

The Saigon International
University



Khóa luận
tốt nghiệp

KHÓA LUẬN TỐT NGHIỆP

Ngành

Luật Kinh tế

Đề tài:

Bảo hộ Quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo - Xu hướng mới trên thế giới và bài học kinh nghiệm cho Việt Nam

Giảng viên hướng dẫn

TS. Nguyễn Thị Thu Hương

Sinh viên

Lê Nguyễn Thanh Trúc

MSSV: 97381901666



**The Saigon
International
University**

Lewis Campus

Email: admission@siu.edu.vn
Website: www.siu.edu.vn

Lời cam đoan

Tôi xin cam đoan khóa luận về đề tài “**Bảo hộ Quyền tác giả đối với tác phẩm từ TTNT - Xu hướng mới trên thế giới và bài học kinh nghiệm cho Việt Nam**” là công trình nghiên cứu cá nhân của tôi trong thời gian qua, được thực hiện dưới sự hỗ trợ từ Giảng viên TS Nguyễn Thị Thu Hương. Các nội dung nghiên cứu và kết quả trong đề tài này là trung thực, chưa được ai công bố trong bất cứ công trình nào. Mọi số liệu sử dụng phân tích trong khóa luận là do tôi tự tìm hiểu, phân tích một cách khách quan, trung thực, có nguồn gốc rõ ràng. Nếu có bất kỳ sự gian lận nào, tôi xin chịu trách nhiệm trước Hội đồng cũng như kết quả khóa luận của mình.

Do những hạn chế về trình độ chuyên môn cùng với thời gian nghiên cứu có hạn, khóa luận sẽ không thể tránh khỏi có những thiếu sót, em kính mong nhận được những ý kiến đóng góp từ các thầy, các cô và các bạn sinh viên khác để đề tài nghiên cứu được hoàn thiện hơn.

Xin chân thành cảm

ơn!

Thành phố. Hồ Chí Minh, ngày 25/04/2023

Sinh viên

Lê Nguyễn Thanh Trúc

Lời cảm ơn

Đề tài "**Bảo hộ Quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo - Xu hướng mới trên thế giới và bài học kinh nghiệm cho Việt Nam**" là nội dung em chọn nghiên cứu và làm khóa luận tốt nghiệp sau thời gian theo học ngành Luật kinh tế tại Đại học Quốc tế Sài Gòn. Trong suốt quá trình học tập tại trường và được sự cho phép và phê duyệt của bản giám hiệu và thầy/cô trưởng khoa để tiến hành thực hiện khóa luận tốt nghiệp. Em đã nhận được rất nhiều sự giúp đỡ, hỗ trợ từ quý thầy cô, gia đình và bạn bè.

Lời cảm ơn đầu tiên. Em xin chân thành cảm ơn giáo viên hướng dẫn của mình là của cô TS. Nguyễn Thị Thu Hương vì sự hỗ trợ và hướng dẫn không ngừng của cô trong suốt thời gian thực hiện khóa luận. Chuyên môn và sự kiên nhẫn của cô đối với em và đã đóng một vai trò quan trọng trong sự thành công của khóa luận này.

Trong quá trình làm khóa luận tốt nghiệp, em thấy rằng mình đã có cơ hội để học tập mở mang rất nhiều kiến thức hữu ích. Từ đó để em học hỏi và áp dụng và hành trang rất giá trị cho quá trình làm việc sau này của mình.

Khóa luận của em tất nhiên sẽ không thể tránh được những hạn chế, thiếu sót. Em rất mong nhận được những nhận xét và góp ý từ quý thầy cô và các bạn học để giúp khóa luận hoàn thiện hơn. Em xin chân thành cảm ơn.

Trân trọng, ngày 1 tháng 7 năm 2023

Sinh Viên

Nhận xét của giáo viên hướng dẫn

Mục Lục

Lời cam đoan	2
Lời cảm ơn	1
Nhận xét của giáo viên hướng dẫn	2
Mục Lục	3
Danh mục các ký hiệu, chữ viết tắt	5
Danh sách các bảng, biểu đồ, đồ thị, sơ đồ, hình ảnh	6
LỜI NÓI ĐẦU	7
Chương 1: Một số vấn đề lý luận về bảo hộ Quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo	10
1.1 Khái quát về công nghệ Trí tuệ nhân tạo	10
1.1.1. Lịch sử hình thành và phát triển của công nghệ Trí tuệ nhân tạo	10
1.1.2. Khái niệm và đặc điểm của công nghệ Trí tuệ nhân tạo	11
1.1.3. Phân loại các Trí tuệ nhân tạo hiện nay	14
1.1.4. Tác phẩm từ Trí tuệ nhân tạo	18
1.2. Khái quát về Quyền tác giả	19
1.2.1. Khái niệm của Quyền tác giả	19
1.2.2. Nội dung của Quyền tác giả	21
1.2.3 Đặc điểm của Quyền tác giả	24
1.3. Pháp luật về bảo hộ Quyền tác giả liên hệ đối với tác phẩm từ Trí tuệ nhân tạo.	25
1.3.1. Các loại tác phẩm được bảo hộ	25
1.3.2. Điều kiện bảo hộ Quyền tác giả đối với tác phẩm.	26
1.3.3. Cơ quan bảo hộ quyền tác giả với tác phẩm	27
1.3.4. Chế tài hành vi xâm phạm Quyền tác giả	28
Kết luận chương 1	29
Chương 2: Thực trạng pháp luật cùng những thách thức với việc bảo hộ tác phẩm từ Trí tuệ nhân tạo tại một số quốc gia và tại Việt Nam hiện nay	30
2.1 Thực trạng áp dụng pháp luật về quyền tác giả đối với các tác phẩm từ TTNT hiện nay tại Việt Nam và thế giới	30

