**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**ĐỒ ÁN TỐT NGHIỆP**

**Website chia sẻ video với bộ lọc nội dung không phù hợp**

**NGÔ QUỐC DŨNG**

dung.nq183896@sis.hust.edu.vn

**Ngành CNTT**

**Chuyên ngành CNTT**

|  |  |
| --- | --- |
| **Giảng viên hướng dẫn:** | Ths. Nguyễn Tiến Thành |
| **Bộ môn:** | Công nghệ phần mềm |
| **HÀ NỘI, 6/2022** | |

# Lời cam kết

Họ và tên sinh viên: Ngô Quốc Dũng. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Điện thoại liên lạc: 0967916836 . . . . . . . . . . Email: dung.nq183896@sis.hust.edu.vn. . . . .

Lớp: CNTT 01 – K63. . . . . . . . . . . . . . . . . . . Hệ đào tạo: Đại học đại trà. . . . . . . . . . . . . .

Tôi – *Ngô Quốc Dũng* – cam kết Đồ án Tốt nghiệp (ĐATN) là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *Thạc sĩ Nguyễn Tiến Thành*. Các kết quả nêu trong ĐATN là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong ĐATN – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.

|  |  |
| --- | --- |
|  | *Hà Nội, ngày 24 tháng 6 năm 2022*  Tác giả ĐATN  *Họ và tên sinh viên* |

# Lời cảm ơn

Lời đầu tiên, em xin chân thành cảm ơn Trường Công nghệ Thông tin và Truyền thông – Đại học Bách Khoa Hà Nội đã tạo điều kiện cho em và các bạn được học tập và tiếp xúc với những kiến thức quan trọng của chương trình học để chúng em tự tin và sẵn sàng với những công việc sau này.

Em xin gửi lời cảm ơn sâu sắc đến ThS. Nguyễn Tiến Thành – người đã dìu dắt chúng em từ những ngày đầu vào trường cho đến tận bây giờ. Đặc biệt, thầy còn là người đồng hành hướng dẫn em trong quá trình nghiên cứu, xây dựng và hoàn thành Đồ án Tốt nghiệp. Không có sự đóng góp của thầy, em đã không thể hoàn thiện đồ án này như hiện giờ.

Ngoài ra, em cũng xin cảm ơn các thầy cô trong trường đã dạy bảo tận tình và giúp đỡ em trong suốt quá trình học tập tại trường, tạo nền tảng vững chắc để em trở thành nguồn nhân lực chất lượng cao trong những công việc sau này.

Cuối cùng, em xin cảm ơn gia đình, bạn bè và đồng nghiệp – những người đã luôn bên cạnh động viên và hỗ trợ em khi em gặp khó khăn và tạo động lực để em theo đuổi những thành tựu mới trong quá trình học tập tại Trường Công nghệ Thông tin và Truyền thông – Đại học Bách Khoa Hà Nội.

Do thời gian có hạn, cũng như kiến thức chuyên môn còn non kém, ĐATN của em sẽ không tránh khỏi những thiết sót và hạn chế. Em rất mong nhận được những đánh giá, nhận xét từ thầy cô và các bạn, để từ đấy nhận thấy những ưu, nhược điểm của bản thân, đúc kết và xây dựng kinh nghiệm làm việc một cách vững vàng hơn.

Em xin chân thành cảm ơn!

# Tóm tắt

Hiện nay với sự phổ biến của các thiết bị thông minh và Internet, việc chia sẻ các video giải trí và học tập trên các nền tảng lớn như Youtube, Tiktok hay thậm chí là Facebook,… đang ngày một phát triển. Tuy nhiên, việc tràn lan những nội dung như vậy trên mạng xã hội cũng tiềm ẩn những nguy cơ như tin giả, nạn bắt nạt trên mạng và đặc biệt hơn là tổn thương tâm lý đối với người xem.

Để tránh phát tán những nội dung không phù hợp trên không gian mạng, nhiều nền tảng lớn như Youtube đã triển khai cơ chế riêng để lọc các video được đăng tải bằng machine learning cùng với sự giám sát của con người. Đây là một biện pháp khá hữu hiệu khi đa số các video không phù hợp được loại bỏ nhanh trước khi tiếp xúc với người dùng. Tuy nhiên, do số lượng người dùng lớn cộng với các video có độ dài khác nhau nên việc kiểm duyệt tự động thường được thực hiện không đồng bộ (tức là có thể đăng video xong thì mới được kiểm duyệt) dẫn đến vẫn còn khả năng số ít người dùng có thể đã xem video trước khi nó bị xóa bỏ.

Nhận thấy hạn chế của hướng tiếp cận này, em quyết định xây dựng một website chia sẻ video giả lập sử dụng chính công nghệ deep learning và computer vision để giúp người quản trị kiểm soát các nội dung không phù hợp, với sự khác biệt cơ bản là các video khi được đăng tải sẽ phải trải qua các bước sàng lọc rõ ràng trước khi được đăng tải.

Sở dĩ em chọn cách giải quyết này bởi nó sẽ tăng tính phòng ngừa với các video nguy hiểm và đảm bảo các video được chọn là phù hợp với tiêu chuẩn chung của trang web. Để làm được điều này, em sẽ sử dụng một Deep learning Model giải quyết bài toán Image classification sau đó chia video thành từng ảnh để phân loại. Video sẽ chỉ được đăng tải khi không có chứa nội dung sexy, porn hoặc hentai.

Qua đồ án này, em đã đề xuất một hướng tiếp cận chặt chẽ hơn trong việc kiểm duyệt nội dung, kết quả đạt được tuy không hoàn hảo nhưng đã cung cấp cái nhìn chung về việc sử dụng AI trong việc kiểm soát các video trên mạng.

# Abstract

Mục này khuyến khích sinh viên viết lại mục “Tóm tắt” đồ án tốt nghiệp ở trang trước bằng tiếng Anh. Phần này phải có đầy đủ các nội dung như trong phần tóm tắt bằng tiếng Việt.

Sinh viên không nhất thiết phải trình bày mục này. Nhưng nếu lựa chọn trình bày, sinh viên cần đảm bảo câu từ và ngữ pháp chuẩn tắc, nếu không sẽ có tác dụng ngược, gây phản cảm.

# Mục lục

[Lời cam kết ii](#_Toc44922331)

[Lời cảm ơn iii](#_Toc44922332)

[Tóm tắt iv](#_Toc44922333)

[Abstract v](#_Toc44922334)

[Mục lục vi](#_Toc44922335)

[Danh mục hình vẽ ix](#_Toc44922336)

[Danh mục bảng x](#_Toc44922337)

[Danh mục công thức xi](#_Toc44922338)

[Danh mục các từ viết tắt xii](#_Toc44922339)

[Danh mục thuật ngữ xiii](#_Toc44922340)

[Chương 1 Giới thiệu đề tài 1](#_Toc44922341)

[1.1 Đặt vấn đề 1](#_Toc44922342)

[1.2 Mục tiêu và phạm vi đề tài 2](#_Toc44922343)

[1.3 Định hướng giải pháp 2](#_Toc44922344)

[1.4 Bố cục đồ án 2](#_Toc44922345)

[Chương 2 Khảo sát và phân tích yêu cầu 4](#_Toc44922346)

[2.1 Khảo sát hiện trạng 4](#_Toc44922347)

[2.2 Tổng quan chức năng 4](#_Toc44922348)

[2.2.1 Biểu đồ use case tổng quan 4](#_Toc44922349)

[2.2.2 Biểu đồ use case phân rã XYZ 5](#_Toc44922350)

[2.2.3 Quy trình nghiệp vụ 5](#_Toc44922351)

[2.3 Đặc tả chức năng 5](#_Toc44922352)

[2.3.1 Đặc tả use case A 5](#_Toc44922353)

[2.3.2 Đặc tả use case B 5](#_Toc44922354)

[2.4 Yêu cầu phi chức năng 5](#_Toc44922355)

[Chương 3 Công nghệ sử dụng 6](#_Toc44922356)

[Chương 4 Phát triển và triển khai ứng dụng 7](#_Toc44922357)

[4.1 Thiết kế kiến trúc 7](#_Toc44922358)

[4.1.1 Lựa chọn kiến trúc phần mềm 7](#_Toc44922359)

[4.1.2 Thiết kế tổng quan 7](#_Toc44922360)

[4.1.3 Thiết kế chi tiết gói 8](#_Toc44922361)

[4.2 Thiết kế chi tiết 8](#_Toc44922362)

[4.2.1 Thiết kế giao diện 8](#_Toc44922363)

[4.2.2 Thiết kế lớp 9](#_Toc44922364)

[4.2.3 Thiết kế cơ sở dữ liệu 9](#_Toc44922365)

[4.3 Xây dựng ứng dụng 9](#_Toc44922366)

[4.3.1 Thư viện và công cụ sử dụng 9](#_Toc44922367)

[4.3.2 Kết quả đạt được 9](#_Toc44922368)

[4.3.3 Minh hoạ các chức năng chính 10](#_Toc44922369)

[4.4 Kiểm thử 10](#_Toc44922370)

[4.5 Triển khai 10](#_Toc44922371)

[Chương 5 Các giải pháp và đóng góp nổi bật 11](#_Toc44922372)

[Chương 6 Kết luận và hướng phát triển 12](#_Toc44922373)

[6.1 Kết luận 12](#_Toc44922374)

[6.2 Hướng phát triển 12](#_Toc44922375)

[Tài liệu tham khảo 13](#_Toc44922376)

[Phụ lục A-1](#_Toc44922377)

[A Hướng dẫn viết đồ án tốt nghiệp A-1](#_Toc44922378)

[A.1 Quy định chung A-1](#_Toc44922379)

[A.2 Ngành học A-2](#_Toc44922380)

[A.3 Tạo đề mục A-2](#_Toc44922381)

[A.4 Bảng biểu A-3](#_Toc44922382)

[A.5 Hình vẽ A-4](#_Toc44922383)

[A.6 Tài liệu tham khảo A-4](#_Toc44922384)

[A.7 Công thức toán học A-5](#_Toc44922385)

[A.8 Tham chiếu chéo A-5](#_Toc44922386)

[A.9 Cập nhật mục lục và tham chiếu chéo A-6](#_Toc44922387)

[A.10 In quyển đồ án tốt nghiệp A-6](#_Toc44922388)

[B Đặc tả use case B-8](#_Toc44922389)

[B.1 Đặc tả use case “Thống kê tình hình mượn sách” B-8](#_Toc44922390)

[B.2 Đặc tả use case “Đăng ký làm thẻ mượn” B-8](#_Toc44922391)

[C Công nghệ sử dụng C-8](#_Toc44922392)

[C.1 Công nghệ bảo mật dữ liệu C-8](#_Toc44922393)

[C.2 Công nghệ blockchain C-8](#_Toc44922394)

[D Thiết kế gói D-8](#_Toc44922395)

[D.1 Thiết kế gói cho kiến trúc tổng quan D-8](#_Toc44922396)

[D.2 Thiết kế gói cho chức năng “Trả sách” D-9](#_Toc44922397)

[E Thiết kế lớp E-9](#_Toc44922398)

# Danh mục hình vẽ

# Danh mục bảng

# Danh mục công thức

# Danh mục các từ viết tắt

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Danh mục thuật ngữ

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Giới thiệu đề tài

## Đặt vấn đề

Trong bối cảnh mạng xã hội đang ngày càng phát triển đến mức bão hòa, các trang mạng xã hội đang dần chuyển hướng từ nhu cầu chia sẻ và giao tiếp thông thường bằng chữ viết và hình ảnh sang chia sẻ các nội dung dạng video. Bằng chứng là Tiktok và Youtube là hai mạng xã hội lớn nhất hiện nay có nội dung chính là các video. Hay thậm chí các trang như Facebook, Instagram cũng đang dần chuyển sang làm nhiều các nội dung video hơn với các chức năng Story, Shorts và Watch.

