

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT

BÁO CÁO TỔNG KẾT
ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG
TỐI ƯU HÓA MÔ HÌNH DỰ BÁO ĐỘ BỀN CỦA
BÊ TÔNG TÍNH NĂNG CAO

Mã số: T2023-06-08

Chủ nhiệm đề tài: ThS. Lê Thị Thùy Linh

Đơn vị: Khoa Sư phạm Công nghiệp

Chương trình đào tạo: Kỹ thuật xây dựng

Đà Nẵng, 6/2025

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT

BÁO CÁO TỔNG KẾT
ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG

TỐI ƯU HÓA MÔ HÌNH DỰ BÁO ĐỘ BỀN CỦA
BÊ TÔNG TÍNH NĂNG CAO

Mã số: T2023-06-08

Xác nhận của cơ quan chủ trì đề tài

KT. HIỆU TRƯỞNG
PHÓ HIỆU TRƯỞNG

PGS.TS. Võ Trung Hùng

Chủ nhiệm đề tài

ThS. Lê Thị Thùy Linh

DANH SÁCH THÀNH VIÊN THAM GIA NGHIÊN CỨU

STT	Họ và tên	Đơn vị công tác và lĩnh vực chuyên môn
1	Trần Lê Nhật Hoàng	Khoa Sư phạm công nghiệp, Trường Đại học Sư phạm Kỹ thuật, Đại học Đà Nẵng – chuyên ngành Kỹ thuật điện
2	Trương Thị Thu Hà	Khoa Kỹ thuật Xây dựng, Trường Đại học Sư phạm Kỹ thuật, Đại học Đà Nẵng – chuyên ngành Xây dựng

MỤC LỤC

PHẦN MỞ ĐẦU.....	1
CHƯƠNG I. TỔNG QUAN NGHIÊN CỨU	2
1.1. Ngành công nghiệp xây dựng và bê tông hiệu suất cao.....	2
1.1.1. Vai trò ngành công nghiệp xây dựng.....	2
1.1.2. Bê tông hiệu suất cao và vai trò của nó trong xây dựng.....	3
1.1.3. Các ứng dụng của HPC trên thế giới	7
1.2. Ứng dụng kỹ thuật học máy trong dự báo cho HPC.....	9
1.2.1. Công nghệ trí thông minh nhân tạo.....	9
1.2.2. Ứng dụng kỹ thuật học máy trong dự báo cho HPC.....	11
1.3. Các thuật toán tối ưu hóa dựa trên kỹ thuật học máy	14
1.3.1 Kỹ thuật học máy	14
1.3.2 Các thuật toán tối ưu hóa dựa trên kỹ thuật học máy	18
CHƯƠNG II. PHƯƠNG PHÁP NGHIÊN CỨU	22
2.1. Các mô hình đơn và mô hình tổng hợp.....	22
2.2. Thuật toán tối ưu hóa để điều chỉnh các siêu tham số	25
2.4. Mô hình dự báo lai ED-SEM.....	27
2.5. Đánh giá hiệu suất dự báo.....	33
2.6. Phần mềm lập trình.....	36
CHƯƠNG III. THU THẬP DỮ LIỆU VÀ KẾT QUẢ THỰC NGHIỆM.....	39
3.1 Thu thập dữ liệu.....	39
3.2 Xử lý dữ liệu và kết quả thực nghiệm.....	46
KẾT LUẬN VÀ KIẾN NGHỊ	54

DANH MỤC BẢNG BIỂU

Số hiệu	Tên bảng	Trang
Bảng 3.1	Nguồn dữ liệu	39
Bảng 3.2	Các thành phần chính của HPC trong bộ dữ liệu	39
Bảng 3.3	Thống kê mẫu dữ liệu	44
Bảng 3.4	Cài đặt tham số cho các mô hình LSSVR, RBFNN, LSSVR (RBFNN, LSSVR) và ED-SEM.	46
Bảng 3.5	Cài đặt tham số cho các thuật toán tối ưu hóa nổi tiếng để tối ưu hóa SEM	47
Bảng 3.6	Hiệu suất dự đoán của mô hình đề xuất	48
Bảng 3.7	Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 1.	50
Bảng 3.8	Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 2.	51
Bảng 3.9	Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 3.	51

DANH MỤC CÁC HÌNH

Số hiệu	Tên bảng	Trang
Hình 1.1	Tòa nhà Burj Khalifa - Dubai, UAE	8
Hình 1.2	Đập Tam Hiệp - Trung Quốc	9
Hình 1.3	Cầu Mỹ Thuận - Việt Nam	9
Hình 2.1	Cấu trúc tổng thể của mô hình tổ hợp	24
Hình 2.2	Sơ đồ mô hình tổ hợp xếp chồng (SEM)	25
Hình 2.3	Phân loại các thuật toán siêu thuật toán lấy cảm hứng từ thiên nhiên	26
Hình 2.4	Vòng lặp phát triển của một công ty	27
Hình 2.5	Mô phỏng bước cấu trúc	29
Hình 2.6	Mô hình dự báo lai ED-SEM	31
Hình 2.7	Pseudocode của ED-SEM.	33
Hình 2.8	K-fold Cross Validation	34
Hình 2.9	Giao diện MATLAB	39
Hình 3.1	Ma trận tương quan tính năng của bộ dữ liệu 1	42
Hình 3.2	Ma trận tương quan tính năng của bộ dữ liệu 2	43
Hình 3.3	Ma trận tương quan tính năng của bộ dữ liệu 3	44

DANH MỤC CÁC CHỮ VIẾT TẮT

ACO	High-Performance Concrete
AI	Artificial intelligence
ANN	
BPNN	Teaching-Learning-Based Optimization
CPU	ant colony optimization
DM	Artificial Neural Network
ED	Back Propagation Neural Network
ED-SEM	Central processing unit
GA	Data mining
GA	Enterprise Development Optimizer
GANN	Enterprise Development Optimizer Ensemble Model
GRNN	Genetic algorithm
HPC	Genetic Algorithm
KNN	Genetic neural network
LSSVR	Generalized neural network
MAE	K-nearest neighbor
MAPE	Least squares support vector regression
MFA	Mean absolute error
ML	Mean absolute percentage error
MSE	Meta Firefly
PSO	Machine Learning
R	Mean square error
RBF	Particle Swarm Optimization
RBFNN	Correlation coefficient

RF	Radial Basis Function
RMSE	Radial basis function neural network
SA	Random Forest
SCMs	Root Mean Squared Error
SEM	simulated annealing
SOS	Supplementary Cementitious Materials
SVM	Symbiotic Organisms Search
SVR	Support VectorMachine
TLBO	Support Vector Regression
TS	Tabu search
VAR	vector autoregressive

PHẦN MỞ ĐẦU

Trong bối cảnh ngành xây dựng đang ngày càng phát triển mạnh mẽ cả về quy mô lẫn chất lượng, yêu cầu về đặc tính kỹ thuật, độ bền và tính bền vững của các công trình ngày càng trở nên khắt khe hơn. Điều này đặt ra nhu cầu cấp thiết phải ứng dụng những loại vật liệu xây dựng tiên tiến, có khả năng đáp ứng được các tiêu chuẩn cao về kỹ thuật, hiệu quả kinh tế và bảo vệ môi trường. Trong số đó, bê tông hiệu suất cao (High-Performance Concrete – HPC) đã và đang khẳng định được vai trò thiết yếu trong việc nâng cao chất lượng và độ bền của các công trình xây dựng hiện đại.

Với các đặc tính vượt trội như cường độ chịu nén cao, khả năng chống thấm tốt, kháng hóa chất mạnh và độ bền lâu dài trong điều kiện môi trường khắc nghiệt, HPC cho phép thiết kế các kết cấu bền vững hơn, hiệu quả hơn và tiết kiệm chi phí thi công, bảo trì trong suốt vòng đời công trình. Ngoài ra, việc sử dụng HPC còn góp phần tích cực trong việc hiện thực hóa các mục tiêu xây dựng xanh thông qua việc tận dụng phụ gia khoáng và vật liệu tái chế, giảm thiểu phát thải khí nhà kính và bảo vệ môi trường.

Trong các tính chất cơ học của HPC, cường độ chịu nén được xem là chỉ số quan trọng hàng đầu, phản ánh khả năng chịu tải của bê tông và là cơ sở thiết yếu để đánh giá chất lượng vật liệu, thiết kế kết cấu và kiểm soát thi công. Việc xác định cường độ nén một cách chính xác không chỉ giúp đảm bảo an toàn cho công trình mà còn tối ưu hóa thiết kế và tiết kiệm chi phí. Tuy nhiên, quá trình xác định cường độ nén thông qua thí nghiệm thực nghiệm thường tốn nhiều thời gian, chi phí và tài nguyên.

Do đó, việc nghiên cứu và phát triển các mô hình dự báo chính xác cường độ nén của HPC là hết sức cần thiết và có ý nghĩa thực tiễn sâu sắc. Nghiên cứu này đề xuất một tích hợp mô hình máy học Ensemble với trình tối ưu hóa Metaheuristic để dự đoán cường độ nén của hỗn hợp bê tông hiệu suất cao. Mô hình này là công cụ đầy hứa hẹn hỗ trợ cho các phương pháp thử nghiệm truyền thống, hỗ trợ kỹ sư và nhà thiết kế trong việc đánh giá nhanh hiệu suất vật liệu, từ đó rút ngắn thời gian thiết kế và thi công.

CHƯƠNG I. TỔNG QUAN NGHIÊN CỨU

1.1. Ngành công nghiệp xây dựng và bê tông hiệu suất cao

1.1.1. *Vai trò ngành công nghiệp xây dựng*

Ngành xây dựng là một trong những ngành công nghiệp trọng yếu, giữ vai trò nền tảng và đóng góp thiết yếu trong tiến trình công nghiệp hóa, hiện đại hóa của mọi quốc gia. Về bản chất, đây là ngành tổng hợp bao gồm toàn bộ các hoạt động liên quan đến khảo sát địa hình, thiết kế kỹ thuật, lập kế hoạch, thi công và hoàn thiện các công trình cơ sở hạ tầng kỹ thuật và xã hội. Các công trình này có thể bao gồm nhà ở, trường học, bệnh viện, nhà máy, trung tâm thương mại, sân bay, hệ thống giao thông (đường bộ, đường sắt, cầu, hầm), hệ thống cấp - thoát nước, công trình thủy điện, nhiệt điện, năng lượng tái tạo và hàng loạt hạ tầng khác phục vụ trực tiếp cho sự phát triển kinh tế và nâng cao chất lượng sống của người dân.