2.2. Những thách thức pháp lý của Trí tuệ nhân tạo với hệ thống pháp luật Sở hữu trí tuệ Việt nam	32
2.2.1. Xác định tư cách pháp lý	35
2.2.2. Trí tuệ nhân tạo và vấn đề bảo hộ Quyền tác giả	37
2.2.3. Trách nhiệm bồi thường thiệt hại liên quan đến Trí tuệ nhân tạo	37
2.3 Thực trạng đặt ra với nhu cầu bảo hộ Quyền tác giả đối với các tác phẩm từ Trí tuệ nhân tạo.	39
Kết luận chương 2	40
Chương 3: Kinh nghiệm tại một số quốc gia về bảo hộ quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo và những giải pháp thúc đẩy bảo hộ quyền tại Việt Nam	41
3.1 Kinh nghiệm quốc tế về bảo hộ tác phẩm từ Trí tuệ nhân tạo	41
3.1.1. Về xác định tư cách pháp lý của Trí tuệ nhân tạo	41
3.1.2. Về xác định tác giả với tác phẩm từ Trí tuệ nhân tạo.	43
3.1.3. Về trách nhiệm bồi thường thiệt hại	46
3.2. Khuyến nghị, giải pháp bảo hộ quyền tác giả với tác phẩm từ Trí tuệ nhân tạo tại Việt Nam	48
3.2.1 Tư cách pháp lý của Trí tuệ nhân tạo.	48
3.2.2. Xác định tác giả tác phẩm từ Trí tuệ nhân tạo.	49
3.2.3. Bồi thường thiệt hại do Trí tuệ nhân tạo gây ra.	52
Kết luận chương 3	56
Kết Luận	57
Tài liệu tham khảo	59

Danh mục các ký hiệu, chữ viết tắt

STT	Ký hiệu chữ viết tắt	Chữ viết đầy đủ
1.	TTNT	Trí tuệ nhân tạo
2.	AI	Trí tuệ nhân tạo
3.	CNTT	Công nghệ thông tin
4.	ANN	Artificial Neural Networks
5.	DLT	<i>Sổ cái phân tán</i>
6.	<i>IoT</i>	<i>vạn vật kết nối</i>
7.	AAAI	<i>Hiệp hội vì sự tiến bộ của trí tuệ nhân tạo</i> (Association for the Advancement of Artificial Intelligence)
8.	ANN	Mạng lưới thần kinh nhân tạo
9.	Luật SHTT 2005	Luật Sở hữu trí tuệ 2005
10.	tác phẩm CAW	TTNT hỗ trợ tạo ra tác phẩm
11.	tác phẩm CGW	TTNT thực hiện và tạo ra tác phẩm
12.	BLDS 2015	Bộ luật dân sự 2015
13.	DABUS	Device for the Autonomous Bootstrapping of Unified Sentience
14.	EPO	Cơ quan Sáng chế châu Âu
15.	ASEAN	Hiệp hội các quốc gia Đông Nam Á
16.	WIPO	Tổ chức Sở hữu trí tuệ thế giới
17.	EVFTA	Hiệp định Thương mại Tự do Việt Nam – Liên minh châu Âu
18.	USCO	Văn phòng Bản quyền Hoa Kỳ
19.	CDPA	<i>Đạo luật Bản quyền, thiết kế và bằng sáng chế Vương quốc Anh</i>
20.	TNNN	Trách nhiệm nghiêm ngặt
21.	BTTH	Bồi thường thiệt hại

Danh sách các bảng, biểu đồ, đồ thị, sơ đồ, hình ảnh

1. Hình 1: Hỗ trợ từ chính phủ, đầu tư cho AI của doanh nghiệp
2. Hình 2: số liệu công bố khoa học WoS của 10 quốc gia Đông Nam Á về sáng chế

LỜI NÓI ĐẦU

1. Tính cấp thiết của đề tài

Xuôi theo dòng chảy lịch sử, nhân loại đã không ngừng đổi mới và phát triển để đạt tới chân thiện mỹ. Qua ngàn năm cải tiến, sự xuất hiện, thăng hoa và đi qua của các cuộc cách mạng công nghiệp 1,2,3 đã tạo tiền đề cho sự xuất hiện của cách mạng công nghiệp lần thứ 4 xuất hiện như một quá trình tất yếu của lịch sử. Cuộc cách mạng công nghiệp lần thứ tư bắt đầu hình thành từ năm 2000 tại Đức. Sau đó lan nhanh sang các nước khác như: Anh, Mỹ, Nhật,..Cuộc cách mạng công nghiệp lần thứ tư được xây dựng tập trung vào 3 ngành chính là: công nghệ sinh học, kỹ thuật số và vật lý. Sau hơn 20 năm hội nhập, cuộc cách mạng này đã thay đổi cả bộ mặt của xã hội trong suốt nhiều thập kỷ qua.

Lĩnh vực công nghệ được xem như đại diện cho cuộc cách mạng công nghiệp 4.0 với sự bùng nổ của xu hướng tự động hóa. Nếu như bánh xe, bóng đèn, la bàn đã thay đổi xã hội trước. Thì giờ đây, internet hay kết nối không dây đã thay đổi hoàn toàn xã hội. Nếu đã đề cập đến lĩnh vực công nghệ thì không thể bỏ qua những “ngôi sao” rực rỡ trên bầu trời kỹ thuật số đó chính là: Công nghệ số cái phân tán (DLT), Internet vạn vật (IoT), học máy (ML), vạn vật kết nối (IoT), dữ liệu lớn (Big Data). Với sự gia tăng nhanh chóng về mức độ phức tạp của các nhiệm vụ mà AI có thể thành thạo bất kể lĩnh vực nào chỉ cần chú trọng vào dữ liệu đầu vào. Và khi này khả năng bảo vệ các kết quả đầu ra, nó ngày càng có ý nghĩa thương mại, đóng góp to lớn cho nền kinh tế toàn cầu và là vấn đề được các nhà đầu tư hết sức quan tâm.