Tuy các video mang đến nhiều thông tin hơn cho người xem nhưng chúng cũng tiềm ẩn những nguy cơ mất an toàn trên không gian mạng nếu nội dung không được sàng lọc kỹ càng. Không như những bài viết thông thường, các video có chứa lượng lớn thông tin trực quan từ hình ảnh đến âm thanh. Những video vì vậy sẽ cho người xem thấy được sự việc rõ ràng và sinh động hơn như thể chính mình đang ở hiện trường. Đây là một công cụ vô cùng mạnh mẽ để tuyên truyền và phổ biến thông tin từ giáo dục, giải trí lành mạnh đến bạo lực, tội phạm... và nhiều hơn nữa. Khi chúng ta tiếp cận với các video trên mạng, ta có nguy cơ bị ảnh hưởng bởi các nội dung tiêu cực, sai trái gây tổn thương tâm lý hay tệ hơn là làm sai lệch nhân cách.

Nếu đảm bảo được các video là phù hợp với tiêu chuẩn cộng đồng và không gây ảnh hưởng đến an ninh, an toàn trên không gian mạng, ta có thể bảo vệ được người dùng trước các hành vi quấy phá. Mạng xã hội tuy có nhiều mặt tích cực nhưng cũng đem đến những rủi ro lớn đối với người dùng, một trong số đố là các nội dung độc hại. Bằng cách loại bỏ những video không phù hợp, ta sẽ góp phần xây dựng một cộng đồng văn minh, lịch sự hơn và tận dụng được những lợi ích của mạng xã hội một cách tối đa.

## Mục tiêu và phạm vi đề tài

Hiện nay, với các nền tảng chính chia sẻ nội dung video như Youtube đã có áp dụng trí tuệ nhân tạo hay cụ thể hơn là deep learning vào phân loại và lọc nội dung không phù hợp một cách chặt chẽ. Tuy vậy, thuật toán của Youtube không phải là hoàn hảo và dù đã có sự phối hợp của con người nhưng vẫn không khó để lừa được bộ lọc này. Tiktok là một trong những mạng xã hội lớn nhất thế giới nhưng những nội dung trên Tiktok được kiểm duyệt một cách lỏng lẻo theo quy định của trang mạng. Theo đó những nội dung nhạy cảm ở mức độ vừa phải vẫn có thể xuất hiện trên trang của người dùng tùy theo mục đích của người đăng tải. Đối với Facebook hay các mạng xã hội khác không tập trung chính vào video thì phần lớn thuật toán và trí tuệ nhân tạo sẽ được áp dụng vào những hành vi không phù hợp của người dùng thay vì chính nội dung của các bài đăng. Ví dụ như Facebook kiểm soát những tin nhắn và comments chặt chẽ hơn nhiều so với Youtube.

Hạn chế chính của những ứng dụng trên là hướng tiếp cận giải quyết nhanh chóng vi phạm thay vì phòng ngừa trước khi vi phạm xảy ra. Bộ lọc của Youtube không tự động lọc các video trước khi đăng tải mà chỉ quét những video xuất hiện trên trang chủ. Tuy công nghệ của Google đã phát triển rất nhiều dẫn đến tốc độ xử lý video nhanh hơn đáng kể so với trước đây nhưng sự tăng trưởng của Youtube cũng đang trở thành một thách thức khi mỗi giờ có rất nhiều các đoạn clip được đăng tải với tổng thời lượng khoảng 300 tiếng. Như vậy xác suất để một video không phù hợp đến được với người dùng là vẫn còn có thể xảy ra.

Để khắc phục hạn chế này, em muốn phát triển trang Web chia sẻ video với chức năng chính là lọc các đoạn clip trước khi được đăng tải theo quy trình rõ ràng. Như vậy các nội dung trên Web sẽ được đảm bảo bởi bộ lọc và cả sự giám sát của con người, giảm thiểu tối đa vi phạm và phòng chống được sai phạm trước khi xảy ra.

## Định hướng giải pháp

Để xây dựng bộ lọc video, em sẽ giải quyết bài toán Image Classification sử dụng mạng CNN rồi áp dụng từ hình ảnh vào video bằng cách quét qua từng khung hình trong đoạn phim và tính toán xác suất lớn nhất của từng nhãn. Các nhãn sẽ được chia thành Neutral - nội dung không có vi phạm (không phải tranh vẽ hay hoạt hình), Drawings – nội dung bằng tranh vẽ, hoạt hình hay anime không có vi phạm, Hentai – truyện tranh hay phim hoạt hình 18+, Sexy – nội dung người thật có xu hướng nhạy cảm 13+, Porn – nội dung người thật 18+. Model của em sẽ được Transfer Learning theo mạng InceptionV3 với các lớp Dense ở cuối và được train với số lượng ảnh là khoảng 140000 ảnh.

Khi tích hợp vào trang Web để lọc nội dung, em sẽ sử dụng bộ lọc một cách tự động mỗi khi có video mới được đăng tải. Video sẽ chỉ được đăng lên trang chủ sau khi đã được gán nhãn bởi bộ lọc. Nếu video là phù hợp (Drawings hoặc Neutral) sẽ được tự động đăng lên trang chủ, ngược lại nếu video ko phù hợp (Hentai, Sexy hoặc Porn) sẽ được gán nhãn tương ứng và không được đăng. Trong trường hợp bộ lọc phát hiện nhầm hoặc đánh nhãn sai người dùng có thể yêu cầu phê duyệt bằng tay đến Admin.

Như vậy, qua đồ án này em đã đề xuất một hướng giải quyết đơn giản nhưng an toàn cho bài toán kiểm duyệt nội dung video trên không gian mạng. Kết quả đạt được cho thấy chất lượng của các đoạn clip được chia sẻ phụ thuộc rất nhiều đến model và phương pháp được sử dụng để lọc video. Model hiện tại tuy chưa được tốt nhưng cũng đã cho thấy được sức mạnh của Deep learning trong xử lý và phân loại video. Tuy vậy, do quy mô bài toán nhỏ nên giải pháp có phần đơn thuần, chưa áp dụng được vào thực tiễn khi số lượng video và người dùng đăng tải lớn.

## Bố cục đồ án

Phần còn lại của báo cáo đồ án tốt nghiệp này được tổ chức như sau.

Chương 2 trình bày về quá trình khảo sát và phân tích yêu cầu, trong đó có quá trình phân tích thực trạng và định hình các nhu cầu, chức năng cần thiết của chương trình. Chương này có quan hệ mật thiết với bài toán được giới thiệu ở trên.

Trong Chương 3, em giới thiệu về các công nghệ sử dụng, đặc biệt là mạng CNN và tổng quan về Deep Learning trong Computer Vision. Ngoài ra em cũng có để cập qua đến framework Django được sử dụng để xây dựng trang Web.

Ở chương 4, em sẽ trình bày quá trình phát triển và xây dựng trang Web cũng như Train và Fine-tune model.

# ­­Khảo sát và phân tích yêu cầu

## Khảo sát hiện trạng

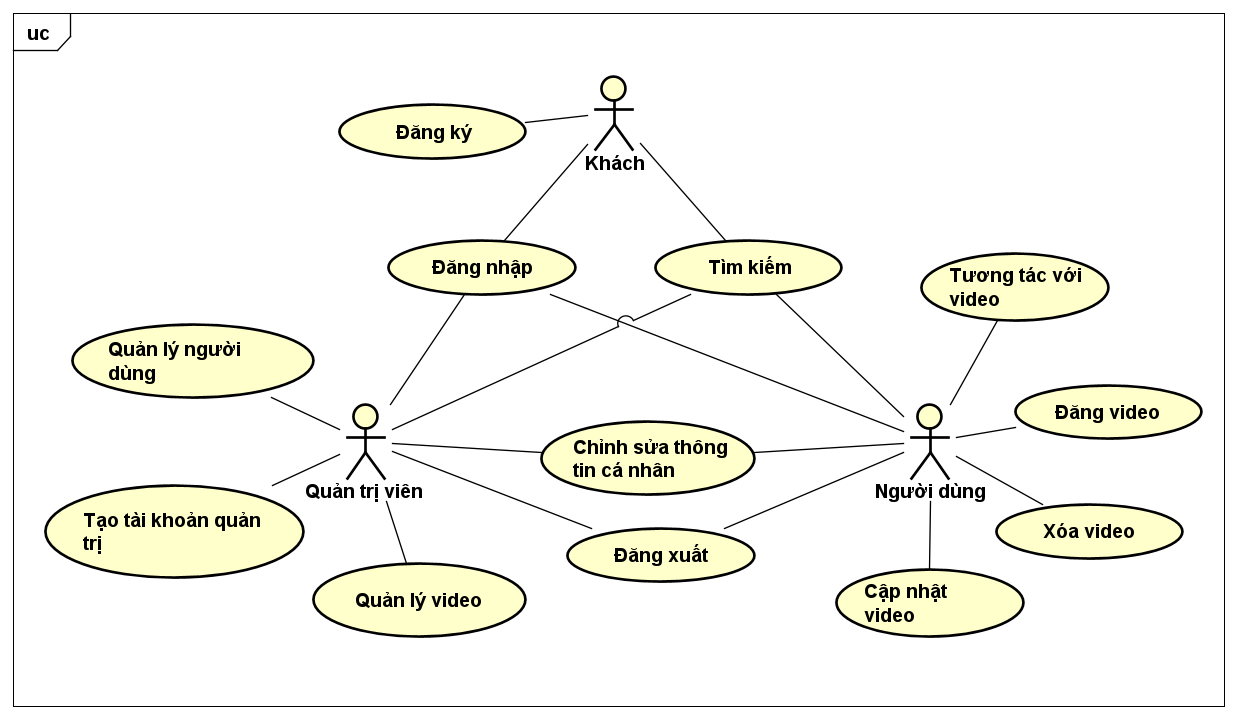
Với các nền tảng chia sẻ video trực tuyến hiện nay, yêu cầu cơ bản là phải cho phép người dùng đăng tải và chia sẻ video cũng như xem các video được chia sẻ trên trang chủ. Ngoài ra, để tiện lợi cho người dùng và đẩy mạnh tương tác giữa người với người hay người dùng với trang web, các hệ thống hiện nay cũng yêu cầu thêm các chức năng như Like/Dislike, Share, Comment và Follow/Subscribe. Để đề cao tính thời sự và tăng tương tác người dùng, các mạng xã hội còn sử dụng chức năng Livestream để phát trực tiếp video. Tuy nhiên khi phát trực tiếp video thì việc quản lý nội dung cũng trở nên khó khan hơn và sẽ phụ thuộc lớn vào báo cáo của các người dùng khác để nhận biết vi phạm. Cuối cùng, để tạo lợi nhuận cho người sáng tạo nội dung cũng như nền tảng chia sẻ video cần có phương thức thanh toán riêng cho phép các nhà tài trợ đăng quảng cáo hoặc người dùng nhận tiền lãi từ mạng xã hội. Các phương thức thanh toán này có khác biệt lớn giữa các mạng xã hội khác nhau và rất khó đi sâu vào tìm hiểu.