Công trình xây dựng là sản phẩm cuối cùng của các hoạt động kỹ thuật, quản lý và tài chính phức tạp, mang tính đa ngành. Chúng có thể tồn tại dưới dạng các kết cấu kỹ thuật đơn lẻ hoặc được tổ chức thành hệ thống hoàn chỉnh, liên kết với nhau để tạo nên một mạng lưới hạ tầng phục vụ phát triển vùng, khu đô thị hoặc quốc gia. Để hoàn thiện một công trình xây dựng đòi hỏi sự phối hợp giữa nhiều yếu tố: kiến thức chuyên môn kỹ thuật, kinh nghiệm thực tiễn, kỹ năng quản lý dự án, hiểu biết về pháp luật xây dựng, đánh giá tác động môi trường và khả năng tối ưu nguồn lực tài chính.

Không chỉ đóng vai trò như một trụ cột của nền kinh tế quốc dân, ngành xây dựng còn có tác động lan tỏa mạnh mẽ đến hầu hết các lĩnh vực sản xuất và dịch vụ khác. Theo thống kê của Tổng cục Thống kê Việt Nam, năm 2023, ngành xây dựng đóng góp khoảng 5,94% vào GDP cả nước, đồng thời là một trong những ngành có tốc độ tăng trưởng ổn định, với mức tăng trưởng trung bình từ 7–9%/năm trong giai đoạn 2016–2022. Ngành này cũng thúc đẩy sự phát triển của các ngành công nghiệp phụ trợ như vật liệu xây dựng (chiếm trên 60% đầu vào công trình), cơ khí, hóa chất, logistics, và đặc biệt là bất động sản – lĩnh vực đang chiếm tới 20–25% tổng đầu tư toàn xã hội.

Tại Việt Nam, ngành xây dựng hiện đang sử dụng trực tiếp và gián tiếp khoảng 4–5

triệu lao động, chiếm gần 8% tổng lực lượng lao động toàn quốc. Đây là lĩnh vực có khả năng tạo việc làm cao nhất trong nhóm ngành công nghiệp, từ kỹ sư, kiến trúc sư, kỹ thuật viên, công nhân lành nghề đến lao động phổ thông. Người lao động có thể tham gia làm việc trong các công ty xây dựng, nhà thầu, đơn vị tư vấn, các doanh nghiệp cung ứng thiết bị và vật tư, hay cơ quan quản lý nhà nước về đô thị – hạ tầng. Ngành này cũng góp phần quan trọng trong xóa đói giảm nghèo, đặc biệt tại các khu vực nông thôn, nơi các công trình xây dựng cơ bản như đường, trường, trạm được triển khai mạnh mẽ nhờ các chương trình đầu tư công.

Sự phát triển hạ tầng xây dựng là nền tảng không thể thiếu để thu hút vốn đầu tư nước ngoài (FDI), mở rộng sản xuất, tăng cường kết nối giao thương và cải thiện năng lực cạnh tranh quốc gia. Năm 2023, tổng vốn đầu tư công được phân bổ cho xây dựng hạ tầng tại Việt Nam đạt hơn 700.000 tỷ đồng, chiếm gần 30% tổng chi đầu tư ngân sách nhà nước, phản ánh tầm quan trọng mà Chính phủ dành cho lĩnh vực này trong chiến lược phát triển bền vững.

1.1.2. Bê tông hiệu suất cao và vai trò của nó trong xây dựng

a/ Bê tông.

Bê tông là hỗn hợp nguyên vật liệu được tạo thành từ các cốt liệu cơ bản trong xây dựng bao gồm cát, sỏi, xi măng và nước với tỉ lệ nhất định. Các cốt liệu này sau khi được đảo đều, khuấy đều hỗn hợp từ máy trộn hay xe trộn, trạm trộn sẽ tạo thành hỗn hợp đồng nhất có độ bền, độ kết dính cao cung cấp cho các công trình xây dựng. Vì vậy có thể nói bê tông là một trong các vật liệu xây dựng được sử dụng rộng rãi và quan trọng tại các công trình. Đặc biệt khi hỗn hợp này cứng lại, nó tạo thành một vật liệu rắn chắc, có khả năng chịu lực tốt và độ bền cao. Bê tông được dùng để xây dựng nền móng, tường, cột, và các cấu trúc khác trong xây dựng [1].

Các đặc tính cơ bản của bê tông:

- Ưu điểm:
 - + Độ chịu lực rất cao.
 - + Tính ổn định, bền vững theo thời gian.

- + Tạo hình phong phú, đa dạng (đối với bê tông tươi).
- + Giá thành khá rẻ

Chính nhờ những ưu điểm này bê tông thường được sử dụng cho các cấu trúc nền móng và kết cấu chịu lực của công trình như cột, dầm, sàn, tường. Chính đặc tính chịu nén cao của bê tông giúp nó chịu được tải trọng lớn, giúp cho các công trình trở nên vững chắc và bền bỉ.

Bê tông có thể được đổ vào các khuôn với nhiều hình dạng và kích thước khác nhau. Phù hợp với các thiết kế đa dạng của công trình từ nhà ở cơ bản, cầu đường đến các tòa nhà cao tầng và cơ sở hạ tầng phức tạp. Bên cạnh đó bê tông có khả năng chống cháy tốt, không cháy và không phát sinh khói độc khi gặp lửa. Điều này làm tăng mức độ an toàn cho các công trình xây dựng. Đặc biệt là các tòa nhà cao tầng và khu vực công cộng.

Nhờ những đặc tính vượt trội này, bê tông đã trở thành vật liệu xây dựng phổ biến. Không thể thiếu trong ngành xây dựng hiện đại.

- Bên cạnh đó, bê tông cũng tồn tại một số nhược điểm so với các vật liệu khác:
 - + Khối lượng riêng nặng.
 - + Khả năng cách âm, cách nhiệt, chống ăn mòn kém.

Trong bối cảnh của sự nóng lên toàn cầu và cạn kiệt tài nguyên trên thế giới, đòi hỏi đó sự phát triển mạnh mẽ của nhu cầu về ngành xây dựng và công nghiệp, việc cải tiến và tìm kiếm những vật liệu thay thế nhằm đáp ứng bối cảnh mới là việc hết sức cần thiết.

b/ Bê tông hiệu suất cao và vai trò của nó trong ngành xây dựng.

Bê tông hiệu suất cao hay còn gọi là bê tông tính năng cao, viết tắt là HPC (High Performance Concrete), là một trong những bước tiến quan trọng trong ngành vật liệu xây dựng hiện đại. Đây là loại bê tông được thiết kế với tính năng vượt trội so với bê tông thông thường, cả về khả năng chịu lực, độ bền lâu dài lẫn hiệu suất thi công. HPC không chỉ đơn thuần là vật liệu xây dựng, mà còn là giải pháp kỹ thuật tiên tiến.

HPC đang ngày càng được sử dụng rộng rãi không chỉ ở các công trình quy mô lớn như đường sắt cao tốc, cầu, hầm, nhà cao tầng, sân bay mà còn trong các thiết kế kiến trúc hiện đại đòi hỏi yếu tố thẩm mỹ cao và chống lại các điều kiện môi trường khắc nghiệt. HPC không chỉ mang lại độ bền vượt trội mà còn đảm bảo yếu tố bền vững về môi trường, giảm chi phí bảo trì và tuổi thọ công trình kéo dài hơn nhiều so với các vật liệu truyền thống. Với các tính năng vượt trội so với bê tông thông thường, HPC không chỉ cần hỗn hợp bốn thành phần, cụ thể là xi măng portland, nước, cốt liệu mịn và cốt liệu thô, mà còn sử dụng thêm hai thành phần bổ trợ khác, cụ thể là phụ gia khoáng (ví dụ: tro bay, xỉ lò cao, khói silic) và phụ gia hóa học (chất siêu dẻo). Do đó, sự khác biệt chính giữa HPC và bê tông thông thường về cơ bản là việc sử dụng phụ gia khoáng và hóa học [2].

Sự khác biệt cơ bản giữa bê tông thông thường và bê tông hiệu suất cao chủ yếu nằm ở việc sử dụng các loại phụ gia hóa học và phụ gia khoáng trong thành phần hỗn hợp. Những phụ gia này không chỉ cải thiện đặc tính cơ lý của bê tông mà còn góp phần nâng cao độ bền lâu dài và khả năng thi công [3].

Phụ gia hóa học, đặc biệt là chất siêu dẻo (superplasticizer), được sử dụng để giảm đáng kể tỷ lệ nước/xi măng (w/c) mà vẫn duy trì tính công tác tốt. Việc giảm hàm lượng nước trong hỗn hợp dẫn đến giảm độ rỗng trong cấu trúc bê tông đã đông cứng, qua đó cải thiện đáng kể tính chất cơ học và độ bền. Tuy nhiên, việc sử dụng liều lượng lớn chất siêu dẻo để đạt được tỷ lệ nước rất thấp có thể dẫn đến các vấn đề kỹ thuật, chẳng hạn như giảm tính ổn định hỗn hợp hoặc tăng nguy cơ tách nước. Hơn nữa, hiệu quả của chất siêu dẻo phụ thuộc vào nhiệt độ môi trường, đặc điểm khoáng hóa của xi măng và độ mịn của xi măng – những yếu tố có thể làm thay đổi đáng kể khả năng phân tán hạt xi măng và tương tác giữa các pha rắn.

Phụ gia khoáng – còn được gọi là vật liệu thay thế xi măng (Supplementary Cementitious Materials – SCMs) như tro bay (fly ash), xỉ lò cao, muội silic, hoặc metakaolin – hoạt động như chất puzolan và chất độn mịn, góp phần lấp đầy các khoảng rỗng trong nền xi măng, tạo nên vi cấu trúc đặc hơn và bền hơn. Ở nhiệt độ tiêu chuẩn, các phản ứng puzolan giữa phụ gia khoáng và canxi hydroxit (Ca(OH)_2)

trong quá trình hydrat hóa diễn ra với tốc độ tương đối chậm. Tuy nhiên, mức độ hoạt tính puzolan tăng lên đáng kể nếu vật liệu có độ mịn cao và cấu trúc hóa học đồng nhất.

Trong bối cảnh các công trình yêu cầu độ bền lâu dài, việc sử dụng phụ gia khoáng như tro bay hoặc xỉ lò cao là một giải pháp hữu hiệu, vì chúng giúp làm giảm tốc độ phát triển nhiệt thủy hóa, giảm ứng suất nhiệt nội tại, và tăng cường tính ổn định thể tích trong giai đoạn đầu. Ngoài ra, do phần lớn phụ gia khoáng là sản phẩm phụ của công nghiệp (ví dụ: tro bay từ nhà máy nhiệt điện, xỉ lò cao từ công nghiệp luyện kim), việc tận dụng các vật liệu này không chỉ mang lại lợi ích kinh tế mà còn góp phần thúc đẩy phát triển bền vững và giảm thiểu phát thải carbon trong ngành xây dựng.

Tuy nhiên, hiệu quả của vật liệu thay thế xi măng trong hỗn hợp bê tông phụ thuộc lớn vào bản chất hóa học, độ mịn, và nguồn gốc của từng loại phụ gia. Do đó, việc lựa chọn và thiết kế thành phần hỗn hợp cần được thực hiện dựa trên điều kiện cụ thể của công trình, cũng như yêu cầu kỹ thuật và môi trường thi công.