Đây không phải là lần đầu tiên con người tạo ra máy móc, nhưng lần này máy móc trở lại với phần mềm máy tính được cải tiến vượt bậc, được đặt tên là “**Trí tuệ nhân tạo**” (sau đây gọi tắt là “TTNT” hoặc “AI”). Sự trở lại lần này không phải với tư cách là một công cụ, người làm thuê mà với tư cách là một kẻ chinh phạt, một bậc thầy sáng tạo. TTNT không những đóng vai trò “kiến tạo” những thành tựu lớn trong xã hội mà còn cụ thể hóa những kiến tạo đó bằng con số GDP tăng trưởng suốt nhiều năm qua nhờ sự đóng góp không nhỏ của TTNT. Những lợi ích mà AI mang lại là điều không cần phải bàn cãi, đi cùng những lợi ích AI đồng thời cũng gây ra nhiều lo lắng, quan ngại.

Vấn đề mà các nhà làm luật quan ngại nhất trong những năm gần đây chính là mối quan hệ giữa các tác phẩm được hình thành bởi TTNT với pháp luật Sở hữu trí tuệ. TTNT vốn đã là một vấn đề phức tạp và pháp luật thế giới chỉ mới từng bước điều chỉnh được. Thì ngày nay, sự xuất hiện của những tác phẩm được hình thành bởi TTNT đã đặt ra một thách thức cho các nhà làm luật, tạo nên một mảng xám lớn tồn tại trong hệ thống pháp luật nói chung cũng như luật sở hữu trí tuệ nói riêng. Những vấn đề khiến xã hội và giới chuyên môn quan tâm là khi

những tác phẩm được tạo ra bởi TTNT thì vấn đề quyền tác giả sẽ thuộc về ai ? hay TTNT có địa vị như thế nào với hệ thống pháp luật?. Trước bối cảnh đó, đề tài tôi lựa chọn để tìm hiểu và đưa ra phân tích là “**Bảo hộ Quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo - Xu hướng mới trên thế giới và bài học kinh nghiệm cho Việt Nam**”

2. Đối tượng nghiên cứu và phạm vi nghiên cứu

Đối tượng nghiên cứu

Khóa luận này dựa trên Luật sở hữu trí tuệ theo pháp luật Việt Nam và một số quốc gia trên cơ sở về vấn đề bản quyền. Tìm hiểu cấu trúc, bản chất hoạt động của công nghệ TTNT. Mỗi quan hệ pháp lý giữa quyền tác giả đối với những tác phẩm được hình thành bởi TTNT. Cùng với hướng tiếp cận của một số nước trên thế giới với vấn đề bảo hộ quyền tác giả đối với tác phẩm được hình thành bởi TTNT.

Phạm vi nghiên cứu

Khóa luận này giới hạn nghiên cứu trong phạm vi Quyền tác giả tập trung xoay quanh, làm rõ 3 vấn đề chính là: địa vị pháp lý của TTNT, xác định Quyền tác giả các tác phẩm được tạo ra bởi TTNT và trách nhiệm bồi thường thiệt hại nếu TTNT gây ra thiệt hại. Khóa luận được thực hiện dựa trên pháp luật Việt Nam về vấn đề sở hữu trí tuệ, cụ thể là pháp luật về quyền tác giả. Cùng với đó là học hỏi, tham khảo từ pháp luật, kinh nghiệm quốc tế về việc bảo hộ quyền tác giả với tác phẩm từ TTNT.

3. Câu hỏi nghiên cứu

Khóa luận được thực hiện dựa trên một số câu hỏi nghiên cứu trọng tâm như:

- Pháp luật Việt Nam hiện nay quy định như thế nào về quyền tác giả với tác phẩm từ TTNT.
- Những thách thức pháp lý nào được đặt ra giữa mối quan hệ quyền tác giả và những tác phẩm được hình thành bởi TTNT
- Những quốc gia tiên tiến tiếp cận vấn đề bảo hộ quyền tác giả các tác phẩm từ TTNT như thế nào ?

4. Mục tiêu và nhiệm vụ nghiên cứu

- Trình bày cơ sở pháp lý hiện hành quy định về quyền tác giả nói chung và quyền tác giả với những tác phẩm từ TTNT nói riêng.

- Phân tích, làm rõ những thách thức về mặt pháp lý với vấn đề quyền tác của những tác phẩm từ TTNT.

- Nêu lên thực trạng của vấn đề hiện nay và đưa ra những cách tiếp cận vấn đề hiện nay tại một số quốc gia tiên tiến.

5. Phương pháp nghiên cứu

Để thực hiện đề tài có sử dụng một số phương pháp nghiên cứu khoa học luật cụ thể như:

Chương 1: Phương pháp phân loại và hệ thống hóa lý thuyết

Chương 2: Phương pháp phân tích và tổng hợp, Phương pháp lịch sử. Phương pháp phân tích nhằm hệ thống thực trạng tồn tại trong việc thực thi các quy định pháp luật

Chương 3: Phương pháp phân tích và tổng kết kinh nghiệm

Ngoài ra, Đề tài còn sử dụng một số phương pháp nghiên cứu khác nhằm phân tích quy phạm pháp luật thực định, phương pháp dự báo pháp luật, để giải quyết các nhiệm vụ nghiên cứu đặt ra trong đề tài.

6. Bố cục của khóa luận

Ngoài các phần như lời cảm ơn, danh mục viết tắt, kết luận, danh mục tài liệu tham khảo, chương tổng quan thì nội dung chính của khóa luận được chia thành ba chương sau:

Chương 1: Một số vấn đề lý luận về bảo hộ quyền tác giả đối với tác phẩm từ TTNT

Chương 2: Thực trạng pháp luật cùng những thách thức với việc bảo hộ tác phẩm từ Trí tuệ nhân tạo tại một số quốc gia và tại Việt Nam hiện nay.

