

TRƯỜNG ĐẠI HỌC VĂN LANG  
**ĐƠN VỊ: KHOA CÔNG NGHỆ THÔNG TIN**

**ĐỀ THI/ĐỀ BÀI, RUBRIC VÀ THANG ĐIỂM**  
**THI KẾT THÚC HỌC PHẦN**  
**Học kỳ 2, năm học 2023-2024**

**I. Thông tin chung**

Tên học phần:	Xử lý ngôn ngữ tự nhiên		
Mã học phần:	71ITAI40703	Số tín chỉ:	3
Mã nhóm lớp học phần:	232_71ITAI40703_01/232_71ITAI40703_02		
Hình thức thi: <b>Đồ án</b>	Thời gian làm bài:	<b>20</b>	phút
<input type="checkbox"/> Cá nhân	<input checked="" type="checkbox"/> Nhóm		
<b>Quy cách đặt tên file</b>	<b>Mã SV_Ho và ten SV_XNNNTN_nhomX</b>		

Giảng viên nộp đề thi, đáp án bao gồm cả **Lần 1 và Lần 2 trước ngày 15/03/2024.**

**1. Format đề thi**

- Font: Times New Roman
- Size: 13
- Quy ước đặt tên file đề thi/đề bài:
- + **Mã học phần**\_Tên học phần\_Mã nhóm học phần\_TIEUL\_De 1

**2. Giao nhận đề thi**

Sau khi kiểm duyệt đề thi, đáp án/rubric. **Trưởng Khoa/Bộ môn** gửi đề thi, đáp án/rubric về Trung tâm Khảo thí qua email: [khaothivanlang@gmail.com](mailto:khaothivanlang@gmail.com) bao gồm file word và file pdf (**nén lại và đặt mật khẩu file nén**) và nhắn tin + họ tên người gửi qua số điện thoại **0918.01.03.09** (Phan Nhất Linh).

## II. Các yêu cầu của đề thi nhằm đáp ứng CLO

(Phần này phải phối hợp với thông tin từ đề cương chi tiết của học phần)

Ký hiệu CLO	Nội dung CLO	Hình thức đánh giá	Trọng số CLO trong thành phần đánh giá (%)	Câu hỏi thi số	Điểm số tối đa	Lấy dữ liệu đo lường mức đạt PLO/PI
(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>CLO 1</b>	Hiểu rõ các khái niệm trong XLNNTN, phân tích được các vấn đề của bài toán XLNNTN, lựa chọn được các mô hình/thuật toán áp dụng cho bài toán	Đồ án	20%	- Trình bày được mục tiêu, phạm vi và phương pháp thực hiện đề tài, lý do chọn mô hình/thuật toán phù hợp với đề tài	10	
<b>CLO 2</b>	Vận dụng được các mô hình học sâu trên các bài toán XLNNTN thực tế	Đồ án	40%	- Lựa chọn và giải thích được mô hình/thuật toán được chọn để hiện thực bài toán	10	
<b>CLO 3</b>	Sử dụng thành thạo thư viện Python trong để hiện thực được các mô hình học sâu trên các bài toán XLNNTN thực tế	Đồ án	20%	- Cài đặt các thuật toán hiện thực mô hình, giải thích được thuật toán, kết quả của bài toán	10	
<b>CLO 5</b>	Thể hiện tinh thần tự học, chủ động cập nhật các xu hướng công nghệ mới trong lĩnh vực công nghệ thông tin để nâng cao năng lực bản thân và ứng dụng vào công việc	Đồ án	20%	-Phân chia công việc trong nhóm rõ ràng, hợp lý. Phối hợp tốt với các thành viên trong nhóm. Trình bày báo cáo đầy đủ, thuyết phục, trả lời câu hỏi khi	10	

				báo cáo đầy đủ, chính xác		
--	--	--	--	---------------------------	--	--

### Chú thích các cột:

(1) Chỉ liệt kê các CLO được đánh giá bởi đề thi kết thúc học phần (tương ứng như đã mô tả trong đề cương chi tiết học phần). Lưu ý không đưa vào bảng này các CLO không dùng bài thi kết thúc học phần để đánh giá (có một số CLO được bố trí đánh giá bằng bài kiểm tra giữa kỳ, đánh giá qua dự án, đồ án trong quá trình học hay các hình thức đánh giá quá trình khác chứ không bố trí đánh giá bằng bài thi kết thúc học phần). Trường hợp một số CLO vừa được bố trí đánh giá quá trình hay giữa kỳ vừa được bố trí đánh giá kết thúc học phần thì vẫn đưa vào cột (1)

(2) Nêu nội dung của CLO tương ứng.