Như đã nêu ở Chương 1, các bộ lọc nội dung video trên mạng xã hội thường có sự kết hợp giữa AI và con người ở các mức độ khác nhau. Nếu so sánh giữa các nền tảng phổ biến thì có thể thấy bộ lọc của Youtube có chất lượng vượt trội hơn hẳn khi lọc các video. Tuy nhiên, để đảm bảo tiện lợi cho người dùng thì các video có khả năng không được phân loại trước khi đăng tải nên chưa giải quyết được nguy cơ tiềm ẩn khi một video vi phạm được đăng lên.

Từ những quan sát trên, em hướng tới tạo nên một trang Web mô phỏng tập trung chính vào tính năng lọc nội dung. Trong đó khi đăng tải video sẽ phải được bộ lọc tự động phân loại trước khi thông qua, đồng thời có cho phép kiểm tra lại bởi con người để đảm bảo không có trường hợp loại nhầm. Trang Web cũng phải đảm bảo được yêu cầu chính của một nền tảng chia sẻ Video là có chức năng đăng tải, chia sẻ video cũng như xem video. Tuy nhiên do bộ lọc tập trung vào các nội dung video nên em sẽ lược bỏ một vài chức năng tương tác với người dùng như Comment, Share… Tính năng Livestream do khó kiểm soát trong kiểm duyệt nội dung nên cũng sẽ không được thêm vào.

## Tổng quan chức năng

### Biểu đồ use case tổng quan



**Hình 1** Biểu đồ UC tổng quan

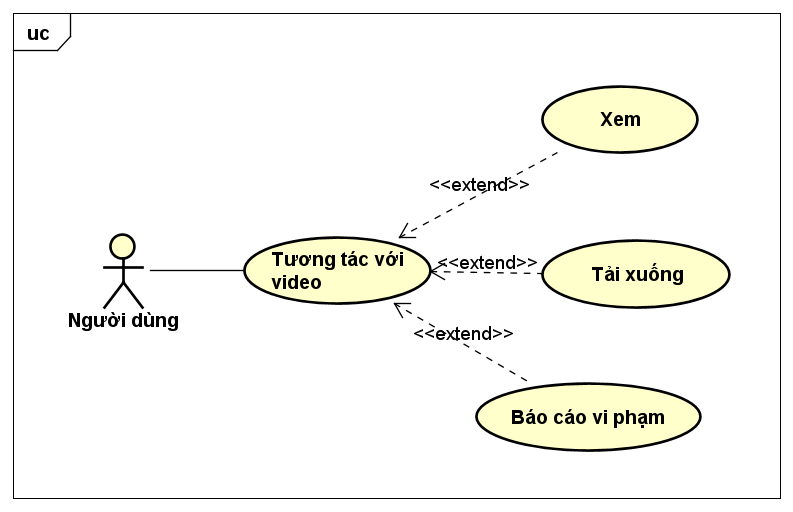
Các tác nhân chính bao gồm Khách vãng lai, Người dùng và Quản trị viên. Trong đó:

Khách vãng lai là đối tượng truy cập vào trang web mà không cần tài khoản. Khách vãng lai có quyền tìm kiếm video, kênh của người dùng và xem các video trên web. Ngoài ra khách vãng lai còn có thể đăng ký tài khoản người dùng hoặc đăng nhập với tài khoản đã có.

Người dùng là đối tượng chính sử dụng trang web khi đã đăng nhập vào tài khoản người dùng. Người dùng có quyền đăng video, chỉnh sửa video, xóa video, tìm kiếm video, kênh, xem video và tương tác với video.

Quản trị viên là đối tượng quản lý trang web sử dụng tài khoản quản trị viên. Quản trị viên có các quyền điều hành trang web như xử lý yêu cầu đăng video, chặn video, khóa/mở khóa tài khoản người dùng và tạo tài khoản quản trị viên.

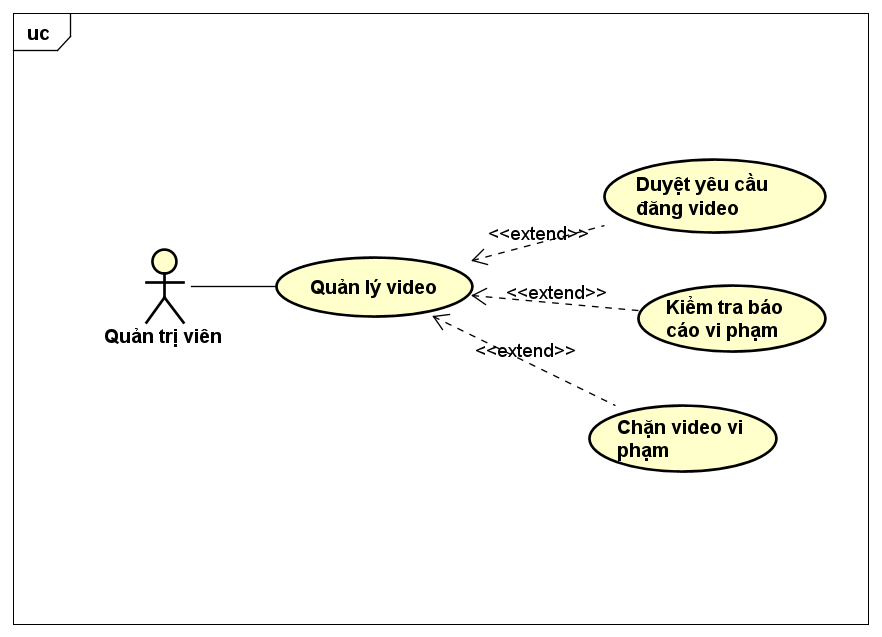
### Biểu đồ use case phân rã tương tác với video



**Hình 2** Biểu đồ UC tương tác với video

Khi tương tác với video người dùng có thể xem video, người dùng cũng có thể tải xuống file video tương ứng hoặc báo cáo vi phạm để đưa yêu cầu đến quản trị viên khi thấy nội dung không phù hợp.

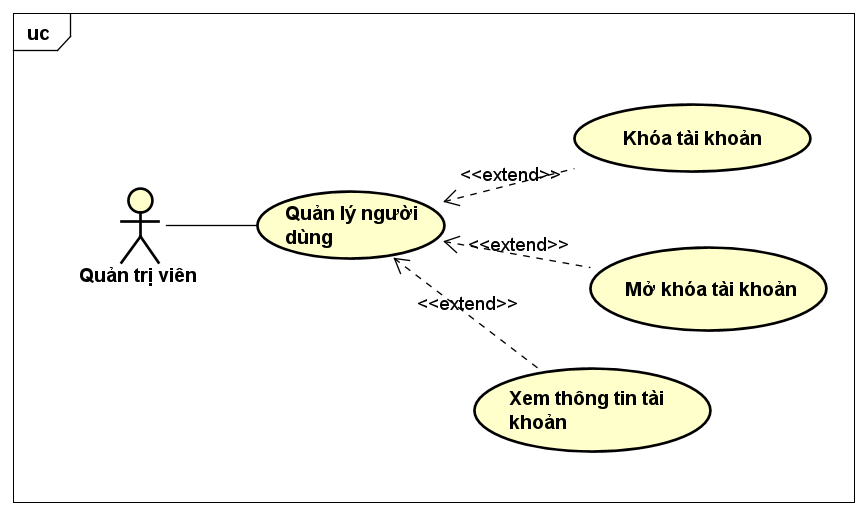
### Biểu đồ use case phân rã quản lý video



**Hình 3** Biểu đồ UC quản lý video

Quản trị viên có quyền giám sát các video trên trang web. Khi có yêu cầu đăng video không được duyệt tự động, người dùng có thể yêu cầu trực tiếp đến quản trị viên. Mặt khác khi có người dùng báo cáo có video vi phạm không được phân loại chính xác, quản trị viên cũng có thể kiểm tra yêu cầu và quyết định có chặn video đó hay không. Quản trị viên cũng có thể chặn video vi phạm bất kì sau khi xem trên trang chủ.

### Biểu đồ use case phân rã quản lý người dùng



**Hình 4** Biểu đồ UC quản lý người dùng

Quản trị viên có quyền xem thông tin người dùng, khóa tài khoản người dùng trong trường hợp có hành vi vi phạm và cũng có thể mở khóa tài khoản đã khóa. Khi tài khoản người dùng bị khóa người dùng đó sẽ không thể đăng video cho đến khi tài khoản được mở khóa bởi người quản trị.

### Quy trình nghiệp vụ

Diagram, schematic

Description automatically generated

**Hình 5** Biểu đồ hoạt động đăng video

Quy trình đăng nghiệp vụ video như sau: Người dùng tạo bài đăng mới và đăng lên Web, hệ thống sẽ tự động quét video và phân loại nội dung, nếu phù hợp video sẽ được tự động đăng, nếu bị lọc người dùng có thể yêu cầu duyệt bằng tay đến quản trị viên.