Những ưu điểm nổi bật của bê tông tính năng cao [2, 3]:

- Cường độ chịu nén cao: HPC có cường độ nén từ 60 đến 120 MPa, cao hơn nhiều lần so với bê tông thông thường (~20–40 MPa). Điều này cho phép sử dụng HPC ở những kết cấu chịu tải trọng lớn như cầu vượt, sàn siêu mỏng, hoặc các tòa nhà chọc trời.
- Khả năng chịu kéo khi uốn vượt trội: Tính linh hoạt cao của HPC thể hiện ở khả năng chống lại sự nứt gãy dưới tác động uốn cong, rất phù hợp với các công trình có kết cấu phức tạp hoặc phải chịu tải trọng thay đổi thường xuyên.
- Môđun đàn hồi cao: Với giá trị từ $4,5 \times 10^6$ đến $6,0 \times 10^6$ DaN/cm², HPC cho thấy khả năng chống biến dạng tốt dưới tải trọng lâu dài, giúp duy trì hình dáng và độ ổn định kết cấu theo thời gian.
- Cường độ ban đầu phát triển nhanh: Khả năng đạt được cường độ cao trong thời gian ngắn giúp rút ngắn thời gian thi công, tăng tốc độ hoàn thiện dự

án và giảm chi phí cấp pha, giàn giáo.

- Tính thẩm mỹ cao: Bề mặt của bê tông HPC có độ mịn và đồng nhất cao, màu sắc ổn định và dễ tạo hình theo các thiết kế phức tạp. Nhiều loại HPC còn có khả năng tự làm sạch hoặc chống bám bẩn, giúp giữ gìn vẻ ngoài công trình theo thời gian mà không cần bảo trì thường xuyên.
- Dễ dàng thi công, tạo hình: HPC có độ linh hoạt cao trong đổ khuôn và tạo hình, kể cả với các chi tiết mảnh, độ dày nhỏ hoặc hình dạng uốn lượn phức tạp. Điều này đặc biệt quan trọng trong các công trình kiến trúc nghệ thuật, nội thất, hoặc những thiết kế mang tính biểu tượng.
- Khả năng kháng mài mòn và ăn mòn tốt: HPC có độ đặc chắc cao, ít lỗ rỗng, giúp chống lại hiện tượng mài mòn cơ học và các tác nhân ăn mòn như hóa chất, khí hậu biển, môi trường công nghiệp...
- Tính chống thấm và bền bỉ trước môi trường khắc nghiệt: Với độ hút nước rất thấp, hệ số khuếch tán ion thấp và khả năng chống băng giá, HPC là lựa chọn lý tưởng cho các công trình ngoài trời, dưới nước, hoặc trong điều kiện môi trường thay đổi liên tục.

Tóm lại, về bản chất, HPC vẫn là một loại bê tông xi măng truyền thống, nhưng được cải tiến mạnh mẽ nhằm nâng cao các đặc tính như: cường độ chịu nén, độ dẻo dai, tính linh hoạt trong tạo hình, khả năng chống thấm và khả năng chống lại các tác nhân ăn mòn hoặc thời tiết khắc nghiệt.

1.1.3. Các ứng dụng của HPC trên thế giới

Bê tông hiệu suất cao (HPC) là vật liệu xi măng tiên tiến có cường độ cao và độ bền tuyệt vời. Nó có tiềm năng trở thành giải pháp thực tế để cải thiện tính bền vững của các tòa nhà và các thành phần cơ sở hạ tầng khác [4]. Trong hai thập kỷ qua, HPC ngày càng được nhiều quốc gia quan tâm khi ứng dụng trong nhiều lĩnh vực, từ các thành phần xây dựng, cầu, đặc điểm kiến trúc, sửa chữa và phục hồi, các thành phần thẳng đứng như tháp cối xay gió và tháp tiện ích cho đến các ứng dụng trong ngành dầu khí , các công trình ngoài khơi, công trình thủy lực và vật liệu phủ [5]. Trong số tất cả các ứng dụng này, xây dựng đường bộ và cầu là ứng dụng HPC phổ biến nhất.

Việc sử dụng HPC cho cầu và các thành phần cầu có thể được thấy ở nhiều quốc gia bao gồm Úc, Áo, Canada, Trung Quốc, Cộng hòa Séc, Pháp, Đức, Ý, Nhật Bản, Malaysia, Hà Lan, New Zealand, Slovenia, Hàn Quốc, Thụy Sĩ và Hoa Kỳ (Mỹ) [6]. Hầu hết các dự án ở các quốc gia đã đề cập đều được các cơ quan chính phủ thúc đẩy như các dự án trình diễn ban đầu nhằm khuyến khích triển khai thêm. Cả các cơ quan tư nhân và chính phủ hiện đang tăng cường sự chú ý và thúc đẩy nhiều nỗ lực hơn nữa để sử dụng vật liệu tiên tiến và đầy hứa hẹn này.

Quỹ nghiên cứu Đức đã khởi xướng một chương trình nghiên cứu trị giá 12 triệu euro (15 triệu đô la Mỹ) vào năm 2005, bao gồm 34 dự án nghiên cứu tại hơn 20 viện nghiên cứu địa phương [6]. Mục tiêu chính của chương trình này là mở rộng nhận thức về HPC để biến nó thành vật liệu đáng tin cậy, phổ biến, khả thi về mặt kinh tế và được áp dụng thường xuyên. Sáng kiến nghiên cứu lớn về HPC cũng có thể được nhìn thấy ở Hàn Quốc. Năm 2007, Viện Công nghệ Xây dựng Hàn Quốc đã tài trợ cho chương trình Super 200 trị giá 11 triệu đô la Mỹ để nghiên cứu việc sử dụng HPC trong cầu dây văng [7]. Việc giới thiệu HPC tại Malaysia bắt đầu vào năm 2006 bởi Dura Technology (DURA) với ứng dụng cầu đầu tiên vào năm 2010. Trong nhiều năm nghiên cứu và phát triển chuyên sâu, DURA cùng với Sở Công trình Công cộng Malaysia đã hợp tác thiết kế và xây dựng các cây cầu HPC tập trung vào khu vực nông thôn. Sau khi thành lập, DURA đã thống trị thị trường HPC tại Malaysia. Từ năm 2010 đến nay, tổng cộng 90 cây cầu HPC đã được hoàn thành với hơn 20 cây cầu đang trong giai đoạn xây dựng [5].



Hình 1.1. Tòa nhà Burj Khalifa - Dubai, UAE



Hình 1.2. Đập Tam Hiệp - Trung Quốc



Hình 1.3. Cầu Mỹ Thuận - Việt Nam

1.2. Ứng dụng kỹ thuật học máy trong dự báo cho HPC

1.2.1. Công nghệ trí thông minh nhân tạo

Trí tuệ nhân tạo (AI) là một nhánh của ngành khoa học máy tính, hỗ trợ việc nghiên cứu, thiết kế và ứng dụng dựa trên máy tính thông minh. Các phương pháp truyền thống để mô hình hóa và tối ưu hóa các hệ thống cấu trúc phức tạp phải đòi hỏi một lượng lớn tài nguyên điện toán. Trong khi đó, trí tuệ nhân tạo thường có thể cung cấp các giải pháp thay thế hữu hiệu giúp giải quyết hiệu quả các vấn đề trong nhiều lĩnh vực của nhiều ngành khác nhau kể cả khoa học công nghệ, kỹ thuật, kinh tế, môi trường, y học v.v...

Nghiên cứu về trí tuệ nhân tạo đã được phát triển từ năm 1956, khi thuật ngữ Trí thông minh nhân tạo, được sử dụng tại cuộc họp ở Đại học Dartmouth. Trí tuệ nhân tạo được xem là một ngành học toàn diện, được phát triển dựa trên sự tương tác của một số loại ngành, như khoa học máy tính, điều khiển học, lý thuyết thông tin, tâm lý học, ngôn ngữ học và sinh lý thần kinh. Theo tiến sỹ Stuart Russell, một chuyên gia trong lĩnh vực máy tính tại đại học California, Mỹ đã khẳng định trong nghiên cứu của mình và đã phát biểu luận điểm này trong hội thảo khoa học quốc gia về công nghệ tại Mỹ năm 2002 rằng trí tuệ nhân tạo là một nhánh của khoa học máy tính, tham gia nghiên cứu, thiết kế và ứng dụng máy tính thông minh .

Mục tiêu của lĩnh vực này là khám phá cách bắt chước và thực hiện một số chức năng

thông minh của bộ não con người, để mọi người có thể phát triển các sản phẩm công nghệ và thiết lập các lý thuyết liên quan. Trí tuệ nhân tạo đã tạo ra một sự bùng nổ về khoa học vào cuối những năm 1960 đến những năm 1970. Trí tuệ nhân tạo đã thúc đẩy nền khoa học công nghệ phát triển và ngày càng chứng minh tầm quan trọng của mình trong các lĩnh vực của đời sống. Vào những năm 1990, có một bước tiến mới trong nghiên cứu về trí tuệ nhân tạo với sự phát triển của công nghệ mạng, đặc biệt là công nghệ internet quốc tế, nghiên cứu trí tuệ nhân tạo trở nên gần gũi và thực tế hơn [8].

Trí tuệ nhân tạo đã nhận được sự chú ý ngày càng tăng trong những năm gần đây. Sự đổi mới, được thực hiện thông qua Internet, đã đưa AI đến gần hơn với cuộc sống hàng ngày của chúng ta. Những tiến bộ này, bên cạnh sự quan tâm đến các tác động kinh tế xã hội và đạo đức tiềm năng của công nghệ, đưa AI lên hàng đầu trong nhiều cuộc tranh luận đương đại. Đầu tư của các lĩnh vực vào AI đang tăng nhanh từng ngày. Việc phát triển siêu dữ liệu Big Data và việc mở rộng Internet vạn vật (IoT), đã tạo ra một môi trường hoàn hảo cho các ứng dụng và dịch vụ AI mới phát triển. Các ứng dụng dựa trên AI đã có thể nhìn thấy trong chẩn đoán chăm sóc sức khỏe, điều trị nhắm mục tiêu, vận chuyển, an toàn công cộng, robot dịch vụ, giáo dục và giải trí, nhưng sẽ được áp dụng trong nhiều lĩnh vực hơn trong những năm tới. Cùng với Internet, AI thay đổi cách chúng ta trải nghiệm thế giới và có tiềm năng trở thành động cơ mới cho tăng trưởng kinh tế.

Trí thông minh nhân tạo được định nghĩa xa hơn là AI hẹp “narrow AI” hay AI cơ bản “general AI”. AI hẹp mà chúng ta tương tác với ngày nay, được thiết kế để thực hiện các tác vụ cụ thể trong một miền (ví dụ: dịch ngôn ngữ). AI nói chung là giả thuyết và không phải là miền cụ thể, nhưng có thể học và thực hiện các nhiệm vụ ở bất cứ đâu. Những tiến bộ trong AI hẹp, đặc biệt là phát triển các thuật toán và mô hình mới trong một lĩnh vực khoa học máy tính được gọi là học máy.