Chương 3: Kinh nghiệm tại một số quốc gia về bảo hộ quyền tác giả đối với tác phẩm từ AI và những giải pháp thúc đẩy bảo hộ quyền tại Việt Nam

Chương 1: Một số vấn đề lý luận về bảo hộ Quyền tác giả đối với tác phẩm từ Trí tuệ nhân tạo

1.1 Khái quát về công nghệ Trí tuệ nhân tạo

1.1.1. Lịch sử hình thành và phát triển của công nghệ Trí tuệ nhân tạo

Nói đến công nghệ TTNT hay còn gọi là AI. Chúng ta có thể nghĩ rằng AI là sự phát triển của khoa học công nghệ trong thập niên gần đây. Tuy nhiên, sự thật là một ý tưởng về “TTNT” đã có từ hàng ngàn năm trước, khi các nhà phát minh cổ đại đã tạo ra những thứ được gọi là “máy tự động” có tính cơ học và di chuyển độc lập với sự can thiệp của con người. Từ “ máy tự động ” xuất phát từ tiếng Hy Lạp cổ đại, và có nghĩa là “hành động theo ý muốn của chính mình”. Một trong những ghi chép sớm nhất về máy tự động có từ năm 400 trước Công nguyên và đề cập đến một con chim bồ câu máy do một người bạn của nhà triết học Plato tạo ra. Nhiều năm sau, một trong những cỗ máy tự động nổi tiếng nhất được tạo ra bởi Leonardo da Vinci vào khoảng năm 1495 .

Theo dòng lịch sử đến thế kỷ 20, nền tảng của TTNT đã bắt đầu từ trước những năm 1900. Và sự xuất hiện của chương trình TTNT đầu tiên được đặt tên là “Nhà lý thuyết logic” do Allen Newell và Herbert A.Simon lập trình. Một năm sau đó đánh dấu sự công nhận dành cho TTNT khi khái niệm “TTNT” lần đầu được nhà khoa học người Mỹ - John McCarthy sử dụng tại Hội nghị Dartmouth¹. Lần đầu tiên, AI được coi là một lĩnh vực học thuật. Sau thời điểm này, TTNT đã có những bước tiến quan trọng từ đó tạo thành cơ sở phát triển sau này. Nếu giai đoạn từ năm 1950 - 1960 là thời điểm của sự sáng tạo thì giai đoạn từ 1970 đến 1980 là giai đoạn TTNT phát triển mạnh mẽ. Vào năm 1972, Người máy thông minh hình người đầu tiên được chế tạo tại Nhật Bản có tên là WABOT-1. Và đến năm 1979 Hiệp hội TTNT Hoa Kỳ, hiện được gọi là Hiệp hội vì sự tiến bộ của TTNT (AAAI) được thành lập.

TTNT đã không ít lần bị các nhà đầu tư và chính phủ từ chối thậm chí là rút vốn ngừng tài trợ cho các dự án nghiên cứu TNTT vì chi phí cao. Tuy nhiên, với tiềm lực phát triển của mình mà TTNT vẫn là mối quan tâm lớn của giới khoa học và thế giới nói chung. TTNT vượt qua những thăng trầm mà phát triển hơn thế. TTNT không còn là những chương trình học máy được lập trình sẵn phụ thuộc hoàn toàn vào con người. Khi vào năm 1997, chương trình học máy TNTT - IBM Deep Blue đánh bại nhà vô địch cờ vua thế giới là Gary Kasparov, và trở thành máy tính đầu tiên đánh bại nhà vô địch cờ vua thế giới. Trong gần 2 thập

¹Rockwell Anyoha (2017), Harvard University The Graduate School of Arts and Sciences, The History of Artificial Intelligence, <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

niên qua các nhà khoa học đã cho ra mắt những chương trình công nghệ như nhận dạng giọng nói, tự động hóa quy trình bằng rô-bốt (RPA), rô-bốt nhảy múa, nhà thông minh và các cải tiến khác. Năm 2020, giai đoạn virus Covid-19 còn diễn biến phức tạp, Baidu-trang mạng hàng đầu Trung Quốc phát hành thuật toán LinearFold AI cho các nhóm y tế, khoa học và y tế đang phát triển vắc-xin trong giai đoạn đầu của đại dịch SARS-CoV-2 (COVID-19). Thuật toán có thể dự đoán trình tự RNA của virus chỉ trong 27 giây, nhanh hơn 120 lần so với các phương pháp khác. Hay gần đây nhất có thể kể đến là sự xuất hiện của chương trình siêu máy tính mang tên ChatGPT đã tạo nên cơn sốt toàn cầu.

Có thể thấy, CNTT giờ đây không còn là những chương trình máy tính cơ bản để chơi cờ caro hay giải toán. Mà giờ đây CNTT đang dần phát triển và phát triển đến mức có thể thay thế một số hoạt động của con người trong đời sống. Sự phát triển của CNTT trong tương lai là điều không cần tranh cãi và sự phát triển đó còn có thể đạt mức độ nào trong thời gian tới vẫn còn là một dấu chấm hỏi lớn với cả nhân loại. Không ai trong chúng ta có thể biết được đâu là giới hạn cho sự phát triển của CNTT. Nhưng chúng ta có thể chắc chắn rằng CNTT sẽ vẫn phát triển và vẫn mang đến những lợi ích to lớn hơn trong đời sống của chúng ta trong tương lai.