## Đặc tả chức năng

### Đặc tả use case Đăng nhập

|  |  |
| --- | --- |
| Mã usecase | UC001 |
| Tác nhân | Khách |
| Mục đích sử dụng | Khách đăng nhập vào hệ thống |
| Sự kiện kích hoạt | Khách chọn chức năng đăng nhập |
| Tiền điều kiện | Không |
| Luồng sự kiện chính  (Thành công) | 1. Khách chọn chức năng Đăng nhập 2. Hệ thống hiển thị giao diện đăng nhập 3. Khách nhập username và password 4. Khách yêu cầu đăng nhập 5. Hệ thống kiểm tra Khách đã nhập các trường bắt buộc nhập hay chưa 6. Hệ thống kiểm tra username và password có hợp lệ do Người dùng đăng kí với hệ thống hay không, Hệ thống kiểm tra vai trò của Khách trong hệ thống. 7. Hệ thống thông báo đăng nhập thành công và chuyển sang giao diện trang chủ, giao diện phù hợp với vai trò trong hệ thống. |
| Luồng sự kiện thay thế | 6a. Hệ thống thông báo lỗi: cần nhập các trường bắt buộc nếu khách nhập thiếu.  7a. Hệ thống thông báo lỗi: Username hoặc mật khẩu không đúng nếu không tìm thấy username hoặc password trong hệ thống. |
| Hậu điều kiện | Hệ thống thông báo đăng nhập thành công |

**Bảng 1** Đặc tả UC Đăng Nhập

### Đặc tả use case Đăng video

|  |  |
| --- | --- |
| Mã usecase | UC002 |
| Tác nhân | Người dùng |
| Mục đích sử dụng | Người dùng đăng video lên trang chủ |
| Sự kiện kích hoạt | Người dùng chọn tạo bài đăng mới |
| Tiền điều kiện | Không |
| Luồng sự kiện chính  (Thành công) | 1. Người dùng chọn chức năng tạo bài mới 2. Hệ thống hiển thị giao diện đăng bài 3. Người dùng nhập tên bài đăng, mô tả và gán file video 4. Người dùng yêu cầu đăng bài 5. Hệ thống kiểm tra Người dùng đã nhập các trường bắt buộc nhập hay chưa 6. Hệ thống kiểm tra thông tin bài đăng có hợp lệ không và file đã đúng định dạng chưa 7. Hệ thống tải lên video và thông báo với người dùng video đang chờ xử lý 8. Quản trị viên kích hoạt bộ lọc tự động 9. Hệ thống tiến hành quét video và cập nhật trạng thái bài đăng thành đang xử lý 10. Hệ thống thông báo video hợp lệ và cập nhật trạng thái bài đăng lên trang chủ |
| Luồng sự kiện thay thế | 5a. Hệ thống thông báo lỗi: cần nhập các trường bắt buộc nếu khách nhập thiếu.  6a. Hệ thống thông báo lỗi: định dạng file không phải video.  10a. Hệ thống thông báo video không phù hợp.  10b. Người dùng chọn yêu cầu duyệt bằng tay đến quản trị viên  10c. Hệ thống thông báo bài đăng dược duyệt hoặc bị từ chối |
| Hậu điều kiện | Hệ thống thông báo bài đăng phù hợp và đã được đăng tải |

**Bảng 2** Đặc tả UC Đăng Video

### Đặc tả use case Duyệt yêu cầu đăng Video

|  |  |
| --- | --- |
| Mã usecase | UC003 |
| Tác nhân | Quản trị viên |
| Mục đích sử dụng | Quản trị viên kiểm tra lại yêu cầu đăng video đã bị lọc bởi hệ thống |
| Sự kiện kích hoạt | Quản trị viên chọn xem yêu cầu đăng bài |
| Tiền điều kiện | Người dùng yêu cầu duyệt bài đăng bằng tay |
| Luồng sự kiện chính  (Thành công) | 1. Quản trị viên chọn xem yêu cầu đăng bài cần duyệt 2. Hệ thống hiển thị giao diện bài đăng và vi phạm 3. Quản trị viên xem nội dung video và chọn Đồng ý hay Từ chối yêu cầu 4. Hệ thống tự động đăng bài nếu được Đồng ý và trả yêu cầu về người dùng nếu từ chối 5. Hệ thống thông báo bài đăng đã được Duyệt hoặc Từ chối |
| Luồng sự kiện thay thế | Không |
| Hậu điều kiện | Hệ thống cập nhật tình trạng bài đăng đến người dùng. Nếu Đồng ý, bài đăng được đăng lên trang chủ |

**Bảng 3** Đặc tả UC Duyệt yêu cầu đăng Video

### Đặc tả use case Báo cáo vi phạm

|  |  |
| --- | --- |
| Mã usecase | UC004 |
| Tác nhân | Người dùng |
| Mục đích sử dụng | Người dùng báo cáo có video vi phạm và yêu cầu quản trị viên kiểm tra |
| Sự kiện kích hoạt | Người dùng chọn báo cáo vi phạm |
| Tiền điều kiện | Bài đăng đã được đăng tải trên trang chủ |
| Luồng sự kiện chính  (Thành công) | 1. Người dùng chọn chức năng báo cáo vi phạm 2. Hệ thống hiển thị giao diện báo cáo vi phạm 3. Người dùng chọn lý do vi phạm 4. Người dùng yêu cầu gửi báo cáo 5. Hệ thống kiểm tra Người dùng đã nhập lý do vi phạm hay chưa 6. Hệ thống thông báo báo cáo đã được gửi đến Quản trị viên |
| Luồng sự kiện thay thế | 5a. Hệ thống thông báo lỗi: cần nhập lý do vi phạm |
| Hậu điều kiện | Hệ thống tạo báo cáo vi phạm mới đến Quản trị viên |

**Bảng 4** Đặc tả UC Báo cáo vi phạm

### Đặc tả use case Chặn Video vi phạm

|  |  |
| --- | --- |
| Mã usecase | UC005 |
| Tác nhân | Quản trị viên |
| Mục đích sử dụng | Quản trị viên chặn video vi phạm trên trang chủ |
| Sự kiện kích hoạt | Quản trị viên chọn chặng video vi phạm |
| Tiền điều kiện | Bài đăng đã được đăng tải trên trang chủ |
| Luồng sự kiện chính  (Thành công) | 1. Quản trị viên chọn chức năng chặn video vi phạm 2. Hệ thống hiển thị giao diện báo cáo vi phạm 3. Quản trị viên chọn lý do vi phạm 4. Quản trị viên yêu cầu chặn video 5. Hệ thống kiểm tra Quản trị viên đã nhập lý do vi phạm hay chưa 6. Hệ thống cập nhật trạng thái video thành bị chặn 7. Hệ thống thông báo video vi phạm đã bị chặn |
| Luồng sự kiện thay thế | 5a. Hệ thống thông báo lỗi: cần nhập lý do vi phạm |
| Hậu điều kiện | Hệ thống cập nhật trạng thái video thành bị chặn |

**Bảng 5** Đặc tả UC Chặn Video vi phạm

## Yêu cầu phi chức năng

Chương trình yêu cầu tính chặt chẽ trong kiểm duyệt nội dung nên cần có tính bảo mật cao, phân quyền rõ ràng: Quản trị viên có quyền cao nhất, tiếp đến là Người dùng và Khách vãng lai.

Vì bài toán giả lập nên quy mô Server nhỏ chỉ đáp ứng được 2-3 yêu cầu đăng video một lúc với thời gian xử lý tùy thuộc vào độ dài video và nội dung (nếu có vi phạm). Ngoài ra cơ sở dữ liệu Sqlite3 là một cơ sở dữ liệu nhúng nên số lượng người dùng và bài đăng cũng bị giới hạn ở mức vài trăm. Sqlite3 cũng không cho phép truy cập đồng thời vào cơ sở dữ liệu.

# Công nghệ sử dụng

## Deep Learning

Trong lĩnh vực nghiên cứu AI, Machine Learning đã đạt được thành công đáng kể trong những năm gần đây – cho phép máy tính vượt qua hoặc tiến gần đến việc kết hợp hiệu suất của con người trong các lĩnh vực từ nhận dạng khuôn mặt đến nhận dạng giọng nói và ngôn ngữ.

Machine Learning là quá trình dạy máy tính thực hiện một nhiệm vụ, thay vì lập trình nó làm thế nào để thực hiện nhiệm vụ đó từng bước một. Khi kết thúc đào tạo, một hệ thống Machine Learning sẽ có thể đưa ra dự đoán chính xác khi được cung cấp dữ liệu.

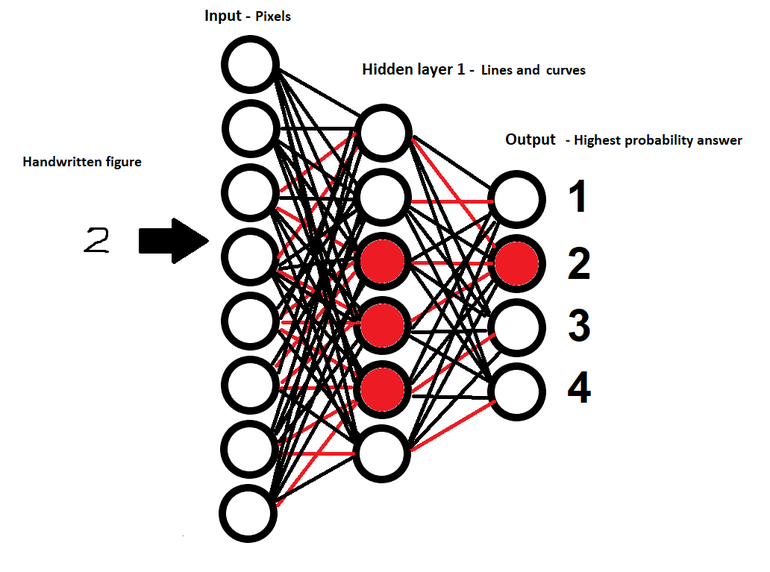
Machine Learning thường được chia thành học có giám sát, trong đó máy tính học bằng ví dụ từ dữ liệu được gắn nhãn và học không giám sát, trong đó các máy tính nhóm các dữ liệu tương tự và xác định chính xác sự bất thường.

Deep Learning là một tập hợp con của Machine Learning, có khả năng khác biệt ở một số khía cạnh quan trọng so với Machine Learning nông truyền thống, cho phép máy tính giải quyết một loạt các vấn đề phức tạp không thể giải quyết được.

Để giải quyết vấn đề này, máy tính cần phải có khả năng đối phó với sự đa dạng lớn trong cách thức trình bày dữ liệu. Mỗi chữ số từ 0 đến 9 có thể được viết theo vô số cách: kích thước và hình dạng chính xác của mỗi chữ số viết tay có thể rất khác nhau tùy thuộc vào người viết và trong hoàn cảnh nào.

Đối phó với sự biến đổi của các tính năng này và sự lộn xộn tương tác lớn hơn giữa chúng, là nơi học tập sâu và mạng nơ-ron sâu trở nên hữu ích.

Mạng nơ-ron là các mô hình toán học có cấu trúc được lấy cảm hứng lỏng lẻo từ bộ não. Mỗi nơ-ron trong mạng nơ-ron là một hàm toán học lấy dữ liệu thông qua đầu vào, biến đổi dữ liệu đó thành dạng dễ điều chỉnh hơn và sau đó phun ra thông qua đầu ra. Bạn có thể nghĩ về các nơ-ron trong một mạng lưới thần kinh như được sắp xếp theo lớp, như hình dưới đây.