Trong kỹ thuật học máy, các thuật toán là một chuỗi các hướng dẫn được sử dụng để giải quyết vấn đề. Các thuật toán, được phát triển bởi các lập trình viên để hướng dẫn máy tính trong các nhiệm vụ mới, là các khối xây dựng của thế giới kỹ thuật số tiên.

Các thuật toán máy tính tổ chức một lượng lớn dữ liệu vào thông tin và dịch vụ, dựa trên các hướng dẫn và quy tắc nhất định .

Thay vì lập trình máy tính từng bước, phương pháp này đưa ra các hướng dẫn máy tính cho phép nó học từ dữ liệu mà không cần hướng dẫn từng bước mới của lập trình viên. Điều này có nghĩa là máy tính có thể được sử dụng cho các tác vụ mới, phức tạp không thể lập trình thủ công. Những thứ như ứng dụng nhận dạng ảnh cho người khiếm thị, hoặc dịch hình ảnh thành lời nói.

1.2.2. Ứng dụng kỹ thuật học máy trong dự báo cho HPC

Tốc độ phát triển của các quốc gia trên thế giới ngày càng vượt bậc là điều tất yếu, kéo theo nhu cầu phát triển hạ tầng ngày càng tăng. Đặc biệt như cầu lát lớn trong việc phát triển những công trình có kết cấu phức tạp cũng như có khả năng chống chịu với thời tiết và môi trường làm việc khắc nghiệt. Việc phát triển bê tông có hiệu suất cao nhằm đáp ứng các nhu cầu ngày càng tăng là tất yếu. Tuy nhiên, để có được sự kết hợp đặc biệt của các yêu cầu về hiệu suất và tính đồng nhất, tỷ lệ trộn gần như tối ưu của bê tông hiệu suất cao là rất quan trọng. Tuy nhiên, chưa có bất kỳ thông số kỹ thuật hướng dẫn nào về tỷ lệ trộn của bê tông hiệu suất cao. Do đó, tỷ lệ trộn cần thiết được thu được bằng phương pháp thử và loại trừ dựa trên dữ liệu hiện có và bằng phương pháp dựa trên các thông số kỹ thuật cho hỗn hợp bê tông thông thường. Các phương pháp như vậy để định lượng hỗn hợp đòi hỏi số lượng lớn các hỗn hợp thử nghiệm để lựa chọn sự kết hợp vật liệu mong muốn, trong khi một quy trình định lượng hỗn hợp tốt phải giảm thiểu số lượng các hỗn hợp thử nghiệm và đạt được hỗn hợp kinh tế và thỏa đáng với các đặc tính mong muốn [2].

Đặc biệt, vì nó rất phức tạp, nên việc mô hình hóa cường độ nén của HPC là một thách thức. Mối quan hệ phi tuyến tính cao giữa các thành phần và tính chất bê tông làm phức tạp việc mô hình hóa toán học về cường độ nén khi sử dụng dữ liệu có sẵn. Việc dự đoán sớm cường độ bê tông là điều cần thiết để lên lịch tháo ván khuôn bê tông và gia cố lại bản sàn. Một mô hình dự đoán có thể giúp giảm thời gian và chi phí căng sau bằng cách cung cấp dữ liệu cần thiết cho các nhà thiết kế và kỹ sư kết cấu. Do đó, việc dự đoán chính xác và sớm cường độ bê tông là một vấn đề quan trọng

trong thi công bê tông [20].

Các phương pháp thông thường để dự đoán cường độ nén của bê tông sử dụng phương pháp hồi quy tuyến tính hoặc phi tuyến tính [16]. Các cách tiếp cận này áp dụng một lượng dữ liệu thực nghiệm hạn chế để xác định hệ số chưa biết. Tuy nhiên, nhiều mô hình hồi quy đã chỉ ra rằng việc thu được một phương trình hồi quy chính xác là khó khăn. Hơn nữa, một số yếu tố ảnh hưởng đến cường độ nén của HPC khác với các yếu tố của bê tông thông thường. Do đó, để dự đoán cường độ nén, phân tích hồi quy có thể không phù hợp [14]. Vì các mô hình thông thường không đủ để phân tích các vật liệu phức tạp, phi tuyến tính và không chắc chắn, nên nhiều nghiên cứu đã sử dụng trí tuệ nhân tạo (AI) trong các hệ thống tiến hóa hoặc lai để phát triển các mô hình chính xác và hiệu quả có thể dự đoán cường độ nén của bê tông. Tuy nhiên, không có mô hình nào được chứng minh là vượt trội một cách nhất quán.

Nhiều nhà nghiên cứu đã sử dụng các phương pháp mô phỏng dựa trên cơ học để định lượng cường độ bê tông, chẳng hạn như Rabczuk và Belytschko đã mô hình hóa bê tông cốt thép bằng các phần tử hữu hạn và được kết hợp với phương pháp hạt bằng hệ số nhân Lagrange, kết quả của họ cho thấy sự phù hợp tốt với dữ liệu thực nghiệm [9]. Rabczuk và cộng sự đã đề xuất một phương pháp tiếp cận hai chiều để mô hình hóa sự gãy của các kết cấu bê tông cốt thép trong điều kiện tải tĩnh (tăng dần), phương pháp tiếp cận này được áp dụng cho các dầm bê tông ứng suất trước có cơ chế phá hoại khác nhau và kết quả số của họ được so sánh với dữ liệu thực nghiệm và cho thấy sự phù hợp tốt [10]. Drzymala và cộng sự đã nghiên cứu những thay đổi về tính chất vật lý và cơ học của HPC tiếp xúc với tác động của nhiệt độ cao [11]. Zhao và cộng sự đã phát triển một mô hình dự đoán co ngót đã được phát triển trong đó tỷ lệ và tổng phần trăm tro bay (FA) và xỉ lò cao nghiền mịn (GGBS) được đưa vào. Kết quả thử nghiệm của họ cho thấy việc đưa FA và GGBS vào HPC có thể dẫn đến giảm đáng kể độ co ngót tổng thể nhưng lại tăng độ co ngót tự thân và vai trò của FA và GGBS khác nhau trong quá trình tiến triển độ co ngót [12]. Ngoài ra, một số phương pháp tuyến tính và phi tuyến tính đã được thừa nhận để phân tích mối quan hệ giữa các yếu tố chính, chẳng hạn như xi măng, tro bay, nước, chất siêu dẻo và tuổi thử

nghiệm, có thể ảnh hưởng đến cường độ nén của HPC. Tuy nhiên, các phương pháp này không thể đáp ứng độ chính xác cao do mối quan hệ phi tuyến tính với các thành phần HPC, điều này gây ra khó khăn trong việc tính toán cường độ nén của HPC từ dữ liệu có sẵn.

Học máy đã được phát triển để nâng cao độ chính xác và hiệu quả dự đoán, phương pháp này đã xác nhận các giải pháp tiềm năng để giải quyết các vấn đề kỹ thuật trong thế giới thực. Do đó, nhiều mô hình dựa trên ML đã được phát triển để xây dựng cường độ nén của HPC. Để kể tên một số, Chou và cộng sự đã khảo sát các phương pháp học máy khác nhau để tìm ra cường độ HPC dự đoán dựa trên ML tốt nhất [13]. Nguyen et al., đề xuất các công thức bán kinh nghiệm chung liên quan đến các kỹ thuật phi chiều hóa và tối ưu hóa, mô hình của họ đã được thử nghiệm thành công bằng năm cường độ nén của tập dữ liệu bê tông đối với các mô hình ML có sẵn [14]. Song et al., xây dựng các mô hình cường độ HPC dự đoán, kết quả của họ cho thấy thuật toán đóng bao có hệ số tương quan cao (0,95), trong khi các giá trị của GEP, ANN và DT lần lượt là 0,86, 0,81 và 0,75 [15].

Do đó, các kỹ thuật AI này cần được kết hợp với các thuật toán tối ưu hóa và các mô hình lai. Bui và cộng sự đã phát triển một hệ thống chuyên gia dựa trên mô hình mạng nơ-ron nhân tạo (ANN) kết hợp với thuật toán đom đóm đã sửa đổi (MFA), mô hình đề xuất được xây dựng từ dữ liệu thực nghiệm trong khi MFA được sử dụng để tối ưu hóa một tập hợp các trọng số và độ lệch ban đầu của ANN để cải thiện độ chính xác của kỹ thuật trí tuệ nhân tạo này. Kết quả chỉ ra rằng hệ thống lai MFA-ANN có thể dự đoán tốt hơn các đặc tính bê tông hiệu suất cao [16]. Zhang và cộng sự tối ưu hóa các tham số của mạng nơ-ron truyền ngược (BPNN), hồi quy vector hỗ trợ (SVR), cây hồi quy (RT), rừng ngẫu nhiên (RF), k-nearest neighbor (KNN), hồi quy logistic (LR) bằng tối ưu hóa bầy cục bộ (PSO) để xây dựng cường độ nén HPC [17]. Trong một nghiên cứu khác, Nazari và Sanjayan đã nguyên thủy hóa các tham số của máy vector hỗ trợ để ước tính cường độ nén của polyme, vữa và bê tông [18]. Trong nghiên cứu của họ, năm thuật toán bao gồm thuật toán tối ưu hóa đàn kiến, thuật toán di truyền, thuật toán cạnh tranh để quốc, thuật toán tối ưu hóa đàn ong nhân tạo và thuật

toán tối ưu hóa bầy hạt đã được sử dụng để tối ưu hóa các tham số của máy vector hỗ trợ (SVM).

Các mô hình tổng hợp là các thuật toán học tích hợp nhiều trình học được gọi là trình học cơ sở cũng được gọi là trình học yếu (ví dụ: hồi quy vector hỗ trợ bình phương nhỏ nhất, mạng nơ-ron hàm cơ sở xuyên tâm) để nâng cao hiệu suất, tính ổn định và độ mạnh mẽ của trình học cơ sở. Vì các phương pháp tổng hợp cung cấp độ chính xác dự đoán tốt hơn và tối ưu hơn so với các phương pháp ML đơn lẻ [19]. Độ chính xác dự đoán của mô hình dựa trên ML phụ thuộc rất nhiều vào các siêu tham số của nó. Việc lựa chọn siêu tham số phù hợp để cải thiện khả năng dự đoán là một vấn đề tối ưu hóa đầy thách thức.

1.3. Các thuật toán tối ưu hóa dựa trên kỹ thuật học máy

1.3.1 Kỹ thuật học máy

Kỹ thuật học máy (ML) là một lĩnh vực của trí tuệ nhân tạo, theo đó thuật ngữ này đề cập đến khả năng các hệ thống công nghệ thông tin có thể độc lập tìm giải pháp cho các vấn đề bằng cách nhận ra các mẫu trong cơ sở dữ liệu. Nói cách khác, ML cho phép các hệ thống máy tính nhận dạng các mẫu trên cơ sở các thuật toán và bộ dữ liệu hiện có và để phát triển các khái niệm giải pháp phù hợp. ML là kiến thức nhân tạo được tạo ra trên cơ sở kinh nghiệm.