(3) Hình thức kiểm tra đánh giá có thể là: trắc nghiệm, tự luận, dự án, đồ án, vấn đáp, thực hành trên máy tính, thực hành phòng thí nghiệm, báo cáo, thuyết trình, ..., phù hợp với nội dung của CLO và mô tả trong đề cương chi tiết học phần.

(4) Trọng số mức độ quan trọng của từng CLO trong đề thi kết thúc học phần do giảng viên ra đề thi quy định (mang tính tương đối) trên cơ sở mức độ quan trọng của từng CLO. Đây là cơ sở để phân phối tỷ lệ % số điểm tối đa cho các câu hỏi thi dùng để đánh giá các CLO tương ứng, bảo đảm CLO quan trọng hơn thì được đánh giá với điểm số tối đa lớn hơn. Cột (4) dùng để hỗ trợ cho cột (6).

(5) Liệt kê các câu hỏi thi số (câu hỏi số ... hoặc từ câu hỏi số... đến câu hỏi số...) dùng để kiểm tra người học đạt các CLO tương ứng.

(6) Ghi điểm số tối đa cho mỗi câu hỏi hoặc phần thi.

(7) Trong trường hợp đây là học phần cốt lõi - sử dụng kết quả đánh giá CLO của hàng tương ứng trong bảng để đo lường đánh giá mức độ người học đạt được PLO/PI - cần liệt kê ký hiệu PLO/PI có liên quan vào hàng tương ứng. Trong đề cương chi tiết học phần cũng cần mô tả rõ CLO tương ứng của học phần này sẽ được sử dụng làm dữ liệu để đo lường đánh giá các PLO/PI. Trường hợp học phần không có CLO nào phục vụ việc đo lường đánh giá mức đạt PLO/PI thì để trống cột này.

### III. Nội dung đề bài

#### 1. Đề bài: XÂY DỰNG ỨNG DỤNG XỬ LÝ NGÔN NGỮ TỰ NHIÊN

- Mỗi nhóm sinh viên từ 1-4 bạn, đăng ký đề tài theo danh sách gợi ý hoặc có thể tự chọn đề tài phù hợp. Có thể chọn đề tài giống nhau nhưng sử dụng mô hình khác.

- Danh sách đề tài gợi ý:

STT	Tên đề tài	Mô tả	Lưu ý
1	Ứng dụng mô hình học máy để phân loại văn bản	- Tìm hiểu các mô hình sử dụng để phân loại văn bản. Chọn 1 mô hình. - Xây dựng chương trình demo minh họa (ví dụ: lọc email, phân loại bài báo theo các nhóm, tóm tắt văn bản) và đánh giá	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài
2	Phân tích cảm xúc trong văn bản sử dụng mô hình học sâu	- Tìm hiểu các mô hình sử dụng để phân tích cảm xúc	SV ghi rõ sử dụng mô hình gì khi