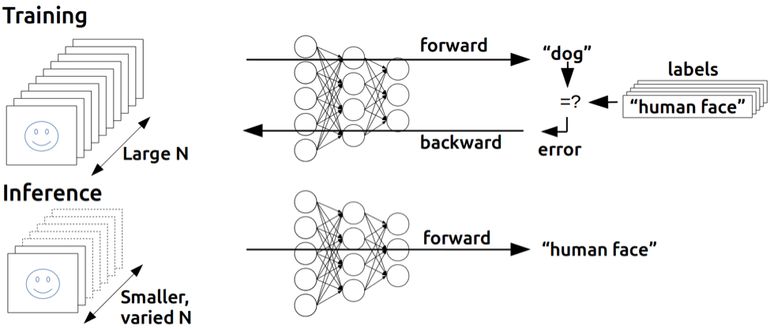


**Hình ảnh: Nick Heath / ZDNet**

Tất cả các mạng nơ-ron đều có một lớp đầu vào, trong đó dữ liệu ban đầu được đưa vào và một lớp đầu ra, tạo ra dự đoán cuối cùng. Nhưng trong một mạng nơ-ron sâu, sẽ có nhiều “lớp tế bào” ẩn giữa các lớp đầu vào và đầu ra, mỗi lớp cho dữ liệu vào nhau. Do đó, thuật ngữ “Deep” trong “Deep Learning” và “mạng nơ-ron sâu”, nó liên quan đến số lượng lớn các lớp ẩn – thường lớn hơn ba – tại trung tâm của các mạng thần kinh này.

Có thể thấy, đầu ra của một lớp là đầu vào của lớp tiếp theo trong mạng, với dữ liệu chảy qua mạng từ đầu vào đến đầu ra.

Quá trình học tập được thực hiện bằng cách mạng có thể thay đổi tầm quan trọng của các liên kết giữa các nơ-ron trong mỗi lớp. Mỗi liên kết có một giá trị đính kèm được gọi là trọng số, nó sẽ sửa đổi giá trị được tạo ra bởi một nơron khi nó truyền từ lớp này sang lớp kế tiếp. Bằng cách thay đổi giá trị của các trọng số này và một giá trị liên quan được gọi là sai lệch, có thể nhấn mạnh hoặc làm giảm tầm quan trọng của các liên kết giữa các nơ-ron trong mạng.



**Một minh họa về cấu trúc của một mạng lưới thần kinh và cách đào tạo hoạt động.**

Mô hình học được các liên kết giữa các nơ-ron rất quan trọng trong việc đưa ra dự đoán thành công trong quá trình đào tạo. Ở mỗi bước trong quá trình đào tạo, mạng sẽ sử dụng một hàm toán học để xác định mức độ chính xác của dự đoán mới nhất của nó so với dự kiến. Hàm này tạo ra một loạt các giá trị lỗi, do đó hệ thống có thể sử dụng để tính toán cách mô hình nên cập nhật giá trị của các trọng số được gắn vào mỗi liên kết, với mục đích cuối cùng là cải thiện độ chính xác của các dự đoán của mạng. Mức độ mà các giá trị này sẽ được thay đổi được tính bởi một chức năng tối ưu hóa, chẳng hạn như giảm độ dốc và những thay đổi đó được đẩy lùi trên toàn mạng vào cuối mỗi chu kỳ đào tạo trong một bước gọi là lan truyền ngược.

Trải qua nhiều, rất nhiều chu kỳ đào tạo và với sự trợ giúp của việc điều chỉnh tham số thủ công không thường xuyên, mạng sẽ tiếp tục nue để tạo dự đoán tốt hơn và tốt hơn cho đến khi nó đạt gần với độ chính xác cao nhất.

Về cơ bản, Deep Learning cho phép Machine Learning giải quyết một loạt các vấn đề phức tạp mới – chẳng hạn như nhận dạng hình ảnh, ngôn ngữ và lời nói – bằng cách cho phép máy móc tìm hiểu cách các tính năng trong dữ liệu kết hợp thành các dạng trừu tượng ngày càng cao hơn.

Trong bài toán nhận dạng ảnh, Deep Learning được sử dụng với các nơ-ron là các pixels hay tập hợp các pixels được đưa qua mạng CNN mà em sẽ đề cập ở phần tiếp theo. Kết quả cuối cùng sẽ dự đoán được xác suất tương ứng của từng nhãn trong ảnh.

Source: https://nordiccoder.com/blog/deep-learning-la-gi/

## Convolutional Neural Network (CNN)

Convolution là một cửa sổ trượt (Sliding Windows) trên một ma trận như mô tả hình dưới:

A picture containing shape

Description automatically generated

Các convolutional layer có các parameter(kernel) đã được học để tự điều chỉnh lấy ra những thông tin chính xác nhất mà không cần chọn các feature.

Trong hình ảnh ví dụ trên, ma trận bên trái là một hình ảnh trắng đen được số hóa. Ma trận có kích thước 5×5 và mỗi điểm ảnh có giá trị 1 hoặc 0 là giao điểm của dòng và cột.

Convolution hay tích chập là nhân từng phần tử trong ma trận 3. Sliding Window hay còn gọi là kernel, filter hoặc feature detect là một ma trận có kích thước nhỏ như trong ví dụ trên là 3×3.

Convolution hay tích chập là nhân từng phần tử bên trong ma trận 3×3 với ma trận bên trái. Kết quả được một ma trận gọi là Convoled feature được sinh ra từ việc nhận ma trận Filter với ma trận ảnh 5×5 bên trái.

Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trọng số trong các node. Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.

Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.

Mô hình này gọi là mạng kết nối đầy đủ (fully connected layer) hay mạng toàn vẹn (affine layer). Còn trong mô hình CNNs thì ngược lại. Các layer liên kết được với nhau thông qua cơ chế convolution.

Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Như vậy mỗi neuron ở lớp kế tiếp sinh ra từ kết quả của filter áp đặt lên một vùng ảnh cục bộ của neuron trước đó.

Mỗi một lớp được sử dụng các filter khác nhau thông thường có hàng trăm hàng nghìn filter như vậy và kết hợp kết quả của chúng lại. Ngoài ra có một số layer khác như pooling/subsampling layer dùng để chắt lọc lại các thông tin hữu ích hơn (loại bỏ các thông tin nhiễu).

Trong quá trình huấn luyện mạng (traning) CNN tự động học các giá trị qua các lớp filter dựa vào cách thức mà bạn thực hiện. Ví dụ trong tác vụ phân lớp ảnh, CNNs sẽ cố gắng tìm ra thông số tối ưu cho các filter tương ứng theo thứ tự raw pixel > edges > shapes > facial > high-level features. Layer cuối cùng được dùng để phân lớp ảnh.

Chart

Description automatically generated

Trong mô hình CNN có 2 khía cạnh cần quan tâm là **tính bất biến** (Location Invariance) và **tính kết hợp** (Compositionality). Với cùng một đối tượng, nếu đối tượng này được chiếu theo các gốc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể.

Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling). Tính kết hợp cục bộ cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter.

Đó là lý do tại sao CNNs cho ra mô hình với độ chính xác rất cao. Cũng giống như cách con người nhận biết các vật thể trong tự nhiên.

Mạng CNN đặc biệt có lợi trong việc xử lý hình ảnh bởi tính không gian của nó. Với bài toán phân loại ảnh, mạng CNN được sử dụng phổ biến bởi tính đơn giản và hiệu quả. Cụ thể là trong bài toán phân loại video theo từng khung hình mà em đang giải quyết, em có thể sử dụng các mạng CNN đã được huấn luyện từ trước để làm lớp trích chọn đặc trưng, sau đó huấn luyện thêm các lớp dense ở cuối để nhận được kết quả phù hợp với bài toán. Đây còn gọi là quá trình Transfer Learning. Sau khi train được model hội tụ, em tiếp tục huấn luyện lại một số lớp của mạng trích chọn đặc trưng để thu được kết quả tốt hơn. Đây là quá trình Fine-tune.

Source: https://topdev.vn/blog/thuat-toan-cnn-convolutional-neural-network/

## Tensorflow

Với sự bùng nổ của lĩnh vực Trí Tuệ Nhận Tạo - AI trong thập kỷ vừa qua, machine learning và deep learning rõ ràng cũng phát triển theo cùng. Và ở thời điểm hiện tại, Tensorflow chính là thư viện mã nguồn mở cho machine learning nổi tiếng nhất thế giới, được phát triển bởi các nhà nghiên cứu từ Google. Việc hỗ trợ mạnh mẽ các phép toán học để tính toán trong machine learning và deep learning đã giúp việc tiếp cận các bài toán trở nên đơn giản, nhanh chóng và tiện lợi hơn nhiều.

Các hàm được dựng sẵn trong thư viện cho từng bài toán cho phép TensorFlow xây dựng được nhiều neural network. Nó còn cho phép bạn tính toán song song trên nhiều máy tính khác nhau, thậm chí trên nhiều CPU, GPU trong cùng 1 máy hay tạo ra các dataflow graph – đồ thị luồng dữ liệu để dựng nên các model. Nếu bạn muốn chọn con đường sự nghiệp trong lĩnh vực A.I. này, nắm rõ những điều cơ bản của TensorFlow thực sự rất quan trọng.

Được viết bằng C++ và thao tác interface bằng Python nên phần performance của TensorFlow cực kỳ tốt. Đối tượng sử dụng nó cũng đa dạng không kém: từ các nhà nghiên cứu, nhà khoa học dữ liệu và dĩ nhiên không thể thiếu các lập trình viên.

Kiến trúc Tensorflow hoạt động được chia thành 3 phần: Tiền xử lý dữ liệu, dựng model, train và ước tính model.

Với bài toán là phân loại ảnh dựa trên các nhãn, em sử dụng Tensorflow để compile, train và fine-tune model dựa trên mạng InceptionV3 với số epoch là 100, số step trong mỗi epoch là 500 và sử dụng metrics là accuracy. Kết quả thu được một model có khả năng gán nhãn ảnh với các nhãn tương ứng.

Từ model thu được, em tiếp tục sử dụng Tensorflow để dự đoán nhãn của các ảnh mới dùng để lọc video theo từng frame.

Source: https://topdev.vn/blog/tensorflow-la-gi/

## OpenCV

OpenCV (Open Computer Vision) là một thư viện mã nguồn mở hàng đầu cho xử lý về thị giác máy tính, machine learning, xử lý ảnh. OpenCV đươc viết bằng C/C++, vì vậy có tốc độ tính toán rất nhanh, có thể sử dụng với các ứng dụng liên quan đến thời gian thực. Opencv có các interface cho C/C++, Python Java vì vậy hỗ trợ được cho Window, Linux, MacOs lẫn Android, iOS OpenCV có cộng đồng hơn 47 nghìn người dùng và số lượng download vượt quá 6 triệu lần.

Opencv có rất nhiều ứng dụng như: Nhận dạng ảnh, xử lý hình ảnh, phục hồi hình ảnh/video, thực tế ảo và các ứng dụng khác.

Trong bài toán phân loại video, em chủ yếu sử dụng OpenCv để xử lý dữ liệu đầu vào, cắt video thành từng frame ảnh và resize hình ảnh, chuyển về dạng mảng.