Sự phát triển của các chương trình máy tính có thể truy cập dữ liệu và sử dụng nó để tự học là những tính năng chính của kỹ thuật học máy. Quy trình học bắt đầu bằng các quan sát hoặc thống kê, chẳng hạn như các ví dụ, hướng dẫn hoặc kinh nghiệm trực tiếp, để tìm kiếm các mẫu trong dữ liệu và đưa ra quyết định nâng cao. Mục đích quan trọng nhất là cho phép máy tính học robot mà không cần sự trợ giúp hay can thiệp của con người và điều chỉnh hành động cho phù hợp.

Kỹ thuật học máy giải thuật các bài toán theo năm bước cơ bản:

- Tìm kiếm, trích xuất và tóm tắt dữ liệu liên quan
- Đưa ra dự đoán dựa trên dữ liệu phân tích
- Tính xác suất cho kết quả cụ thể

- Thích ứng với sự phát triển nhất định một cách tự chủ
- Tối ưu hóa các quy trình dựa trên các mẫu được công nhận

Các thuật toán học máy được giám sát được áp dụng để dự đoán các sự kiện trong tương lai dựa trên những gì đã được học trong quá khứ với dữ liệu mới bằng cách sử dụng các ví dụ được gắn nhãn. Bắt đầu từ nghiên cứu của một tập dữ liệu đào tạo đã biết, một hàm được suy ra để đưa ra dự đoán về giá trị đầu ra được phát triển với thuật toán học tập. Chương trình này có thể cung cấp các mục tiêu sau khi đào tạo đủ cho bất kỳ đầu vào nào. Thuật toán học máy cũng có thể đánh giá đầu ra của nó một cách chính xác, dự định để điều chỉnh mô hình cho phù hợp và hỗ trợ trong việc tìm kiếm lỗi.

Để so sánh, các thuật toán học máy không giám sát được thực hiện khi thông tin được sử dụng để hướng dẫn không được dán nhãn cũng không được phân loại. Học máy không tự giám sát nghiên cứu cách các hệ thống có thể suy ra một chức năng để mô tả một sự sắp xếp che giấu từ dữ liệu không được gắn nhãn. Hệ thống khám phá dữ liệu và có thể rút ra các kết luận từ các bộ dữ liệu để mô tả các cấu trúc ẩn từ dữ liệu chưa được gắn nhãn nhưng không phải lúc nào cũng tìm ra đầu ra đúng. Amazon, Google, Salesforce, Netflix và IBM là một số trong những cái tên nổi tiếng đang đầu tư mạnh vào lĩnh vực này.

Quá trình cơ bản của học máy là cung cấp dữ liệu đào tạo cho một thuật toán học tập. Thuật toán học tập sau đó tạo ra một bộ quy tắc mới, dựa trên các suy luận từ dữ liệu. Đây thực chất là tạo ra một thuật toán mới, chính thức được gọi là mô hình học máy. Bằng cách sử dụng dữ liệu đào tạo khác nhau, cùng một thuật toán học tập có thể được sử dụng để tạo ra các mô hình khác nhau. Ví dụ, cùng một loại thuật toán học tập có thể được sử dụng để dạy cho máy tính cách dịch ngôn ngữ hoặc dự đoán thị trường chứng khoán.

Suy ra các hướng dẫn mới từ dữ liệu là thế mạnh cốt lõi của học máy. Nó cũng nhấn mạnh vai trò quan trọng của dữ liệu: càng có nhiều dữ liệu để đào tạo thuật toán, nó càng học được nhiều hơn. Trên thực tế, nhiều tiến bộ gần đây trong AI không phải là do những đổi mới căn bản trong việc học thuật toán, mà là do lượng dữ liệu khổng lồ

được Internet kích hoạt.

Mặc dù mô hình học máy có thể áp dụng hỗn hợp các kỹ thuật khác nhau, các phương pháp học tập thường có thể được phân loại thành ba loại chung:

- Học có giám sát : Thuật toán học được cung cấp dữ liệu có nhãn và đầu ra mong muốn. Ví dụ, hình ảnh của những con chó được dán nhãn chú chó Wap sẽ giúp thuật toán xác định các quy tắc để phân loại hình ảnh của con chó.
- Học tập không giám sát : Dữ liệu được cung cấp cho thuật toán học tập không được gán nhãn và thuật toán được yêu cầu xác định các mẫu trong dữ liệu đầu vào. Ví dụ: hệ thống đề xuất của một trang web thương mại điện tử nơi thuật toán học phát hiện ra các mặt hàng tương tự thường được mua cùng nhau.
- Học tăng cường : Thuật toán tương tác với một môi trường năng động cung cấp phản hồi về phần thưởng và hình phạt. Ví dụ: xe tự lái được thưởng khi ở trên đường.

Hầu hết mọi ngành công nghiệp đang làm việc với hàm lượng lớn dữ liệu đều nhận ra tầm quan trọng của công nghệ học máy, giúp các tổ chức vận hành hiệu quả hơn hoặc tạo được lợi thế cạnh tranh so với các đối thủ. Ví dụ:

- Các dịch vụ tài chính: Ngân hàng và những doanh nghiệp hoạt động trong lĩnh vực tài chính sử dụng công nghệ Machine Learning với 2 mục đích chính: xác định insights trong dữ liệu và ngăn chặn lừa đảo. Insights sẽ biết được các cơ hội đầu tư hoặc thông báo đến nhà đầu tư thời điểm giao dịch hợp lý. Data mining cũng có thể tìm được những khách hàng đang có hồ sơ rủi ro cao hoặc sử dụng giám sát mạng để chỉ rõ những tín hiệu lừa đảo.
- Chính phủ: Các tổ chức chính phủ hoạt động về an ninh cộng đồng hoặc tiện ích xã hội sở hữu rất nhiều nguồn dữ liệu có thể khai thác insights. Ví dụ, khi phân tích dữ liệu cảm biến, chính phủ sẽ tăng mức độ hiệu quả của dịch vụ và tiết kiệm chi phí. Machine Learning còn hỗ trợ phát hiện gian lận và giảm thiểu khả năng trộm cắp danh tính.

- Chăm sóc sức khỏe: Machine Learning là 1 xu hướng phát triển nhanh chóng trong ngành chăm sóc sức khỏe, nhờ vào sự ra đời của các thiết bị và máy cảm ứng đeo được sử dụng dữ liệu để đánh giá tình hình sức khỏe của bệnh nhân trong thời gian thực (real-time). Công nghệ Machine Learning còn giúp các chuyên gia y tế xác định những xu hướng hoặc tín hiệu để cải thiện khả năng điều trị, chẩn đoán bệnh.
- Kinh doanh: Dựa trên hành vi mua hàng trước đây, các trang web sử dụng Machine Learning phân tích lịch sử mua hàng, từ đó giới thiệu những vật dụng mà bạn có thể sẽ quan tâm và yêu thích. Khả năng tiếp nhận dữ liệu, phân tích và sử dụng những dữ liệu đó để cá nhân hóa trải nghiệm mua sắm (hoặc thực hiện chiến dịch Marketing) chính là tương lai của ngành bán lẻ.
- Dầu khí: Tìm kiếm những nguồn nguyên liệu mới. Phân tích các mỏ dầu dưới đất. Dự đoán tình trạng thất bại của bộ cảm biến lọc dầu. Sắp xếp các kênh phân phối để đạt hiệu quả và tiết kiệm chi phí. Có thể nói, số lượng các trường hợp sử dụng Machine Learning trong ngành công nghiệp này cực kì lớn và vẫn ngày càng mở rộng.
- Vận tải: Phân tích dữ liệu để xác định patterns & các xu hướng là trọng tâm trong ngành vận tải vì đây là ngành phụ thuộc vào khả năng tận dụng hiệu quả trên mỗi tuyến đường và dự đoán các vấn đề tiềm tàng để gia tăng lợi nhuận. Các chức năng phân tích dữ liệu và modeling của Machine Learning đóng vai trò quan trọng với các doanh nghiệp vận chuyển, vận tải công cộng và các tổ chức vận chuyển khác.

Ngoài Microsoft, Google, Facebook, IBM và Amazon, Apple cũng dành nguồn lực tài chính khổng lồ cho việc sử dụng và phát triển hơn nữa của Machine Learning. Siêu máy tính Watson của IBM vẫn là thiết bị nổi tiếng nhất cho Machine Learning. Watson chủ yếu được sử dụng trong lĩnh vực y tế và tài chính. Như đã đề cập, Facebook sử dụng Machine Learning để nhận dạng hình ảnh, Microsoft cho hệ thống nhận dạng giọng nói Cortana, Apple cho Siri. Tất nhiên, Machine Learning cũng được sử dụng tại Google, cả trong lĩnh vực dịch vụ hình ảnh và xếp hạng công cụ tìm kiếm.

Các nhà cung cấp đám mây như Google, Microsoft, Amazon Webservice và IBM hiện đã tạo ra các dịch vụ cho Machine Learning. Với sự giúp đỡ của họ, các nhà phát triển không có kiến thức về Machine Learning cụ thể cũng có thể phát triển các ứng dụng. Các ứng dụng này có thể học hỏi từ một bộ dữ liệu có thể xác định tự do. Tùy thuộc vào nhà cung cấp, các nền tảng này có tên khác nhau:

- IBM: Watson
- Amazon: Amazon Machine Learning
- Microsoft: Azure ML Studio
- Google: Tensorflow

Ngoài các nền tảng được đề cập ở trên, có rất nhiều chương trình nguồn mở miễn phí, chất lượng cao, qua đó Machine Learning đã được tiếp cận với nhiều đối tượng hơn, để bạn với tư cách là nhà phát triển hoặc chuyên gia dữ liệu có thể làm việc với nó.

1.3.2 Các thuật toán tối ưu hóa dựa trên kỹ thuật học máy

Thuật toán là một quá trình tính toán cụ thể, trong đó lấy một hay nhiều giá trị làm đầu vào (input) và cho ra một hay nhiều giá trị kết quả (output). Nói cách khác, thuật toán giống như bản đồ chỉ đường giúp ta giải quyết một vấn đề rõ ràng, cụ thể. Ví dụ, nhiều dòng code để tính dãy Fibonacci là một cách cài đặt của một thuật toán nào đó. Một hàm để cộng hai số cũng được gọi là một thuật toán, mặc dù nó rất đơn giản.

Một số thuật toán, như thuật để tính dãy Fibonacci, khá trực quan và ta có thể suy ra nhờ vào suy luận logic cũng như kỹ năng giải bài. Trong cuộc sống hàng ngày, những việc tưởng chừng như đơn giản như kiểm tra email hay nghe nhạc đều đòi hỏi những thuật toán rất phức tạp.