		trong văn bản. Chọn 1 mô hình. Xây dựng chương trình demo minh họa (ví dụ: đánh giá sản phẩm và dịch vụ: album âm nhạc, phim, sản phẩm mới của một nhãn hàng) . Đánh giá chương trình	đăng ký đề tài
3	Xây Dựng Hệ Thống Tìm Kiếm và Trích Xuất Tự Động Từ Văn Bản	Tìm hiểu các mô hình sử dụng để trích xuất thông tin. Chọn mô hình. Xây dựng ứng dụng demo (ví dụ: tóm tắt sách, tóm tắt văn bản, tóm tắt bài giảng, chatbot...). Đánh giá kết quả thực hiện	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài
4	Xây dựng ứng dụng phát hiện tin tức giả mạo	Tìm hiểu các mô hình phù hợp cho bài toán. Chọn 1 mô hình. Xây dựng ứng dụng demo (Có thể chọn xây dựng mô hình cho 1 lĩnh vực cụ thể: y tế, giáo dục...). Đánh giá kết quả thực hiện	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài
5	Xây dựng ứng dụng xác định sự tương đồng của văn bản	Tìm hiểu các mô hình phù hợp cho bài toán. Chọn 1 mô hình. Xây dựng ứng dụng demo. Đánh giá kết quả thực hiện	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài
6	Xây dựng ứng dụng chuyển văn bản thành tiếng nói	Tìm hiểu các mô hình phù hợp cho bài toán. Chọn 1 mô hình. Xây dựng ứng dụng demo. Đánh giá kết quả thực hiện	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài
7	Xây dựng chương trình dịch ngôn ngữ	Tìm hiểu các mô hình phù hợp cho bài toán. Chọn 1 mô hình. Xây dựng ứng dụng demo (Có thể chọn dịch từ 1 ngôn ngữ nguồn sang 1 ngữ đích hoặc dịch đa	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài

		ngôn ngữ). Đánh giá kết quả thực hiện	
8	Xây dựng ứng dụng chuyển âm thanh thành văn bản	Tìm hiểu các mô hình phù hợp cho bài toán. Chọn 1 mô hình. Xây dựng ứng dụng demo. Đánh giá kết quả thực hiện	SV ghi rõ sử dụng mô hình gì khi đăng ký đề tài

### **Mục tiêu của đồ án**

- Giúp sinh viên tìm hiểu về các công cụ, các mô hình xử lý ngôn ngữ tự nhiên
- Xây dựng ứng dụng xử lý ngôn ngữ tự nhiên, sử dụng các mô hình học sâu

### **Phạm vi đồ án**

- Xây dựng ứng dụng xử lý ngôn ngữ tự nhiên, sử dụng các mô hình học sâu

### **Thời gian thực hiện**

- Từ ngày 08/01/2024 đến ngày 07/04/2024

### **Vai trò & trách nhiệm của các thành viên trong nhóm**

- Mỗi nhóm sinh viên bao gồm 1-4 thành viên.
- Sinh viên cần phân công vai trò và trách nhiệm rõ ràng cho mỗi thành viên, có đánh giá trọng số đóng góp cá nhân của mỗi thành viên trong quá trình thực hiện đồ án

### **Kế hoạch thực hiện đồ án.**

STT	Yêu cầu	Thời gian	Ghi chú	Cách thức đánh giá
1	Giới thiệu chung	Tuần 2,3		Xem mục 3. Rubric và thang điểm
2	Cơ sở lý thuyết	Tuần 4,5		
3	Hiện thực	Tuần 6,7,8		
4	Hoàn thiện đồ án	Tuần 9,10	Nộp báo cáo	

### **Sản phẩm đồ án bao gồm:**

- Báo cáo đồ án theo mẫu: Trình bày các nội dung chính của đồ án (xem ở mục 2)
- Ứng dụng xử lý ngôn ngữ tự nhiên hoàn chỉnh (source code)
- Tài liệu trình bày (slide)
- Đánh giá trọng số cá nhân theo mẫu
- Bảng phân công công việc trong nhóm

## **2. Hướng dẫn thể thức trình bày đề bài**

### **a) Các nội dung chính đồ án**

- o Giới thiệu chung
  - Giới thiệu về đề tài và lý do chọn đề tài
  - Mục tiêu, phạm vi của đề tài
  - Phương pháp thực hiện

- Cơ sở lý thuyết
  - Khái niệm xử lý ngôn ngữ tự nhiên
  - Giới thiệu về mô hình mà đề tài sử dụng
- Hiện thực:
  - Dữ liệu sử dụng trong đồ án (chọn nguồn dữ liệu, tập huấn luyện, tập test...)
  - Cài đặt/huấn luyện mô hình
  - Test
  - Đánh giá mô hình
  - Kết quả thực hiện
- Kết luận
  - Tổng kết đề tài và những kết quả đạt được
  - Những hướng phát triển trong tương lai