1.1.2. Khái niệm và đặc điểm của công nghệ Trí tuệ nhân tạo

Thứ nhất, khái niệm của công nghệ Trí tuệ nhân tạo

Đứng trước sự bùng nổ của ngành công nghệ thông tin, việc nghiên cứu và đưa ra định nghĩa cho đối tượng là vô cùng quan trọng. Hiện nay trên thế giới có nhiều khái niệm khác nhau về CNTT. Chưa có một khái niệm nào được thống nhất để sử dụng trên toàn thế giới. Tuy nhiên vẫn có 2 trường phái về khái niệm CNTT là Strong AI và Weak AI. Theo đó, Strong AI có thể tạo ra thiết bị có trí thông minh và các chương trình máy tính thông minh hơn con người còn Weak AI là chương trình máy tính có thể mô phỏng các hành vi thông minh của con người². Nếu Strong AI có khả năng vẽ những nét phác thảo đầu tiên, định hình cho một bức tranh thì Strong AI sẽ có chức năng vẽ chi tiết và khiến cho bức tranh đẹp nhất có thể. Thường thì Weak AI là mối ưu tiên hàng đầu của các nhà nghiên cứu vì Weak AI có trách nhiệm tập trung vào một nhiệm vụ cụ thể, hoạt động dưới nhiều ràng buộc hơn cả trí thông minh cơ bản nhất của con người để hoàn thiện nhiệm vụ đó và thực hiện nó thậm chí còn tốt hơn cả con người. Một ví dụ điển hình và nổi tiếng nhất cho Weak AI chính là Deep Blue, một máy tính do IBM tạo ra đã đánh bại nhà vô địch cờ vua thế giới Gary Kasparov trong một trận đấu sáu ván vào năm 1997 (Kasparov đã thắng trận đầu tiên một

² TS. Ngô Hữu Phúc (2011), Bài giảng nhập môn CNTT, Học viện Kỹ thuật Quân sự.

năm trước đó). Deep Blue có thể chọn từ hàng trăm triệu nước đi và đã “nhìn thấy” trước đối thủ 20 nước đi — một thành tích mà cho đến nay chưa con người nào có thể đạt được.

Nhưng dù là tiếp cận từ trường phái nào thì định nghĩa của TTNT đều sẽ nói đến khả năng của chương trình máy tính trong khi thực hiện các công việc mà con người thường phải xử lý; và trong quá trình thực hiện lệnh đó sẽ đưa ra kết quả thực hiện của máy là tốt hơn hoặc tương đương với con người thì người ta gọi đó là máy thông minh hay máy có trí thông minh.³ TTNT sẽ được thể hiện dưới một hệ thống chương trình đã được con người lập trình sẵn để chương trình TTNT đó trông như có vẻ biết suy nghĩ và hành động trên cơ sở logic có tính chính xác như con người. Và để đưa ra được kết quả như vậy, con người đã trang bị cho nó các công cụ như thị giác, thính giác, cơ sở tiếp nhận thông tin và xử lý thông tin, cơ chế tự phân giải hay di chuyển như một con người thực thụ.

Thứ hai, đặc điểm của công nghệ TTNT

Nói một cách đơn giản, TTNT (AI) là một công cụ hiệu quả cho phép máy móc học hỏi kinh nghiệm, thích nghi với những thay đổi mới và thực hiện các nhiệm vụ giống như con người. Khái niệm cốt lõi của AI là có quyền truy cập vào tất cả thông tin liên quan đến các đối tượng, thuộc tính, danh mục và mối quan hệ giữa tất cả các trường hợp sử dụng nghiệp vụ để triển khai kỹ thuật đã được lập trình. TTNT có thể được phân thành nhiều loại và được định nghĩa theo nhiều cách khác nhau. Tuy nhiên TTNT sẽ có một số đặc điểm cơ bản mà chúng ta có thể kể đến như:

(i) Xử lý đặc trưng (Feature Engineering)

Xử lý đặc trưng là các bước tiền xử lý chuyển đổi dữ liệu thô thành các tính năng có thể được sử dụng trong các thuật toán máy học, chẳng hạn như các mô hình dự đoán. Trong vũ trụ thông tin mà chương trình học máy nhận được, nhờ có tính năng Xử lý đặc trưng - được ví như bộ lọc khổng lồ sẽ giúp chương trình sau khi tiếp nhận dữ liệu thô sẽ tiến hành sắp xếp, xử lý, thiết kế theo những đặc trưng (feature) nhất định từ dữ liệu đầu sau đó mới tiến hành xây dựng mô hình để giải quyết các yêu cầu được đưa ra. Bằng cách sắp xếp, lập trình đầu vào tốt thông qua Feature Engineering chúng ta có thể tạo ra những đặc trưng hiệu quả nhất giúp huấn luyện mô hình phân loại hoặc dự báo với độ chính xác cao, sai số thấp. Một số phương pháp trong quá trình Xử lý đặc trưng thường được sử dụng có thể kể đến như: phương pháp Trích lọc đặc trưng (Feature Extraction), Biến đổi đặc trưng (Feature Transformation) và Lựa chọn đặc trưng (Feature Selection). Vì vậy có thể nói Xử lý đặc trưng (Feature Engineering) vừa là nghệ

³ George F. Luger - William A. Stubblefield, “Artificial Intelligence”, Wesley Publishing Company, 1997, tr.30.

thuật vừa là bộ môn khoa học khiến các nhà lập trình dữ liệu đặc biệt quan tâm và thường dành 70% thời gian để chuẩn bị dữ liệu trước khi xây dựng mô hình.

(ii) Mạng lưới thần kinh nhân tạo (ANN)

Artificial Neural Networks (ANN) hay được gọi là Mạng thần kinh nhân tạo. Mạng lưới này sẽ dựa trên tập hợp các nút được kết nối được gọi là tế bào thần kinh nhân tạo giống như tế bào não người. Các nút trong ANN được sử dụng để lấy dữ liệu đầu vào. Mỗi kết nối truyền một tín hiệu từ nơ-ron này sang nơ-ron khác sau khi xử lý nó.

Mạng lưới thần kinh nhân tạo (ANN) được ví như bộ não của AI nơi tập hợp các dây thần kinh, một bộ não điện tử thực thực. Mạng lưới nơ-tron của AI sẽ hoạt động như mạng lưới nơ-tron của con người. Trong đó mỗi nơ-tron là một hàm toán học có chức năng thu thập, xử lý và phân loại thông tin theo một cấu trúc nhất định.