Tất nhiên, việc thường xuyên gặp những bài toán chưa được nghiên cứu trước đó. Lúc này ta phải tự nghĩ ra thuật mới, hoặc áp dụng thuật cũ một cách sáng tạo hơn. Càng có kiến thức về thuật toán, ta càng có khả năng giải quyết thành công vấn đề. Trong nhiều trường hợp, một vấn đề mới có thể được đưa về một vấn đề cũ hơn mà không cần quá nhiều sức lực, với điều kiện ta phải có kiến thức đủ sâu về vấn đề cũ này.

Ở một switch nhanh, chúng ta cần gửi các gói dữ liệu này càng nhanh càng tốt ở mỗi quảng nghỉ để chúng không ứ đọng và bị bỏ qua. Vậy mục tiêu của thuật toán là gửi càng nhiều gói data càng tốt ở mỗi quảng nghỉ, và gói nào đến trước thì được chuyển đi trước. Hóa ra, một thuật toán có tên "stable matching" có liên hệ trực tiếp tới vấn đề này và hoàn toàn có thể áp dụng, dù nhìn sơ mối tương quan giữa vấn đề và thuật toán có vẻ không rõ ràng. Chỉ có nền tảng thuật toán vững vàng mới giúp ta đi đến lời giải cho những trường hợp như vậy.

Nhiều nhà nghiên cứu sử dụng học máy tiên tiến, được tối ưu hóa với các thuật toán tối ưu hóa mới, để cải thiện tính chính xác của dự đoán của họ. Trong các nghiên cứu gần đây, thuật toán tối ưu hóa có thể đóng một vai trò quan trọng trong việc giải quyết vấn đề kỹ thuật. Phương pháp tối ưu hóa dựa trên thuật toán di truyền lấy cảm hứng lượng tử đã được đề xuất để cải thiện hiệu quả tính toán bằng cách chuyển đổi vấn đề xác định dấu hiệu nhân tố tỷ lệ thành vấn đề tối ưu hóa trong thử nghiệm tác động di động mới

Ngoài ra, trong tin học kỹ thuật dân dụng, các mô hình lai đã giải quyết hiệu quả các vấn đề dự đoán, cung cấp kết quả có thể khái quát tốt. Chẳng hạn, Chou và Phạm [22] đã phát triển phương pháp hồi quy vector hỗ trợ dựa trên thuật toán nhân tạo đom đóm thông minh (SAFCAS) để giải quyết các vấn đề về kỹ thuật dân dụng và quản lý (CEM). Sau đó, họ đã sử dụng SAFCA để tối ưu hóa các tham số của mô hình LS-SVR cho các ứng dụng kỹ thuật khác nhau, liên quan đến vật liệu xây dựng, cấu trúc mặt đường và các tòa nhà tiết kiệm năng lượng. Nghiên cứu của họ cho thấy một mô hình lai dựa trên các kỹ thuật AI có sức mạnh trong việc giải quyết các vấn đề CEM

Tương tự, trong lĩnh vực kỹ thuật kết cấu, Chou và cộng sự [23] đã sử dụng thuật toán SAFCA để dự đoán cường độ chịu cắt trong dầm bê tông cốt thép (RC), bằng cách kết hợp thuật toán thuộc địa đom đóm thông minh nhân tạo (SFA) mới nhất với hỗ trợ bình phương nhỏ nhất hồi quy véc tơ (LS-SVR). Kết quả thực hiện SAFCA chứng minh rằng mô hình lai được đề xuất là vượt trội so với các mô hình khác trong việc dự đoán cường độ chịu cắt của chùm sâu RC và nó có thể được sử dụng bởi các

kỹ sư dân sự trong việc thiết kế cấu trúc chòm sâu RC.

Trong lĩnh vực phân tích dự báo về các giá trị của chuỗi thời gian, Chou và Ngo [24] đã đề xuất một hệ thống hồi quy tối ưu hóa trí thông minh (SIOR) để ước tính cường độ cắt cực đại của đất gia cố sợi phân tán ngẫu nhiên (FRS). Hệ thống SIOR tích hợp một kỹ thuật học máy với thuật toán trí thông minh cải tiến để có được hiệu suất dự đoán và khái quát tốt. Hệ thống SIOR cải thiện đáng kể độ chính xác dự đoán của cường độ cắt cực đại, xét về sai số phần trăm tuyệt đối trung bình, cung cấp độ chính xác dự đoán là 99,7% và giúp các kỹ sư dân sự ước tính cường độ cắt của FRS.

Các phương pháp lai nói chung vượt trội hơn các phương pháp dựa trên các mô hình AI đơn lẻ và đại diện cho một lĩnh vực nghiên cứu đầy triển vọng trong tương lai gần. Các kỹ thuật lai được đề xuất là phương pháp tiềm năng cho lĩnh vực dự báo, bao gồm cả các giá trị làm cơ sở giúp tham khảo để chuẩn bị, phán đoán, lên kế hoạch cho một chiến lược đầu tư lâu dài.

Mặc dù nhiều loại hình học máy khác nhau được sử dụng trong kỹ thuật, nhưng loại gần đây được chú ý nhất trong các nghiên cứu về dự báo là máy véc tơ hỗ trợ (support vector machines – SVMs). SVMs là một tập hợp các phương pháp học có giám sát liên quan đến nhau để phân loại và phân tích hồi quy được xem là một trong những phương pháp học máy hữu ích của công nghệ AI.

Thuật toán SVMs được phát triển bởi Vapnik vào năm 1995 ban đầu là một thuật toán dạng nhị phân. SVMs xây dựng một siêu phẳng hoặc một tập hợp các siêu phẳng trong một không gian nhiều chiều hoặc vô hạn chiều. Để phân loại tốt nhất thì các siêu phẳng nằm ở càng xa các điểm dữ liệu của tất cả các lớp (hàm lẻ) càng tốt, vì lẻ càng lớn thì sai số tổng quát hóa của thuật toán phân loại càng bé. Để sử dụng cho mục đích hồi quy, máy véc tơ hỗ trợ hồi quy ra đời (Support vector regresion - SVR) [25].

Phương pháp hồi quy vector nhỏ nhất (LSSVR) là một cải tiến của SVMs. Các nhà nghiên cứu gần đây đã sử dụng LSSVR, vì LSSVR là một phương pháp học máy mạnh mẽ thay thế chương trình bậc hai tiêu chuẩn trong SVR bằng hệ thống tuyến tính bình phương nhỏ nhất để giảm thời gian tính toán và độ phức tạp của nó.

Tuy nhiên, nhược điểm của mô hình LSSVR là độ chính xác dự báo phụ thuộc đáng kể vào các tham số của mô hình, cụ thể là tham số kiểm soát (regularization parameter C) và tham số hàm nhân (kernel function parameter σ). Một số nghiên cứu đã thực hiện để tối ưu hóa mô hình học máy nói chung và LSSVR nói riêng. Có thể kể ra một số phương pháp lựa chọn tham số cho mô hình LSSVR như dựa vào kinh nghiệm, thử và chọn (try-and-error), thuật toán tìm kiếm lưới (grid search algorithm). Tuy nhiên những phương pháp này tốn nhiều thời gian để thực hiện và hiệu quả không cao.

Trong những năm gần đây, tối ưu hóa dựa trên các thuật toán Metaheuristic được sử dụng phổ biến. Metaheuristic là cách gọi chung cho các giải thuật heuristic trong việc giải quyết các bài toán tổ hợp khó. Metaheuristic bao gồm những chiến lược khác nhau trong việc khám phá không gian tìm kiếm bằng cách sử dụng những phương thức khác nhau và phải đạt được sự cân bằng giữa tính đa dạng và chuyên sâu của không gian tìm kiếm.

Một số ví dụ về metaheuristic như thuật toán di truyền (genetic algorithm - GA), thuật toán luyện kim (simulated annealing - SA), thuật toán tối ưu đàn kiến (ant colony optimization - ACO),... Các metaheuristic này được xây dựng dựa trên kiến thức từ thực tế; chẳng hạn GA ra đời dựa trên ý tưởng từ cơ chế di truyền trong sinh học và tiến trình tiến hóa trong cộng đồng cá thể của một loài; SA xuất phát từ kỹ thuật luyện kim bao gồm việc nung và điều khiển làm nguội các kim loại để đạt được trạng thái năng lượng nhỏ nhất.

Một thuật toán mới được đề xuất vào năm 2024 bởi Chou và Truong gọi là thuật toán Tối ưu hóa Phát triển Doanh nghiệp (ED). ED là một phương pháp tối ưu hóa thông minh độc đáo lấy cảm hứng từ cơ chế tiến hóa của các doanh nghiệp hiện đại. Trong môi trường thị trường năng động, các doanh nghiệp phấn đấu phát triển bền vững thông qua các điều chỉnh chiến lược, tối ưu hóa cấu trúc, nâng cấp công nghệ và hợp tác nhân sự. Thuật toán ED tóm tắt các hành vi doanh nghiệp và quy trình ra quyết định trong thế giới thực này thành một khuôn khổ tối ưu hóa, nhằm tìm ra con đường phát triển tốt nhất cho doanh nghiệp [21]. Cốt lõi của thuật toán ED là một khuôn khổ

tối ưu hóa dựa trên trạng thái cân bằng động. Nó mô phỏng cách các doanh nghiệp cân bằng giữa việc thăm dò thị trường (tìm kiếm toàn cầu) và tích hợp tài nguyên (khai thác cục bộ) trong điều kiện hạn chế về tài nguyên. Việc thăm dò thị trường giúp doanh nghiệp xác định các cơ hội mới và mở rộng phạm vi tiếp cận của họ, trong khi tích hợp tài nguyên tối ưu hóa việc phân bổ các nguồn lực nội bộ để nâng cao hiệu quả hoạt động. Ngoài ra, thuật toán nhấn mạnh vào việc duy trì tính đa dạng của tổ chức thông qua sự hợp tác giữa các phòng ban (tương tác dân số) và các lần lặp lại chiến lược (giữ chân nhân tài), phản ánh bản chất hợp tác và thích ứng của các doanh nghiệp trong thế giới thực.

Trong khi nhiều nghiên cứu đã sử dụng các mô hình đơn lai để giải quyết các khoảng cách kiến thức với tính mới lạ của khoa học, nghiên cứu này là nghiên cứu đầu tiên tích hợp phương pháp tiếp cận mô hình tổng thể với các thuật toán tối ưu hóa siêu thuật toán để phát triển các mô hình dự đoán cho cường độ nén của HPC. Thuật toán được đánh giá bằng nhiều hàm chuẩn cổ điển, các dữ liệu lịch sử được xử lý trước nhằm tối ưu dữ liệu đầu vào. Hiệu suất của mô hình được đánh giá bằng cách kiểm chứng bởi bộ dữ liệu thực tế từ nhà thực nghiệm.