#### b) Trình bày báo cáo:

- Báo cáo phải được trình bày ngắn gọn, rõ ràng, mạch lạc, sạch sẽ, không được tẩy xóa, có đánh số trang, đánh số bảng biểu, hình vẽ.
- Báo cáo sử dụng bảng mã Unicode, font chữ Times New Roman cỡ 13, dẫn dòng đặt ở chế độ 1.5 line; lề trên 3,5cm; lề dưới 3,0cm; lề trái 3,5cm; lề phải 2cm. Số trang được đánh ở giữa, phía trên đầu mỗi trang, đánh số trang 1 kể từ lời nói đầu.
- Báo cáo in đậm mục số, chữ và tên phần/mục và tuân theo nguyên tắc đánh số a) rập (1, 2, 3,...) (tránh dùng số la mã I, II, III, ...) nhiều cấp (thường tối đa 3 cấp)
- Việc đánh số ảnh, biểu đồ và bảng biểu phải gắn với số chương.
- Tựa hình ảnh, biểu đồ, sơ đồ nằm phía dưới hình ảnh, biểu đồ, sơ đồ - Tựa bảng biểu nằm phía trên bảng biểu.
- Chú thích ảnh, biểu đồ, bảng biểu được bố trí nằm phía dưới ảnh, biểu đồ và bảng biểu
- Danh sách tài liệu được đặt cuối bài viết, được bắt đầu bằng tiêu đề “Tài liệu tham khảo”, tiếp theo là danh mục liệt kê tài liệu tham khảo (sách, bài báo, nguồn ấn phẩm điện tử) được sắp xếp thứ tự Alphabet theo tên tác giả, tên bài viết, không đánh số thứ tự

### 3. Rubric và thang điểm

<b>Rubric: Đồ án</b>					
<b>Tiêu chí chấm điểm</b>	<b>Trọng số (%)</b>	<b>Tốt 100%</b>	<b>Khá 75%</b>	<b>Trung bình 50%</b>	<b>Không đạt &lt;50%</b>
Trình bày chính xác cơ sở lý thuyết	20%	Không sai sót	Sai sót khoảng 25%	Đúng 50% yêu cầu	Đúng dưới 50% yêu cầu
Vận dụng được giải thuật áp dụng cho bài toán	40%	Không sai sót	Sai sót khoảng 25%	Đúng 50% yêu cầu	Đúng dưới 50% yêu cầu

Kết quả thực nghiệm	20%	Không sai sót	Sai sót khoảng 25%	Đúng 50% yêu cầu	Đúng dưới 50% yêu cầu
Chủ động nghiên cứu, tìm hiểu trong quá trình làm bài tập. Phối hợp tốt với các thành viên trong nhóm	10%	Được các thành viên nhóm đánh giá 90% trở lên	Được các thành viên nhóm đánh giá từ 75%-90%	Được các thành viên nhóm đánh giá từ 60% đến 75%	Được các thành viên nhóm đánh giá $\leq 60\%$
Trình bày báo cáo khoa học, đúng yêu cầu.	5%	Đúng mẫu, đầy đủ nội dung	Sai sót khoảng 25%	Đúng 50% yêu cầu	Đúng dưới 50% yêu cầu
Thuyết trình thuyết phục, trả lời câu hỏi khi báo cáo đầy đủ, chính xác, hợp lý	5%	Thuyết trình tự tin, đầy đủ nội dung. Trả lời câu hỏi chính xác	Chưa tự tin, trả lời sai sót 25%	Thuyết trình không tự tin, trả lời sai sót 25%-50%	Thuyết trình không tự tin, trả lời sai sót $> 50\%$
<b>Tổng</b>	<b>100</b>				

TP. Hồ Chí Minh, ngày 26 tháng 03 năm 2024

Người duyệt đề

(Đã duyệt)

**Phùng Thế Bảo**

Giảng viên ra đề

  
**Nguyễn Thị Huyền Trang**