Mạng lưới thần kinh nhân tạo này cũng được ứng dụng nhiều trong đời sống ví dụ như sử dụng ANN trong việc phát hiện những lỗi còn sót lại của máy bay lái tự động trong lĩnh vực không gian vũ trụ. ANN còn giúp phân biệt, nhận dạng giọng nói. Ngoài ra trong lĩnh vực y khoa ANN có khả năng chuẩn đoán các tế bào ung thư, dựa vào ANN mà phân tích điện não đồ và điện tâm đồ. Nhưng lợi ích từ chương trình ANN đem lại trong vài thập niên qua là một kết quả bất ngờ. Tuy nhiên trong thời gian tới, sự phát triển của ANN vẫn là một lĩnh vực rất được trông đợi. Những gã khổng lồ trong ngành công nghiệp AI đang dự đoán rằng lĩnh vực AI đang chuyển dịch từ điện toán đám mây(Cloud Computing) sang điện toán biên (Edge Computing). Lợi ích của sự thay đổi này là kiểm tra tốc độ, tiết kiệm điện, điều phối giao thông và điều khiển xe nâng hàng. Để biến nó thành hiện thực, kích thích mô hình AI cần phải giảm xuống và do đó, một số kỹ thuật đang được phát triển để thu nhỏ mạng lưới thần kinh mà không ảnh hưởng đến hiệu suất. Chính vì vậy, sự phát triển và ứng dụng vượt bậc của ANN trong thời gian tới vẫn là một ẩn số đáng trông chờ.

(iii) Học sâu (Deep learning)

Thế giới hiện đại chứa đầy những dữ liệu thông tin và với sự trợ giúp của cơ chế học sâu, thế giới kỹ thuật số đang biến thành một nơi tuyệt đẹp hơn trước đây. Đó là một kỹ thuật học máy giúp máy tính tự động suy nghĩ giống như con người. Học sâu (Deep learning) là một nhánh trong phạm trù của Machine learning. Nếu Machine learning cho phép CNTT nghiên cứu và xây dựng các kỹ thuật “học” tự động từ những dữ liệu ban đầu để hoàn thành những nhiệm vụ được giao. Thì Deep learning tập trung giải quyết các vấn đề liên quan đến mạng thần kinh nhân tạo nhằm nâng cấp các công nghệ như nhận diện giọng nói, tầm nhìn máy tính và xử lý ngôn ngữ tự nhiên. Nếu như chúng ta ví CNTT như một ngôi nhà 3 tầng thì ngôi nhà điện tử này được cấu thành từ các tầng xếp chồng lên nhau, trong đó mạng thần kinh nhân tạo nằm ở tầng thứ nhất - là nền móng của ngôi nhà,

machine learning nằm ở tầng tiếp và cuối cùng deep learning nằm ở tầng trên cùng của ngôi nhà.

Deep learning đã trở thành một lĩnh vực đáng được kỳ vọng và mang lại nhiều ứng dụng cho đời sống con người. Điển hình như mô hình xe lái tự động Tesla ở chế độ Autopilot, trong đó Học sâu giúp phân biệt giữa tín hiệu dừng hoặc tín hiệu xanh và đưa ra quyết định lái xe hay không lái xe. Các ví dụ khác là cá nhân hóa nguồn cấp dữ liệu trên phương tiện truyền thông xã hội, nhận dạng hình ảnh, nhận dạng văn bản trực tuyến, v.v.

TTNT dù được tạo ra, mô phỏng theo cấu trúc của con người. Tuy nhiên trong quá trình tạo hình, TTNT vẫn mang trong mình những đặc điểm rất riêng biệt. TTNT được phân thành nhiều nhánh với nhiều mô hình, nhiều tính năng khác nhau. Dù là Deep Blue hay Chat GPT thì nói chung TTNT vẫn mang trong mình những đặc điểm chung gắn kết các nhánh của lĩnh vực TTNT.

1.1.3. Phân loại các Trí tuệ nhân tạo hiện nay

AI hay TTNT là chủ đề “nóng sốt” trong suốt những thập niên qua. Đây không chỉ là chủ đề làm đầu đầu các nhà nghiên cứu, các nhà làm luật, là cánh cửa tiềm năng cho các doanh nghiệp. Dù không phải là một nhà nghiên cứu trong lĩnh vực này thì bạn vẫn có thể kể ra vài cái tên tiêu biểu nổi tiếng rầm rộ trong thời gian qua. Nhưng nếu bạn hiểu được cách mà AI được phân loại thì bạn sẽ thật sự hiểu được sự thành công đáng nể phục của những “gã khổng lồ” công nghệ đứng đằng sau những thành phẩm AI đó. Nếu tìm hiểu sâu vào các phân loại của AI thì bạn có lẽ sẽ bất ngờ vì có đến 4,5 loại hay nhiều hơn thế. Xin đừng bất ngờ !. Dù có thành mấy loại thì nó vẫn chính xác và hợp lý. Và lý giải cho vấn đề này chính là do AI được mô phỏng theo trí tuệ con người. Mà đã là con người thì các nhà nghiên cứu có thể tiếp cận từ nhiều khía cạnh khác nhau. Trong phạm vi khóa luận này, tôi sẽ đưa ra 2 hướng được tiếp cận rộng rãi nhất. Hướng tiếp cận thứ nhất chính là phân loại dựa trên sự tương đồng của AI với trí tuệ con người và hướng tiếp cận thứ hai là phân loại dựa trên mức độ thông minh, linh hoạt của AI. Và trong mỗi hướng tiếp cận, các chuyên gia sẽ phân AI thành các loại khác nhau. Chúng ta sẽ sơ lược qua những phân loại AI theo 2 hướng tiếp cận trên.

Một, Phân loại AI dựa trên sự tương đồng của AI với trí tuệ con người.