CHƯƠNG II. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Các mô hình đơn và mô hình tổng hợp

2.1.1. Mô hình đơn

a/ Radial basis function neural network

Radial basis function neural network (RBFNN) được công bố bởi Broomhead và Lowe (1988) [26], RBFNN bao gồm ba lớp: lớp đầu vào, lớp ẩn và lớp đầu ra. Các nơ-ron trong lớp đầu vào bằng với số lượng tham số đầu vào. Lớp ẩn bao gồm số lượng RBFN cần xác định và chuyển giao phi tuyến tính từ lớp đầu vào sang lớp ẩn. Lớp đầu ra phản ứng với một nơ-ron và được tạo ra bằng cách chuyển giao tuyến tính từ lớp ẩn. Phương trình (1) trình bày dạng chuẩn của RBFNN.

$$y = \sum_{j=1}^{N_n} w_j \varphi_j(\|x - v_j\|) + b \quad (1)$$

Trong đó $x = (x_1, x_2, \dots, x_u)$ là vector u chiều; b là hạng tử độ lệch; $\|x - v_j\|$ là khoảng cách Euclid giữa tâm của nút ẩn thứ j và một điểm dữ liệu; w_j là trọng số kết nối từ nút ẩn thứ j đến lớp đầu ra; N_n là số lượng nơ-ron ẩn và $\varphi_j(\cdot)$ là biểu thị RBF của nút ẩn thứ j .

b/ Least squares support vector regression

Least squares support vector regression (LSSVR) được đề xuất bởi Suykens và cộng sự (2002) [27], lý thuyết LSSVR giả định rằng tập dữ liệu $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ có thể được biểu diễn dưới dạng hàm phi tuyến tính và hàm quyết định, được đưa ra trong Công thức (2).

$$y(x) = \omega^T \phi(x) + b \quad (2)$$

Trong đó $x \in R^n, y \in R$, và $\phi(x): R^n \rightarrow R^{n^h}$ là phép ánh xạ tới không gian đặc trưng có chiều cao và b là hạng tử thiên vị.

Một bài toán tối ưu hóa bị ràng buộc được sử dụng để xây dựng công thức hàm R công thức (3), trong đó C_1 là hằng số chính quy, phải được người dùng tối ưu hóa và cung cấp trọng số của phần thứ hai so với phần thứ nhất và ε_i là lỗi dữ liệu đào tạo.

$$\text{Minimize: } R(\omega, \varepsilon, b) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} C_1 \sum_{i=1}^n \varepsilon_i^2 \quad (3)$$

Tùy thuộc vào: $y_i = \omega^T \phi(x_i) + b + \varepsilon_i, i = 1, \dots, n$

Hàm mục tiêu này trở nên không thể giải được khi vô hạn. Do đó, phương pháp lập trình tối ưu Lagrange Multiplier, được đưa ra bởi Công thức (2) được sử dụng để xử lý nhiệm vụ này.

Mô hình LSSVR kết quả để ước tính hàm được đưa ra bởi công thức (4)

$$y(x) = \sum_{i=1}^n \xi_i K(x_i, x_k) + b \quad (4)$$

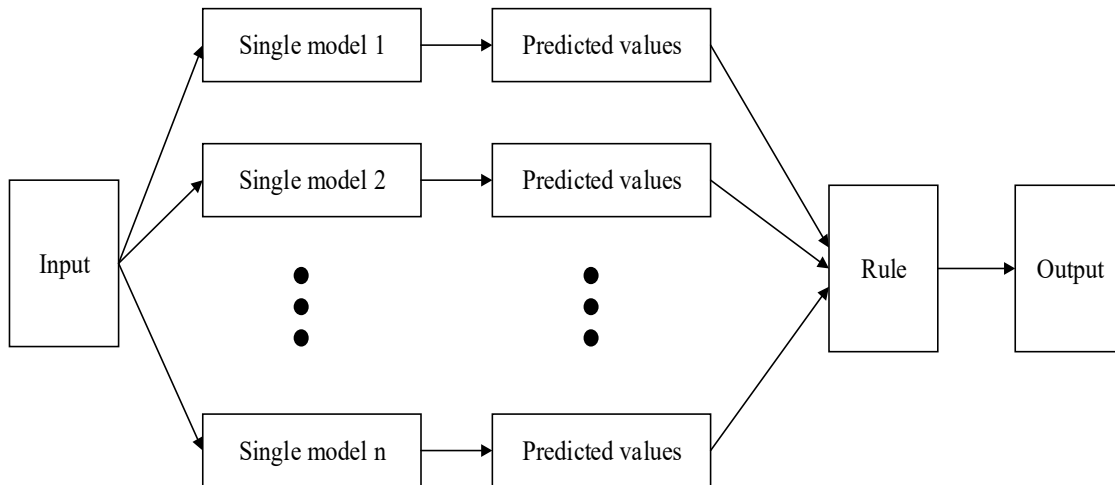
$$K(x_i, x_k) = \exp\left(\frac{\|x_i - x_k\|^2}{2\gamma_1^2}\right) \quad (5)$$

trong đó ξ_i và b là các nghiệm của hệ thống tuyến tính và γ_1 là tham số hàm hạt nhân.

2.1.2. Mô hình hỗn hợp (SEM)

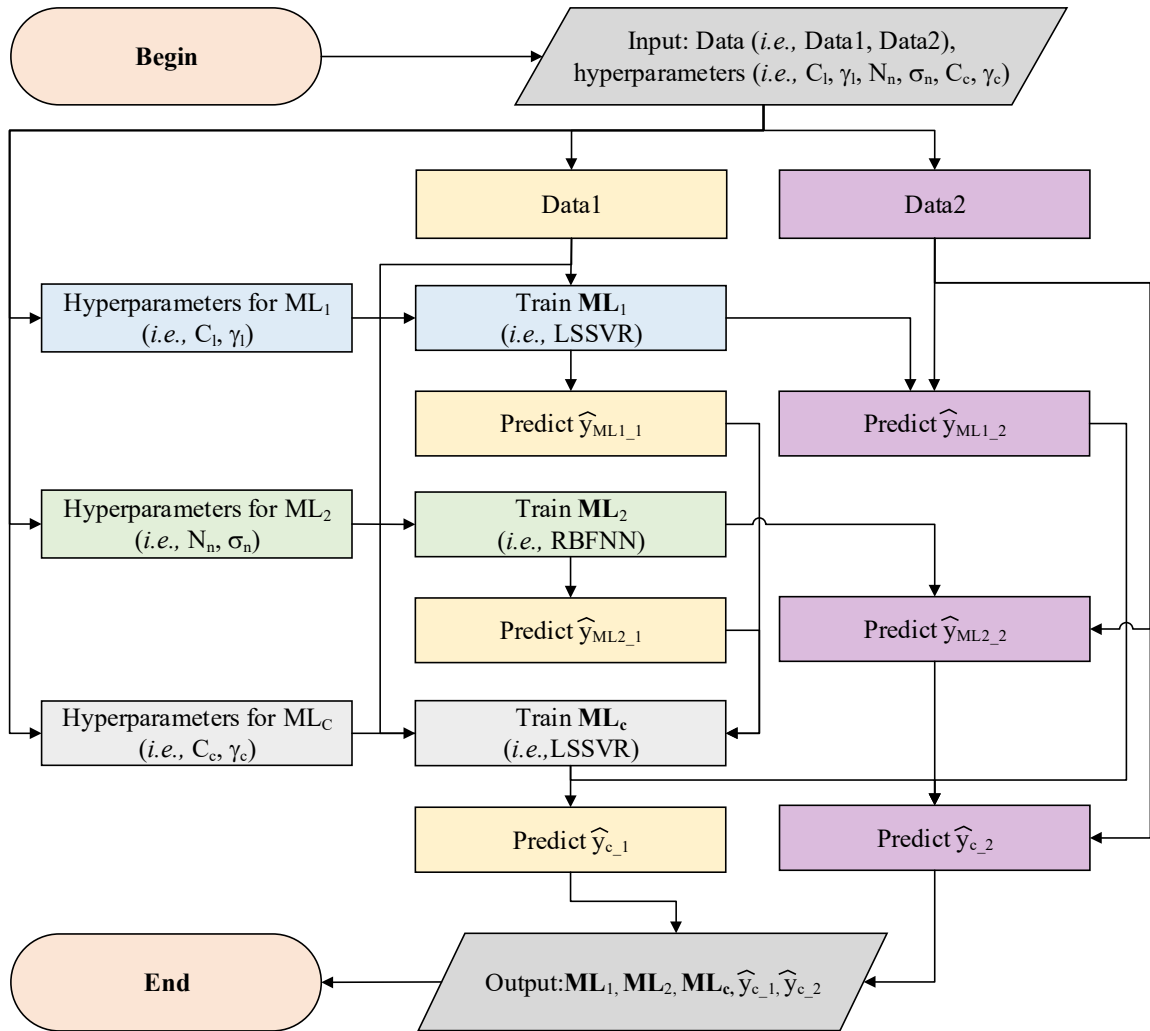
Mô hình hỗn hợp hay mô hình tổ hợp (Ensemble Model - SEM) là một phương pháp trong lĩnh vực học máy và trí tuệ nhân tạo, trong đó nhiều mô hình học máy độc lập (mô hình đơn) được kết hợp lại để tạo thành một mô hình mạnh hơn, ổn định hơn và có khả năng tổng quát hóa cao hơn.

Hình 2.1 mô tả nguyên lý của mô hình tổ hợp được Nisbet và cộng sự [28] áp dụng các quy tắc khác nhau, chúng ta có các mô hình tổ hợp khác nhau và ba bộ học tập tổng hợp thường được sử dụng là đóng gói, chéo và xếp chồng. Trong số đó, xếp chồng đã được áp dụng rộng rãi như một phương pháp tổ hợp với cấu trúc linh hoạt và độ mạnh mẽ cao để tích hợp các kết quả dự đoán của bộ học cơ sở [17].