## Django

Django là một web framework miễn phí mã nguồn mở được viết bằng Python. Django sử dụng mô hình Model-View-Template (MVT). Mô hình MVT được sử dụng trong khi tạo một ứng dụng với Tương tác người dùng. Mô hình này bao gồm code HTML với Django Templage Language (DTL). Controller là mã được viết để kiểm soát sự tương tác giữa Model và View và Django tự động xử lý được công việc này. Bất cứ khi nào người dùng gửi request, nó xử lý request của người dùng đó bằng Model, View và Template. Nó hoạt động như một Controller để kiểm tra xem nó có khả dụng hay không bằng cách ánh xạ URL và nếu URL ánh xạ thành công thì View sẽ bắt đầu tương tác với Model và gửi lại Template cho người dùng dưới dạng respone.

Mục tiêu chính của Django là đơn giản hóa việc tạo các website phức tạp có sử dụng cơ sở dữ liệu. Django tập trung vào tính năng “có thể tái sử dụng” và “có thể tự chạy” của các component, tính năng phát triển nhanh, không làm lại những gì đã làm.

Với bài toán không quá phức tạp và việc tích hợp model deep learning trong phân loại video, em chọn ngôn ngữ lập trình là Python và framework Django để đơn giản hóa công việc lập trình web và tập trung vào bài toán xử lý nhận dạng video.

Source: https://niithanoi.edu.vn/django-la-gi.html

# Phát triển và triển khai ứng dụng

## Thiết kế kiến trúc

### Lựa chọn kiến trúc phần mềm

Vì sử dụng Framework Django nên kiến trúc em sử dụng sẽ là mô hình MVT. Mô hình MVT là biến thể của mô hình MVC. MVT là viết tắt của Model-View-Template. Về bản chất MVT và MVC đều giống nhau, đều chia ứng dụng ra thành ba thành phần: phần xử lý logic, phần hiển thị và phần thao tác với cơ sở dữ liệu.

A screenshot of a computer

Description automatically generated with medium confidence

Sự khác biệt của mô hình MVT và MVC được minh họa trong bảng sau:

|  |  |  |
| --- | --- | --- |
| **Mô hình MVC** | **Mô hình MVT** | **Chức năng** |
| Model | Model | Làm việc với cơ sở dữ liệu |
| View | Template | Hiển thị giao diện |
| Controller | View | Xử lý logic |

Template nghĩa là bản mẫu, là công cụ để tạo ra các giao diện web. Trong một bản mẫu sẽ có những phần cố định (static) và những phần được bổ sung sau (dynamic). Django có hai công cụ để tạo các template là DTL và Jinja2. Để tạo ra các template, bạn sẽ sử dụng cú pháp của một trong hai công cụ là DTL hoặc Jinja2, sau đó chuyển template cho template engine để chuyển sang mã HTML.

Nhiệm vụ của Model là thực hiện quản lý dữ liệu. Nghĩa là Model sẽ kiểm soát mọi sự thay đổi của dữ liệu cũng như việc giao tiếp CRUD giữa BE và Database.

Mục đích của View là xử lý logic nghiệp vụ thông qua các url được định nghĩa. Khi có một request từ client, View sẽ là trung tâm để thực hiện các yêu cầu của request đó và sẽ trả về (response) dữ liệu theo yêu cầu.

Đối với trang Web em thiết kế, do quy mô nhỏ gọn nên em sử dụng sẵn Template Engine của Django và cơ sở dữ liệu nhúng Sqlite3. Ưu điểm của chúng là được hỗ trợ sẵn bởi Django và không cần tùy chỉnh quá nhiều.

### Thiết kế tổng quan

Sinh viên vẽ biểu đồ gói UML (UML package diagram), nêu rõ sự phụ thuộc giữa các gói (package). SV cần vẽ các gói sao cho chúng được phân theo các tầng rõ ràng, không được sắp đặt package lộn xộn trong hình vẽ. Sinh viên chú ý các quy tắc thiết kế (Các gói không phụ thuộc lẫn nhau, gói tầng dưới không phụ thuộc gói tầng trên, không phụ thuộc bỏ qua tầng, v.v.) và cần giải thích sơ lược về mục đích/nhiệm vụ của từng package. SV tham khảo ví dụ minh họa trong Hình 1.

Diagram

Description automatically generated

**Hình 1** Ví dụ biểu đồ phụ thuộc gói

### Thiết kế chi tiết gói

Sinh viên thiết kế và lần lượt vẽ biểu đồ thiết kế cho từng package, hoặc một nhóm các package liên quan để giải quyết một vấn đề gì đó. Khi vẽ thiết kế gói, sinh viên chỉ cần đưa tên lớp, không cần chỉ ra các thành viên phương thức và thuộc tính. SV tham khảo ví dụ minh họa trong Hình 2.

Sinh viên cần vẽ rõ ràng quan hệ giữa các lớp trong biểu đồ. Các quan hệ bao gồm: phụ thuộc (dependency), kết hợp (association), kết tập (aggregation), hợp thành (composition), kế thừa (inheritance), và thực thi (implementation). Các quan hệ này đều đã được minh họa trong Hình 2.

Sau khi vẽ hình minh họa, sinh viên cần giải thích ngắn gọn về thiết kế của mình.

Diagram

Description automatically generated

**Hình 2** Ví dụ thiết kế gói

## Thiết kế chi tiết

### Thiết kế giao diện

Phần này có độ dài từ hai đến ba trang. Sinh viên đặc tả thông tin về màn hình mà ứng dụng của mình hướng tới, bao gồm độ phân giải màn hình, kích thước màn hình, số lượng màu sắc hỗ trợ, v.v. Tiếp đến, sinh viên đưa ra các thống nhất/chuẩn hóa của mình khi thiết kế giao diện như thiết kế nút, điều khiển, vị trí hiển thị thông điệp phản hồi, phối màu, v.v. Sau cùng sinh viên đưa ra một số hình ảnh minh họa thiết kế giao diện cho các chức năng quan trọng nhất. Lưu ý, sinh viên không nhầm lẫn giao diện thiết kế với giao diện của sản phẩm sau cùng.

### Thiết kế lớp

Phần này có độ dài từ ba đến bốn trang. Sinh viên trình bày thiết kế chi tiết các thuộc tính và phương thức cho một số lớp chủ đạo/quan trọng nhất của ứng dụng (từ 2-4 lớp). Thiết kế chi tiết cho các lớp khác, nếu muốn trình bày, sinh viên đưa vào phần phụ lục.

Để minh họa thiết kế lớp, sinh viên thiết kế luồng truyền thông điệp giữa các đối tượng tham gia cho 2 đến 3 use case quan trọng nào đó bằng biểu đồ trình tự (hoặc biểu đồ giao tiếp).

### Thiết kế cơ sở dữ liệu

Phần này có độ dài từ hai đến bốn trang. Sinh viên thiết kế, vẽ và giải thích biểu đồ thực thể liên kết (E-R diagram). Từ đó, sinh viên thiết kế cơ sở dữ liệu tùy theo hệ quản trị cơ sở dữ liệu mà mình sử dụng (SQL, NoSQL, Firebase, v.v.)

## Xây dựng ứng dụng

### Thư viện và công cụ sử dụng

Sinh viên liệt kê các công cụ, ngôn ngữ lập trình, API, thư viện, IDE, công cụ kiểm thử, v.v. mà mình sử dụng để phát triển ứng dụng. Mỗi công cụ phải được chỉ rõ phiên bản sử dụng. SV nên kẻ bảng mô tả tương tự như Bảng 1. Nếu có nhiều nội dung trình bày, sinh viên cần xoay ngang bảng.

**Bảng 1** Danh sách thư viện và công cụ sử dụng

|  |  |  |
| --- | --- | --- |
| **Mục đích** | **Công cụ** | **Địa chỉ URL** |
| IDE lập trình | Eclipse Oxygen 64 bit | http://www.eclipse.org/ |
| v.v. | v.v. | v.v. |

### Kết quả đạt được

Sinh viên trước tiên mô tả kết quả đạt được của mình là gì, ví dụ như các sản phẩm được đóng gói là gì, bao gồm những thành phần nào, ý nghĩa, vai trò?

Sinh viên cần thống kê các thông tin về ứng dụng của mình như: số dòng code, số lớp, số gói, dung lượng toàn bộ mã nguồn, dung lượng của từng sản phẩm đóng gói, v.v. Tương tự như phần liệt kê về công cụ sử dụng, sinh viên cũng nên dùng bảng để mô tả phần thông tin thống kê này.

### Minh hoạ các chức năng chính

Sinh viên lựa chọn và đưa ra màn hình cho các chức năng chính, quan trọng, và thú vị nhất. Mỗi giao diện cần phải có lời giải thích ngắn gọn. Khi giải thích, sinh viên có thể kết hợp với các chú thích ở trong hình ảnh giao diện.

## Kiểm thử

Phần này có độ dài từ hai đến ba trang. Sinh viên thiết kế các trường hợp kiểm thử cho hai đến ba chức năng quan trọng nhất. Sinh viên cần chỉ rõ các kỹ thuật kiểm thử đã sử dụng. Chi tiết các trường hợp kiểm thử khác, nếu muốn trình bày, sinh viên đưa vào phần phụ lục.

Sinh viên sau cùng tổng kết về số lượng các trường hợp kiểm thử và kết quả kiểm thử. Sinh viên cần phân tích lý do nếu kết quả kiểm thử không đạt.

## Triển khai

Sinh viên trình bày mô hình và/hoặc cách thức triển khai thử nghiệm/thực tế. Ứng dụng của sinh viên được triển khai trên server/thiết bị gì, cấu hình như thế nào. Kết quả triển khai thử nghiệm nếu có (số lượng người dùng, số lượng truy cập, thời gian phản hồi, phản hồi người dùng, khả năng chịu tải, các thống kê, v.v.)

# Các giải pháp và đóng góp nổi bật

Chương này có độ dài tối thiểu 5 trang, tối đa không giới hạn. Sinh viên cần trình bày tất cả những nội dung đóng góp mà mình thấy tâm đắc nhất trong suốt quá trình làm ĐATN. Đó có thể là một loạt các vấn đề khó khăn mà sinh viên đã từng bước giải quyết được, là giải thuật cho một bài toán cụ thể, là giải pháp tổng quát cho một lớp bài toán, hoặc là mô hình/kiến trúc hữu hiệu nào đó được sinh viên thiết kế.

Chương này là **cơ sở quan trọng** để các thầy cô đánh giá sinh viên. Vì vậy, sinh viên cần phát huy tính sáng tạo, khả năng phân tích, phản biện, lập luận, tổng quát hóa vấn đề và tập trung viết cho thật tốt.

Mỗi giải pháp hoặc đóng góp của sinh viên cần được trình bày trong một mục độc lập bao gồm ba mục con: (i) dẫn dắt/giới thiệu về bài toán/vấn đề, (ii) giải pháp, và (iii) kết quả đạt được (nếu có).