Dựa trên sự phân loại theo mức độ tương đồng giữa AI với trí tuệ con người. Chúng ta sẽ có 4 loại:

(i) Công nghệ AI phản ứng (Reactive Machine)

Đây là loại có hệ thống cơ bản nhất được lập trình để cung cấp đầu ra có thể dự đoán được dựa trên đầu vào mà nó nhận được. Các cỗ máy phản ứng luôn phản ứng với các tình huống giống hệt nhau theo cùng một cách chính xác mọi lúc và

chúng không thể học các hành động hoặc hình dung về quá khứ hoặc tương lai. AI loại cơ bản này không có khả năng hình thành ký ức cũng như không sử dụng kinh nghiệm trong quá khứ để đưa ra các quyết định hiện tại. Deep Blue, siêu máy tính chơi cờ của IBM, đã đánh bại đại kiện tướng quốc tế Garry Kasparov vào cuối những năm 1990, là ví dụ hoàn hảo cho loại máy này. Deep Blue có thể xác định các quân cờ trên bàn cờ vua và biết cách di chuyển của từng quân cờ. Nó có thể đưa ra dự đoán về những bước đi tiếp theo có thể xảy ra đối với nó và đối thủ của nó. Và nó có thể chọn những nước đi tối ưu nhất trong số các khả năng.

AI có khả năng phản ứng lại thế giới quan đã là một bước tiến của khoa học. Tuy nhiên những loại AI này không thể hoạt động ngoài các nhiệm vụ mà chúng được giao. Và đương nhiên trí tuệ con người không dừng lại ở đó. Loại TTNT tiếp theo đây cũng được phát triển dựa trên nền tảng của TTNT cơ bản này.

(ii) Công nghệ AI với trí nhớ hạn chế (Limited memory machines)

Loại Công nghệ Ai này là phiên bản nâng cấp của loại Công nghệ AI phản ứng. Khắc phục được những khuyết điểm cũ đó chính là ở phiên bản này Công nghệ AI với trí nhớ hạn chế có khả năng ghi nhớ quá khứ, tích lũy kinh nghiệm đã có trong quá khứ kết hợp cùng những dữ liệu đã được lập trình sẵn mà đưa ra kết quả dự đoán có khả thi nhất. Và đây cũng loại hình AI được sử dụng khá phổ biến hiện nay. Một ví dụ điển hình cho mô hình này chính là xe ô tô không người lái. Với mô hình Công nghệ AI với trí nhớ ngắn hạn, chúng được tích hợp cảm biến cùng với dữ liệu có sẵn để tính toán được khoảng cách an toàn của xe phía sau với xe phía trước, từ đó dự đoán nguy hiểm có thể xảy ra nếu khoảng cách an toàn bị phá vỡ và gây ra va chạm để điều chỉnh tốc độ xe, đảm bảo an toàn và tránh gây tai nạn giao thông.

(iii) Lý thuyết về TTNT (Theory of mind AI)

Với loại AI này, máy móc sẽ có được khả năng ra quyết định thực sự tương tự như con người. Những cỗ máy với lý thuyết trí tuệ AI sẽ có thể hiểu và ghi nhớ cảm xúc, sau đó điều chỉnh hành vi dựa trên những cảm xúc đó khi chúng tương tác với con người. Đồ án tiến sĩ robot Kismet do Giáo sư Cynthia Breazeal phát triển, có thể nhận ra các tín hiệu cảm xúc trên khuôn mặt con người và tái tạo những cảm xúc đó trên khuôn mặt của chính nó. Sophia - robot hình người do Hanson Robotics ở Hồng Kông phát triển, có thể nhận dạng khuôn mặt và phản ứng với các tương tác bằng nét mặt của chính mình. Cả hai robot trên đều được xây dựng dựa trên nền tảng công nghệ về Lý thuyết của TTNT. AI theo mô hình “Theory of mind AI” hiện đang được phát triển, nhằm mục đích nhận biết, mô phỏng, theo dõi và phản ứng phù hợp với cảm xúc của con người bằng cách phân tích giọng nói, hình ảnh và các loại dữ liệu khác. Nhưng khả năng này của AI mặc dù có tiềm năng vô giá trong nhiều lĩnh vực đời sống như chăm sóc sức khỏe, quảng cáo và nhiều lĩnh vực khác. Tuy nhiên đây vẫn chưa là phiên bản

hoàn hảo nhất và vẫn còn lâu mới trở thành một AI sở hữu tâm trí thật sự. Bởi vì dù AI có thể vẽ một bức tranh đẹp đến đâu hay có thể viết một bài báo cáo, bài luận chính xác thì AI vẫn không thật sự hiểu được những gì nó viết.

(iv) AI tự nhận thức

Đây là bước phát triển cao nhất của công nghệ trí tuệ AI. Phần cuối cùng cho sự phát triển của AI sẽ là thiết kế các hệ thống có ý thức về bản thân, hiểu biết có ý thức về sự tồn tại của chúng. Tuy nhiên loại AI này chưa tồn tại và chưa khả thi trong giai đoạn hiện tại. Bởi vì con người chưa thật sự kiểm soát được AI vì vậy giai đoạn này vẫn là mục tiêu các nhà nghiên cứu đang hướng tới. Khi đạt tới giai đoạn này AI sẽ không chỉ có thể hiểu được cảm xúc, đặc điểm và suy nghĩ của con người mà còn có sự hiểu biết, cảm xúc và mong muốn tiềm ẩn của riêng nó. Mặc dù điều này nghe có vẻ thú vị, nhưng nhiều nhà nghiên cứu lo sợ về một kịch bản như vậy. Sẽ ra sao nếu một con robot có khả năng tự suy duy và có thể đọc thấu suy nghĩ con người. Không điều gì có thể bảo đảm cho sự phục tùng, tuân theo mệnh lệnh mãi mãi của những con robot đó. Và khi đó một dấu chấm hỏi lớn cho sự tồn tại của con người sẽ xuất hiện. Vì vậy cho đến khi nào con người có thể hoàn toàn kiểm soát được TTNT thì sự xuất hiện của loại công nghệ AI tự nhận thức vẫn còn trong vòng bí mật.

