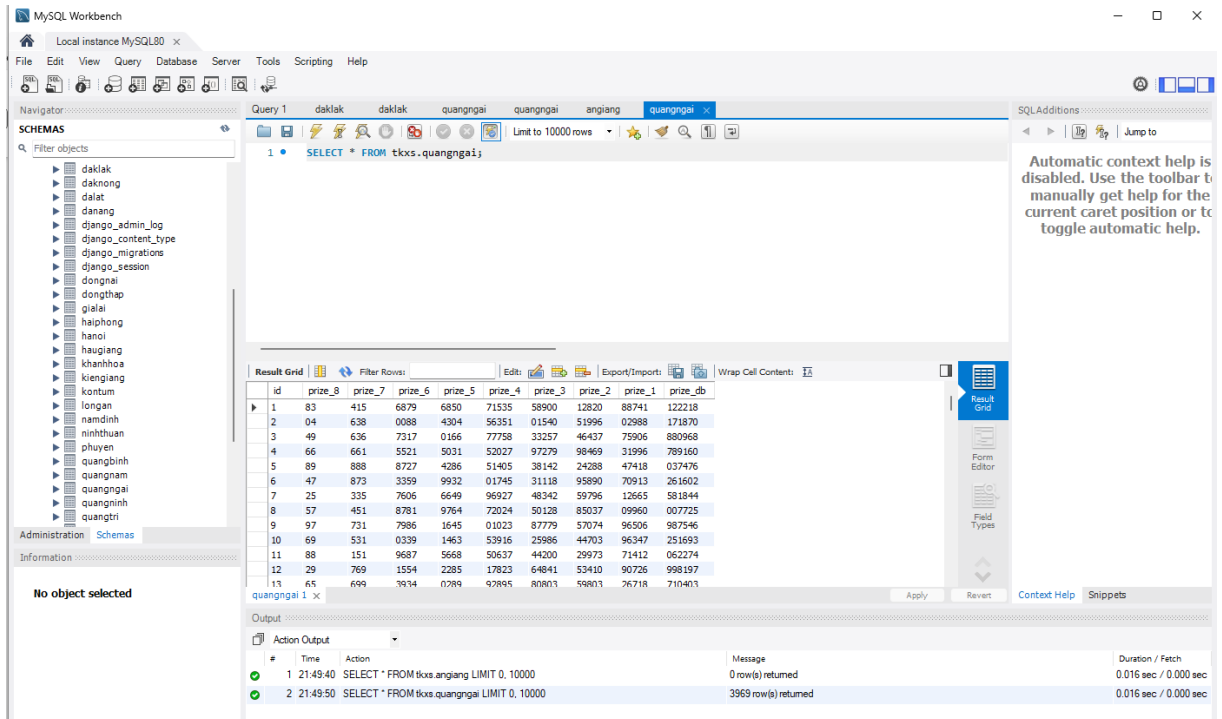
Giao diện trang chủ MySql Workbench

- Tiến hành cài đặt sau khi tải phần mềm và đặt mật khẩu root là 123456.



Giao diện đặt mật khẩu

- Mở phần mềm và nhập mật khẩu khi đã cài đặt xong. Chọn bảng của các tỉnh để xem các số liệu đã thu thập.



Giao diện MySql Workbench