Sinh viên lưu ý **không trình bày lặp lại nội dung**. Những nội dung đã trình bày chi tiết trong các chương trước không được trình bày lại trong chương này. Vì vậy, với nội dung hay, mang tính đóng góp/giải pháp, sinh viên chỉ nên tóm lược/mô tả sơ bộ trong các chương trước, đồng thời tạo tham chiếu chéo tới đề mục tương ứng trong Chương 5 này. Chi tiết thông tin về đóng góp/giải pháp được trình bày trong mục đó.

Ví dụ, trong Chương 4, sinh viên có thiết kế được kiến trúc đáng lưu ý gì đó, là sự kết hợp của các kiến trúc MVC, MVP, SOA, v.v. Khi đó, sinh viên sẽ chỉ mô tả ngắn gọn kiến trúc đó ở Chương 4, rồi thêm các câu có dạng: “Chi tiết về kiến trúc này sẽ được trình bày trong phần 5.1”.

# Kết luận và hướng phát triển

## Kết luận

Sinh viên so sánh kết quả nghiên cứu hoặc sản phẩm của mình với các nghiên cứu hoặc sản phẩm tương tự.

Sinh viên phân tích trong suốt quá trình thực hiện ĐATN, mình đã làm được gì, chưa làm được gì, các đóng góp nổi bật là gì, và tổng hợp những bài học kinh nghiệm rút ra nếu có.

## Hướng phát triển

Trong phần này, sinh viên trình bày định hướng công việc trong tương lai để hoàn thiện sản phẩm hoặc nghiên cứu của mình.

Trước tiên, sinh viên trình bày các công việc cần thiết để hoàn thiện các chức năng/nhiệm vụ đã làm. Sau đó sinh viên phân tích các hướng đi mới cho phép cải thiện và nâng cấp các chức năng/nhiệm vụ đã làm.

# Tài liệu tham khảo

Lưu ý: Sinh viên không được đưa bài giảng/slide, các trang Wikipedia, hoặc các trang web thông thường làm tài liệu tham khảo.

Một trang web được phép dùng làm tài liệu tham khảo **chỉ khi** nó là công bố chính thống của cá nhân hoặc tổ chức nào đó. Ví dụ, trang web đặc tả ngôn ngữ XML của tổ chức W3C https://www.w3.org/TR/2008/REC-xml-20081126/ là TLTK hợp lệ.

Có năm loại tài liệu tham khảo mà sinh viên phải tuân thủ đúng quy định về cách thức liệt kê thông tin như sau. Lưu ý: các phần văn bản trong cặp dấu < > dưới đây chỉ là hướng dẫn khai báo cho từng loại tài liệu tham khảo; sinh viên cần xóa các phần văn bản này trong ĐATN của mình.

<**Bài báo đăng trên tạp chí khoa học**: Tên tác giả, tên bài báo, tên tạp chí, volume, từ trang đến trang (nếu có), nhà xuất bản, năm xuất bản>

1. Hovy E. H., Automated Discourse Generation Using Discourse Structure Relations, Artificial Intelligence, Elsevier Science Publishers, 63: 341-385, 1993.

<**Sách**: Tên tác giả, tên sách, volume (nếu có), lần tái bản (nếu có), nhà xuất bản, năm xuất bản>

1. Peterson L. L. and Davie B. S., Computer Networks: A Systems Approach, 2nd ed., Mogran-Kaufmann, 1999.
2. Nguyễn Thúc Hải, Mạng máy tính và các hệ thống mở, Nhà xuất bản giáo dục, 1999.

<**Tập san Báo cáo Hội nghị Khoa học**: Tên tác giả, tên báo cáo, tên hội nghị, ngày (nếu có), địa điểm hội nghị, năm xuất bản>

1. Poesio M. and Di Eugenio B., Discourse Structure and Anaphoric Accessibility, In Proc. of the ESSLLI Workshop on Information Structure, Discourse Structure and Discourse Semantics, Helsinki, 2001.

<**Đồ án tốt nghiệp, Luận văn Thạc sĩ, Tiến sĩ**: Tên tác giả, tên đồ án/luận văn, loại đồ án/luận văn, tên trường, địa điểm, năm xuất bản>

1. Knott D., A Data-Driven Methodology for Motivating a Set of Coherence Relations, Ph.D. Thesis, University of Edinburgh, UK, 1996.

<**Tài liệu tham khảo từ Internet**: Tên tác giả (nếu có), tựa đề, cơ quan (nếu có), địa chỉ trang web, thời gian lần cuối truy cập trang web>

1. Berners-Lee T., Hypertext Transfer Protocol (HTTP), CERN, ftp:/info.cern.ch/pub/www/doc/http-spec.txt.Z, last visited May 2010.
2. Princeton University, WordNet, http://www.cogsci.princeton.edu/~wn/index.shtml, last visited May 2010.

# Phụ lục

Phần phụ lục là không bắt buộc. Nếu sinh viên không có nhu cầu trình bày thêm, có thể xóa bỏ phần này. Lưu ý là phần phụ lục chỉ được đánh chỉ mục đến cấp 2, sinh viên không được phép chia nhỏ hơn nữa.

Hướng dẫn viết đồ án tốt nghiệp

Quy định chung

Dưới đây là một số quy định và hướng dẫn viết đồ án tốt nghiệp mà bắt buộc sinh viên phải **đọc kỹ và tuân thủ nghiêm ngặt**.

Sinh viên cần đảm bảo tính thống nhất toàn báo cáo (font chữ, căn dòng hai bên, hình ảnh, bảng, margin trang, đánh số trang, v.v.). Để làm được như vậy, sinh viên chỉ cần sử dụng các định dạng theo đúng template ĐATN này. Khi paste nội dung văn bản từ tài liệu khác của mình, sinh viên cần chọn kiểu Copy là “Text Only” để định dạng văn bản của template không bị phá vỡ/vi phạm.

Tuyệt đối cấm sinh viên đạo văn. Sinh viên cần ghi rõ nguồn cho tất cả những gì không tự mình viết/vẽ lên, bao gồm các câu trích dẫn, các hình ảnh, bảng biểu, v.v. Khi bị phát hiện, sinh viên sẽ **không được phép bảo vệ ĐATN**.

Tất cả các hình vẽ, bảng biểu, công thức, và tài liệu tham khảo trong ĐATN nhất thiết phải được SV giải thích và tham chiếu tới ít nhất một lần. Không chấp nhận các trường hợp sinh viên đưa ra hình ảnh, bảng biểu tùy hứng và không có lời mô tả/giải thích nào.

Sinh viên tuyệt đối không trình bày ĐATN theo kiểu viết ý hoặc gạch đầu dòng. ĐATN không phải là một slide thuyết trình; khi người đọc không hiểu sẽ không có ai giải thích hộ. Sinh viên cần viết thành các đoạn văn và phân tích, diễn giải đầy đủ, rõ ràng. Câu văn cần đúng ngữ pháp, đầy đủ chủ ngữ, vị ngữ và các thành phần câu.

Khi thực sự cần liệt kê, sinh viên nên liệt kê theo phong cách khoa học với các ký tự La Mã. Ví dụ, nhiều sinh viên luôn cảm thấy hối hận vì (i) chưa cố gắng hết mình, (ii) chưa sắp xếp thời gian học/chơi một cách hợp lý, (iii) chưa tìm được người yêu để chia sẻ quãng đời sinh viên vất vả, và (iv) viết ĐATN một cách cẩu thả.

Trong một số trường hợp nhất thiết phải dùng các bullet để liệt kê, sinh viên cần thống nhất Style cho toàn bộ các bullet các cấp mà mình sử dụng đến trong báo cáo. Nếu dùng bullet cấp 1 là hình tròn đen, toàn bộ báo cáo cần thống nhất cách dùng như vậy; ví dụ như sau:

* Đây là mục 1 – Thực sự không còn cách nào khác tôi mới dùng đến việc bullet trong báo cáo.
* Đây là mục 2 – Nghĩ lại thì tôi có thể không cần dùng bullet cũng được. Nên tôi sẽ xóa bullet và tổ chức lại hai mục này trong báo cáo của mình cho khoa học hơn. Tôi muốn thầy cô và người đọc cảm nhận được tâm huyết của tôi trong từng trang báo cáo ĐATN.

Ngành học

Sinh viên lưu ý viết đúng ngành/chuyên ngành trên bìa và trên gáy theo đúng quy định của Trường. Ngành học hay chuyên ngành học phụ thuộc vào ngành học mà sinh viên đăng ký. Sinh viên có thể đăng nhập trên trang quản lý học tập của mình để xem lại chính xác ngành học của mình.

Một số ví dụ sinh viên có thể tham khảo dưới đây, trong trường hợp có chuyên ngành thì sinh viên không cần ghi chuyên ngành:

* Đối với kỹ sư chính quy:
  + Từ K61 trở về trước: Ngành Kỹ thuật phần mềm
  + Từ K62 trở về sau: Ngành Khoa học máy tính
* Đối với cử nhân:
  + Ngành Công nghệ thông tin
* Đối với chương trình EliteTech:
  + Chương trình Việt Nhật/KSTN: Ngành Công nghệ thông tin
  + Chương trình ICT Global: Ngành Information Technology
  + Chương trình DS&AI: Ngành Khoa học dữ liệu
* Đối với kỹ sư chương trình đào tạo quốc tế SIE
  + Chương trình LTU: Ngành Công nghệ thông tin/Information Technology
  + Chương trình VUW: Ngành Kỹ thuật Phần mềm
  + Chương trình G-INP: Ngành Hệ thống thông tin

Tạo đề mục

Đề mục giúp tạo bố cục cho tài liệu. Để các tính năng tự động – ví dụ tính năng cập nhật mục lục, hoặc tính năng tham chiếu chéo – của Word hoạt động được, sinh viên cần tuân thủ theo các style đã tạo trong tài liệu này. Để hiển thị các style này, sinh viên vào tab Home trong thanh Ribbon của Word.

Để tạo đề mục cấp 1, 2, 3, 4, 5, sinh viên gõ tiêu đề cho đề mục của mình rồi chọn các style là Heading 1, 2, 3, 4, 5 tương ứng. Sinh viên hạn chế dùng tới đề mục cấp 4, và phải trong trường hợp thực sự cần thiết mới dùng đến đề mục cấp 5.

Phần phụ lục chỉ cho phép có hai cấp tiêu đề. Hai style tương ứng với hai cấp này là “Heading 7, Phụ lục cấp 1” và “Heading 8, Phụ lục cấp 2”.

Bảng biểu

Sinh viên lưu ý không để bảng tràn ra lề (margin) trên, dưới, trái hoặc phải của trang. Do không gian nhỏ hẹp, bảng nên có font là 12pt hoặc nhỏ hơn. Độ dãn dòng của bảng nên là 1 line. Căn lề bảng là căn giữa, nhưng nội dung văn bản trong bảng nên được căn lề trái.