Sau khi tìm hiểu sơ qua 4 loại công nghệ AI nếu được phân loại dựa trên mức độ tương đồng giữa AI với trí tuệ con người. Chúng ta sẽ tìm hiểu 3 loại công nghệ AI nếu chúng được phân loại dựa trên tiêu chí khác, là:

Hai, phân loại theo mức độ thông minh, linh hoạt của TTNT.

Nếu dựa theo tiêu chí này, chúng ta sẽ có 3 loại: TTNT hẹp (ANI), TTNT chung (AGI), Siêu TTNT (ASI).

(i) TTNT hẹp (ANI)

TTNT hẹp (ANI) hay còn gọi là AI yếu hoặc là AI hẹp. Đây là loại TTNT duy nhất mà đến hiện nay chúng ta đã thực hiện thành công. TTNT thu hẹp sẽ hướng đến mục tiêu là được thiết kế để thực hiện các nhiệm vụ đơn lẻ chẳng hạn như: nhận diện khuôn mặt, giọng nói,... AI trong phân loại này được thiết lập tập trung vào hoàn thành những nhiệm vụ đơn lẻ. Chúng rất thông minh trong việc hoàn thành nhiệm vụ cụ thể mà nó được lập trình để thực hiện.

Những cỗ máy TTNT hẹp này có vẻ thông minh. Tuy nhiên chúng vẫn bị ràng buộc trong những giới hạn nhất định, chính vì vậy ở loại này chúng có những khuyết điểm và phạm vi sáng tạo của AI cũng hẹp hơn những AI được phát triển sau này. Một số ứng dụng của TTNT hẹp trong đời sống xã hội có thể kể đến là: Siri của Apple, Alexa của Amazon, Cortana của Microsoft và các trợ lý ảo khác, Watson của IBM, phần mềm nhận dạng hình ảnh / khuôn mặt, công cụ lập bản đồ và dự đoán bệnh, sản xuất và robot bay không người lái. TTNT hẹp được sử dụng khá phổ biến trong cuộc sống, mang lại nhiều lợi ích cho con người. Tuy

nhiên, chúng vẫn còn nhiều hạn chế và phạm vi vẫn còn rất hẹp. Chính vì thế, đây sẽ là cơ sở để những TTNT tân tiên hơn xuất hiện sau này.

(ii) TTNT tổng hợp (AGI)

TTNT tổng hợp (AGI), còn được gọi là AI mạnh hoặc AI sâu. Khái niệm này đề cập đến mô hình AI với trí tuệ thông minh tổng hợp có khả năng bắt chước hành vi của con người. Với trí thông minh được lập trình, công nghệ này cho phép AI áp dụng kiến thức và kỹ năng trong nhiều tình huống khác nhau. Hiện nay, các nhà nghiên cứu vẫn đang cố gắng để tiến đến trình độ TTNT tổng hợp. Tuy nhiên để có một phương pháp để khiến cho máy móc có ý thức, với bộ khả năng nhận thức đầy đủ. Đầu năm 2023, sự xuất hiện của ChatGPT với khả năng tổng hợp thông tin để trao đổi, sản xuất, tóm tắt, phân tích đã đem đến nhiều bất ngờ cho con người. Trước đây, lĩnh vực TTNT tổng hợp - General AI này đã nhận được khoản đầu tư lên đến 1 tỷ USD từ Microsoft thông qua công ty chủ quan là OpenAI. Chat GPT mang đến nhiều bất ngờ hơn so với một cỗ máy AI thông thường, tuy nhiên Chat GPT vẫn chưa thực sự đạt được trình độ như một General AI thực thụ. Bởi vì Chat GPT dù hay đến đâu vẫn phải phụ thuộc và lượng thông tin không lồ ở giai đoạn đầu vào. Chứ Chat GPT chưa có thể tự suy nghĩ, tự nhận thức như bộ não con người. Chính vì vậy, dù là nỗ lực đáng được công nhận nhưng Chat GPT chưa thể được coi là một TTNT tổng hợp hoàn toàn. Chương trình máy tính K do công ty Fujitsu sản xuất, là một siêu máy tính nhanh nhất thế giới và được kế thừa bởi chương trình máy tính Fugaku sau này với năng lực xử lý dữ liệu lớn (big data) nhanh chóng. Tuy nhiên cũng phải mất hơn 30 phút để chiếc siêu máy tính này mô phỏng 1 giây hoạt động của bộ não người. Đây đúng là nỗ lực của giới chuyên môn trong công cuộc tiến đến giai đoạn General AI. TTNT tổng hợp hay được gọi là AI mạnh là kỹ nguyên rất đáng được trông chờ bởi khả năng học hỏi, hoàn thành bất kỳ nhiệm vụ nào con người có thể làm được thì AI mạnh cũng có thể làm được. Tuy nhiên trong một tương lai xa, vẫn chưa thật sự biết được khi nào General AI mới được thực sự hoàn thiện.

(iii) Siêu TTNT - Super AI

Siêu TTNT (Super AI) tương tự như TTNT tự nhận thức, đây được xem là đỉnh cao của công nghệ AI - là nấc thang danh vọng mà tất cả các nhà nghiên cứu đều muốn bước đến. Khi đạt tới trình độ này TTNT sẽ vượt qua cả trí thông minh của con người và khi đó nó sẽ thực hiện tất cả các nhiệm vụ được giao tốt hơn con người. Trong chương trình cấu tạo của Siêu TTNT, có thể thấy AI được cấu tạo để nhằm đạt đến khả năng có được cảm xúc, khả năng, trí tuệ như con người thậm chí là hơn. Ngoài ra, chúng còn có nhu cầu, niềm tin, mong muốn của riêng nó.

Hiện nay, dự đoán về sự xuất hiện của Super AI vẫn còn là một ẩn số. Chưa điều gì có thể khẳng định về sự phát triển và tồn tại của nó. Super AI hiện nay