Sinh viên có thể viết tắt các từ trong bảng để tiết kiệm không gian nhưng phải giải thích các từ viết tắt này ở phần Chú thích bảng. Ví dụ áp dụng được minh họa trong Bảng 2.

**Bảng 2** Ví dụ sử dụng bảng

Chú thích: Y: Year; RS: Risk Set; G: Graduated; AB: Academically Excluded;C: Censored; HRG: Hazard Ratio – Graduated

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Y** | **RS** | **G** | **AE** | **C** | **HRG**  **(%)** |
| **1** | 11.959 | 0 | 725 | 1619 | 0,0 |
| **2** | 10.457 | 0 | 474 | 1513 | 0,0 |
| **3** | 7.365 | 1213 | 335 | 966 | 16,7 |
| **4** | 900 | 599 | 145 | 405 | 55,3 |

Sinh viên được tùy ý lựa chọn định dạng (template) cho các bảng trong ĐATN của mình, nhưng phải đảm bảo tính nhất quán trên toàn ĐATN. Template bảng phải đảm bảo phần heading cột trong bảng có font in đậm và nổi bật (highlighted) hơn so với các nội dung khác trong bảng.

Vì bảng có thể kéo dài nhiều trang, tiêu đề của bảng nên để ở phần đầu của bảng. Sinh viên không thêm tiêu đề bảng bằng tay. Để thêm tiêu đề bảng tự động, sinh viên nhấn chuột phải vào bảng, chọn “Insert Caption”, chọn “Label” là “Bảng”, rồi nhấn nút “OK”. Sau đó, sinh viên nhập vào nội dung tiêu đề và căn chỉnh “Center” cho tiêu đề này. Lưu ý, sinh viên cần bôi đậm bằng tay cụm từ **Bảng n**. Kết quả thu được có dạng như sau “**Bảng 2** Ví dụ sử dụng bảng”. Sinh viên nên xoay ngang trang giấy trong trường hợp bảng có nhiều cột với nhiều nội dung văn bản.

Hình vẽ

Tương tự như bảng, sinh viên không được để hình vẽ tràn lề trang. Căn lề cho hình vẽ là căn giữa (Center). Cách thêm tiêu đề hình vẽ tương tự như cách thêm tiêu đề bảng, nhưng sinh viên chọn “Label” là “Hình” thay vì “Bảng”.

Tiêu đề hình vẽ phải đặt ở dưới hình vẽ. Nếu hình vẽ được copy từ trên mạng, sinh viên bắt buộc phải ghi rõ nguồn. Sinh viên nên thống nhất công cụ sử dụng và style cho hình vẽ trong toàn ĐATN. Các chi tiết trong hình vẽ phải được bố trí gọn gàng; chữ trong hình phải đảm bảo nhìn được rõ nét khi in báo cáo trên giấy A4. Khi resize ảnh, cần giữ nguyên tỷ lệ dài rộng, tránh làm méo hoặc vỡ hình. Ví dụ hình vẽ được minh họa trong Hình 3.

Diagram

Description automatically generated

**Hình 3** Ví dụ hình vẽ

Tài liệu tham khảo

Sinh viên cần hạn chế tối đa dùng trang Web làm tài liệu tham khảo. Chỉ chấp nhận trang Web làm TLTK khi trang đó là nơi công bố chính thức của tổ chức hoặc cá nhân nào đó. Ví dụ, trang web đặc tả ngôn ngữ XML của tổ chức W3C https://www.w3.org/TR/2008/REC-xml-20081126/ là TLTK hợp lệ.

Khi giới thiệu về công cụ, API, thư viện, hoặc nền tảng nào đó, sinh viên có thể đưa ra địa chỉ URL của các tiện ích này. Sinh viên lưu ý địa chỉ URL đó không phải là tài liệu tham khảo. Trong các trường hợp tương tự như vậy, sinh viên nên tạo “Footnote”. Sinh viên tạo “Footnote” bằng cách vào mục “References”, chọn “Insert Footnote”. Ví dụ tạo Footnote như sau: TensorFlow[[1]](#footnote-1) là nền tảng học máy mã nguồn mở đang được sử dụng rộng rãi hiện nay. Lưu ý: số Footnote phải đặt sát với từ được mô tả. Như trong ví dụ trên, số 1 được đặt ngay cạnh chữ TensorFlow (không có dấu cách).

Công thức toán học

Giống như bảng, hình vẽ, và tài liệu tham khảo, công thức toán học cần được đánh số, giải thích, và tham chiếu đầy đủ.

Để thêm tiêu đề (caption) cho công thức, sinh viên đặt con trỏ văn bản vào dòng văn bản dưới công thức. Sau đó vào menu “Insert” (lưu ý là menu Insert, không phải là tab Insert trong thanh Ribbon) và chọn mục “Caption”. Từ Popup Menu hiện ra, sinh viên chọn “Label” là “Công thức”, rồi nhấn nút “OK”. Sinh viên bôi đậm chữ **Công thức n**, và căn tiêu đề ra giữa. Công thức 1 là ví dụ mẫu cho sinh viên tham khảo.

**Công thức 1** Khai triển Newton

Tham chiếu chéo

Tham chiếu chéo (Cross-reference) là tiện ích hữu hiệu cho người viết báo cáo. Nó giúp tạo các liên kết tham chiếu (hyperlink) tới các hình ảnh, bảng biểu, tài liệu tham khảo, và các đề mục một cách tự động. Ví dụ, ngay trong câu này, một tham chiếu đã được tạo ra tới mục 2.1. Người đọc dễ dàng nhấp chuột vào liên kết 2.1 để ngay lập tức chuyển đến mục đó.

Để tạo tham chiếu chéo tới các đề mục (heading), sinh viên vào tab “References”, rồi tìm và nhấn chọn “Cross-refrence”. SV chọn “Refrence type” là “Heading” và chọn “Insert Reference to” là “Heading number (no context)”. Sau đó, SV chọn phần đề mục muốn tham chiếu rồi bấm “Insert”.

Để tạo tham chiếu chéo tới các hình vẽ, bảng biểu và công thức, sinh viên cũng vào tab “References” và chọn “Cross-refrence”. SV chọn “Refrence type” là “Hình”, “Bảng”, hoặc “Công thức’. SV chọn “Insert Reference to” là “Only label and number”, trỏ đến phần muốn tham chiếu rồi bấm “Insert”. Nếu font chữ trong liên kết tham chiếu tạo ra được in đậm (bold), SV chuyển về dạng thường cho chuẩn tắc.

Thực hiện tương tự các bước mô tả ở trên, sinh viên có thể tạo tham chiếu chéo tới các tài liệu tham khảo. Sinh viên chọn “Refrence type” là “Numbered item”, chọn “Insert Reference to” là “Paragraph number”, trỏ đến phần tài liệu tham khảo muốn tham chiếu rồi bấm “Insert”. Ví dụ, tham chiếu chéo tới tài liệu tham khảo [2], [3], [4] đã được tạo.

Cập nhật mục lục và tham chiếu chéo

Trong suốt quá trình viết ĐATN, sinh viên sẽ tạo ra nhiều xáo trộn như thay đổi vị trí hình và bảng, thay đổi thứ tự tài liệu tham khảo, thay đổi tên đề mục, v.v. Vì vậy, để hoàn thành ĐATN, sinh viên cần cập nhật lại các thành phần mục lục, danh sách hình ảnh/bảng/công thức và tham chiếu chéo được tạo tự động trong ĐATN của mình.

Để cập nhật các thành phần này, sinh viên bấm Ctrl+A để chọn toàn báo cáo, nhấn chuột phải và chọn “Update Field”, rồi liên tục chọn mục “Entire table” khi được hỏi trong hộp thoại Popup Menu. Sau đó sinh viên search chữ “Error” trên toàn báo cáo để kiểm tra xem có lỗi đánh chỉ mục hoặc lỗi tham chiếu nào không. Đồng thời, sinh viên nên tự soát lại bằng tay toàn bộ nội dung quyển ĐATN để tránh mọi sai sót.

In quyển đồ án tốt nghiệp

Do hiện nay có nhiều phiên bản Word cho nhiều nền tảng máy tính, sinh viên nhất thiết phải xuất ĐATN ra định dạng PDF rồi mang tới cửa hàng in ấn để tránh sai sót. Quyển ĐATN nên được in một mặt trên các trang giấy A4.

Khi đóng quyển, bìa trước và bìa sau là giấy liền khổ. **Sử dụng keo nhiệt để dán gáy khi đóng quyển thay vì sử dụng băng dính và dập ghim** (Xem Hình 4).



**Hình 4** Quy cách đóng quyển

Phần gáy ĐATN được đóng theo quy cách như Hình 5, với các thông tin cần ghi bao gồm *Kỳ làm ĐATN - Ngành đào tạo - Họ và tên sinh viên - Mã số sinh viên*



**Hình 5** Quy cách ghi chữ phần gáy

Sinh viên có thể in một mặt hoặc hai mặt. Template này mặc định ở chế độ in một mặt. Để chuyển sang chế độ in hai mặt, sinh viên phải thiết lập lại page layout như sau. Sinh viên vào tab Page Layout, chọn Margins, chọn Custom Margin, trong popup hiện ra, ở mục Multiple pages, chọn Mirror margins, và nhấn OK (Xem **Hình 6**). Do trang đầu tiên là trang bìa của đồ án, nên khi in chế độ hai mặt có thể cần chèn (insert) 1 trang trắng sau trang bìa để đảm bảo trang “Lời cam kết” sẽ là trang lẻ. Để tránh sai sót gây lãng phí, sinh viên cần kiểm tra kỹ trang chẵn lẽ trước khi in.

Sinh viên lưu ý, để đóng quyển theo đúng quy cách với keo nhiệt và chữ in trên gáy, độ dày tối thiểu cho tổng các trang là 0.5cm. Sinh viên tự cân nhắc in một mặt hoặc hai mặt.

Graphical user interface, application, Word

Description automatically generated

**Hình 6** Hướng dẫn thiết lập in hai mặt

Đặc tả use case

Nếu trong nội dung chính không đủ không gian cho các use case khác (ngoài các use case nghiệp vụ chính) thì đặc tả thêm cho các use case đó ở đây.

Đặc tả use case “Thống kê tình hình mượn sách”

Cách trình bày tương tự như phần hướng dẫn ở mục 2.3.

Đặc tả use case “Đăng ký làm thẻ mượn”

Cách trình bày tương tự như phần hướng dẫn ở mục 2.3.

Công nghệ sử dụng

Công nghệ bảo mật dữ liệu

Công nghệ blockchain

Thiết kế gói

Thiết kế gói cho kiến trúc tổng quan

Thiết kế gói cho chức năng “Trả sách”

Thiết kế lớp

1. https://www.tensorflow.org/, lần truy cập cuối: 28/06/2018 [↑](#footnote-ref-1)