Hình 2.1. Cấu trúc tổng thể của mô hình tổ hợp [28]

Hình 2.2 trình bày sơ đồ của mô hình tổ hợp xếp chồng. Trong các mô hình dựa trên xếp chồng, tối ưu hóa tham số tăng cường đáng kể hiệu suất dự đoán của các bộ phân loại cơ sở, do đó tăng cường hiệu suất của các mô hình dựa trên xếp chồng (SEM)



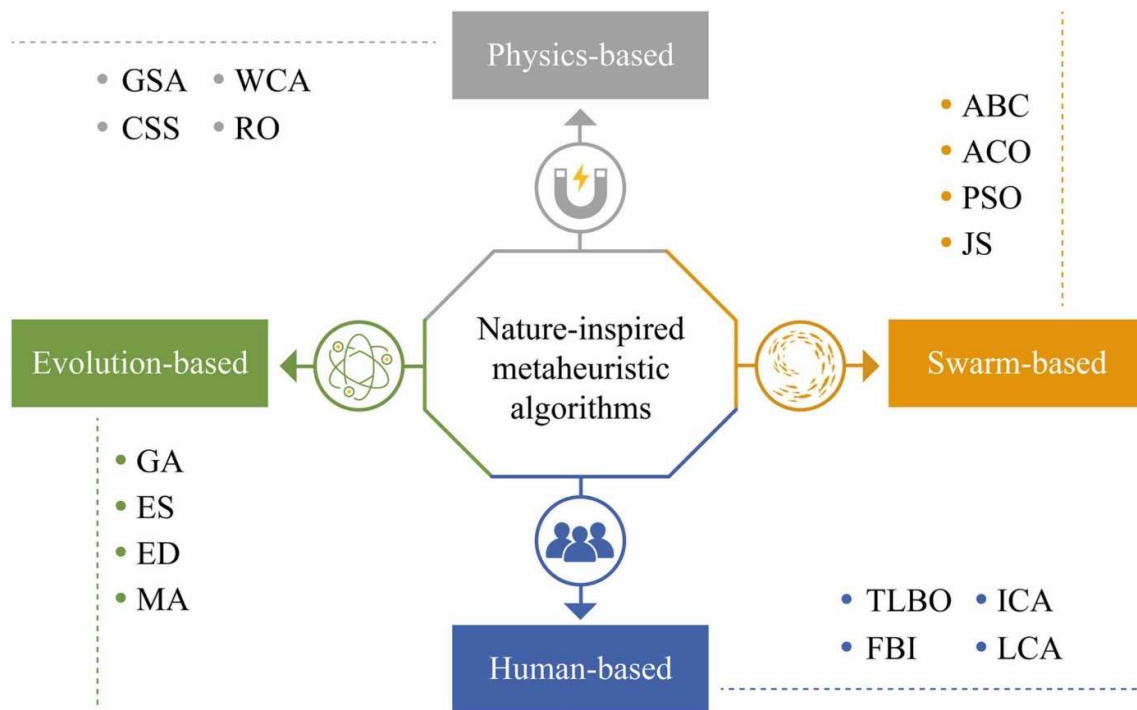
Hình 2.2. Sơ đồ mô hình tổ hợp xếp chồng (SEM)

2.2. Thuật toán tối ưu hóa để điều chỉnh các siêu tham số

Tối đa hóa lợi ích hoặc giảm thiểu tổn thất luôn là mối quan tâm trong các vấn đề kỹ thuật. Đối với các lĩnh vực kiến thức đa dạng, sự phức tạp của các vấn đề tối ưu hóa tăng lên khi khoa học và công nghệ phát triển. Thông thường, các ví dụ về các vấn đề kỹ thuật có thể yêu cầu một phương pháp tối ưu hóa là trong chuyển đổi và phân phối năng lượng, trong thiết kế cơ khí, trong hậu cần và trong việc tải lại các lò phản ứng hạt nhân.

Tối ưu hóa là một nhánh của khoa học toán học trong đó các phương pháp dựa trên độ dốc thường được sử dụng. Tuy nhiên, các vấn đề trong thế giới thực đã trở nên ngày càng phức tạp và không còn có thể dễ dàng giải quyết bằng các phương pháp cổ

điển, các nhà nghiên cứu đã nỗ lực phát triển các chiến lược khác và các thuật toán siêu thuật toán. Các thuật toán này chủ yếu dựa vào cảm hứng tự nhiên. Cảm hứng như vậy có thể đã phát triển trong hàng triệu năm hoặc được thể hiện thông qua các quy luật của tự nhiên. Trong các thuật toán siêu thuật toán, các quy tắc kinh nghiệm được tổ chức sao cho các quy trình được thực hiện ở dạng đơn giản nhất của chúng. Các quy tắc kinh nghiệm này có thể được mô hình hóa để trở thành các phương pháp tính toán hiệu quả bằng cách kiểm tra logic chi phối việc thực hành kinh nghiệm và sử dụng các đặc điểm vốn có của nó. Với những hiện tượng này, các phương pháp dựa trên trí thông minh có thể được sử dụng trong nhiều ứng dụng khác nhau, chẳng hạn như mô phỏng, lập mô hình và tối ưu hóa [21].



Hình 2.3. Phân loại các thuật toán siêu thuật toán lấy cảm hứng từ thiên nhiên [21]

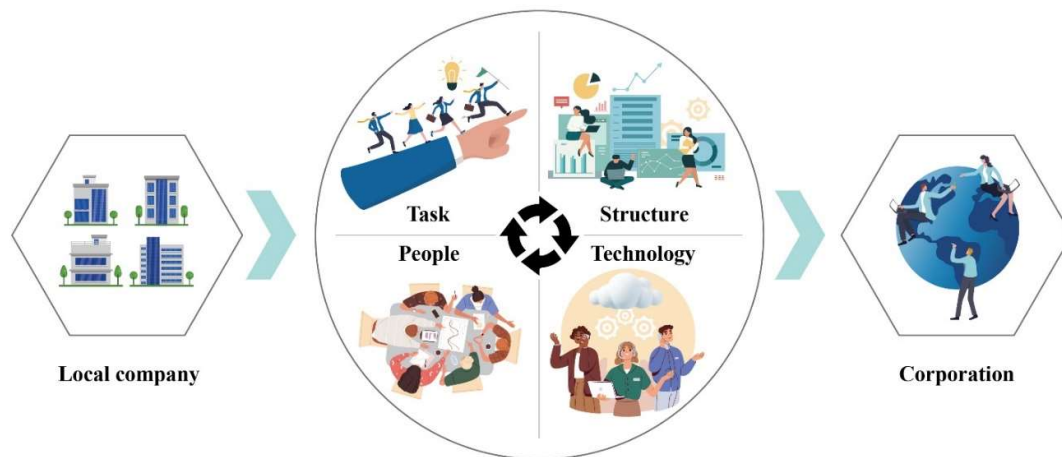
Để tối đa hóa hoặc tối thiểu hóa một chức năng để tìm ra mức tối ưu, có một số cách tiếp cận mà người ta có thể thực hiện. Mặc dù có một loạt các thuật toán tối ưu hóa có thể được sử dụng, không có một thuật toán chính nào được coi là tốt nhất cho mọi trường hợp. Một phương pháp tối ưu hóa phù hợp cho một vấn đề có thể không phải là phương pháp khác; nó phụ thuộc vào một số tính năng, ví dụ, liệu chức năng có khác biệt và độ lõm của nó (lồi hay lõm). Để giải quyết vấn đề, người ta phải hiểu

các phương pháp tối ưu hóa khác nhau để người này có thể chọn thuật toán phù hợp nhất với vấn đề của các tính năng.

2.4. Mô hình dự báo lai ED-SEM

2.4.1. Thuật toán *Enterprise Development Optimizer*

Thuật toán Enterprise Development Optimizer (ED) được Truong và Chou phát triển vào năm 2024 [21]. Thuật toán này lấy cảm hứng từ các tập đoàn tìm kiếm các giải pháp tổ chức thông qua các nhiệm vụ, cấu trúc, công nghệ và hoạt động của con người. Thuật toán ED chuyển đổi giữa các giải pháp này bằng cơ chế chuyển đổi hoạt động. Giai đoạn điều tra bắt đầu khi tổ chức xác định được giải pháp cho một thách thức kinh doanh và thiết lập khuôn khổ của giải pháp đó. Sau đó, tổ chức tập trung vào công nghệ của mình trong các giai đoạn khám phá và khai thác. Cuối cùng, một bước dân số lặp lại được thực hiện để đạt được khai thác (giai đoạn khai thác) và cơ chế chuyển đổi hoạt động xen kẽ giữa bốn loại hoạt động (hoặc giải pháp). Khi vòng lặp được lặp lại, tính tối ưu của tổ chức đạt được, biểu thị giai đoạn tối ưu.



Hình 2.4. Vòng lặp phát triển của một công ty [21].

Ba tính năng mạnh mẽ của ED như sau: (1) nó chỉ có hai tham số bên trong; (2) nó rất dễ dàng để viết mã và (3) nó rất dễ áp dụng. 50 hàm toán học nhỏ / trung bình và 30 hàm toán học quy mô lớn về tính toán tiến hóa (CEC2020) và 24 hàm (CEC2022) được sử dụng để xác thực thuật toán ED. Kết quả tỷ lệ truy cập thành công trên các

hàm nhỏ / trung bình cho thấy rằng ED có thể tìm kiếm các vị trí tối ưu tốt hơn các thuật toán khác.

Đáng chú ý, ED yêu cầu ít thời gian hơn và hội tụ nhanh hơn các thuật toán khác. Đối với các hàm toán học quy mô lớn, ý nghĩa của kiểm định tổng hạng Wilcoxon khẳng định thêm rằng ED có khả năng tìm các giá trị tối ưu trong các hàm quy mô lớn. ED rõ ràng là tốt hơn các thuật toán AVOA (2021), ChOA (2020), GTO (2022), MGO (2022), ROA (2022), SAO (2021) và ba thuật toán chiến thắng trong cuộc thi CEC 2020 bao gồm các thuật toán SASS (2020), sCMaGES (2020) và COLSHADE (2020) và mười thuật toán nổi tiếng ABC (2007), CA (2016), DE (1997), GA (1992), GWO (2014), IWO (2006), PSO (1995), SOS (2014), TLBO (2020) và WOA (2016) trong các vấn đề tối ưu hóa ràng buộc mục tiêu đơn trong thế giới thực. Do đó, ED có khả năng là một thuật toán siêu mô tả tuyệt vời để giải quyết các vấn đề tối ưu hóa kỹ thuật khác nhau.

Để mô phỏng thuật toán ED, ba quy tắc tối ưu phải được thiết lập như sau:

1. Việc thiết lập quy trình ED đòi hỏi bốn hoạt động - nhiệm vụ, cấu trúc, công nghệ và hoạt động của con người - và một cơ chế chuyển đổi giữa các hoạt động này.
2. Các hoạt động có mối tương quan với hiệu suất của tổ chức. Các giải pháp này hấp dẫn hơn đối với các tổ chức có hiệu suất đáng chú ý hơn.
3. Giải pháp và hàm mục tiêu tương ứng đánh giá hiệu suất của tổ chức.

Trong quản lý quy trình kinh doanh, các nhiệm vụ có thể có nhiều hình thức khác nhau hoặc tồn tại dưới dạng thói quen hàng ngày. Để mô phỏng các hoạt động của nhiệm vụ, hoạt động tệ nhất được thay thế bằng:

$$\mathbf{x}_{\text{worst}}(t) = \mathbf{l}_b + \text{rand}(0,1) \times (\mathbf{u}_b - \mathbf{l}_b) \quad (6)$$

trong đó $\mathbf{x}_{\text{worst}}$ là giải pháp tệ nhất trong khi $\mathbf{u}_b$ và $\mathbf{l}_b$ là giới hạn trên và giới hạn dưới của không gian tìm kiếm.

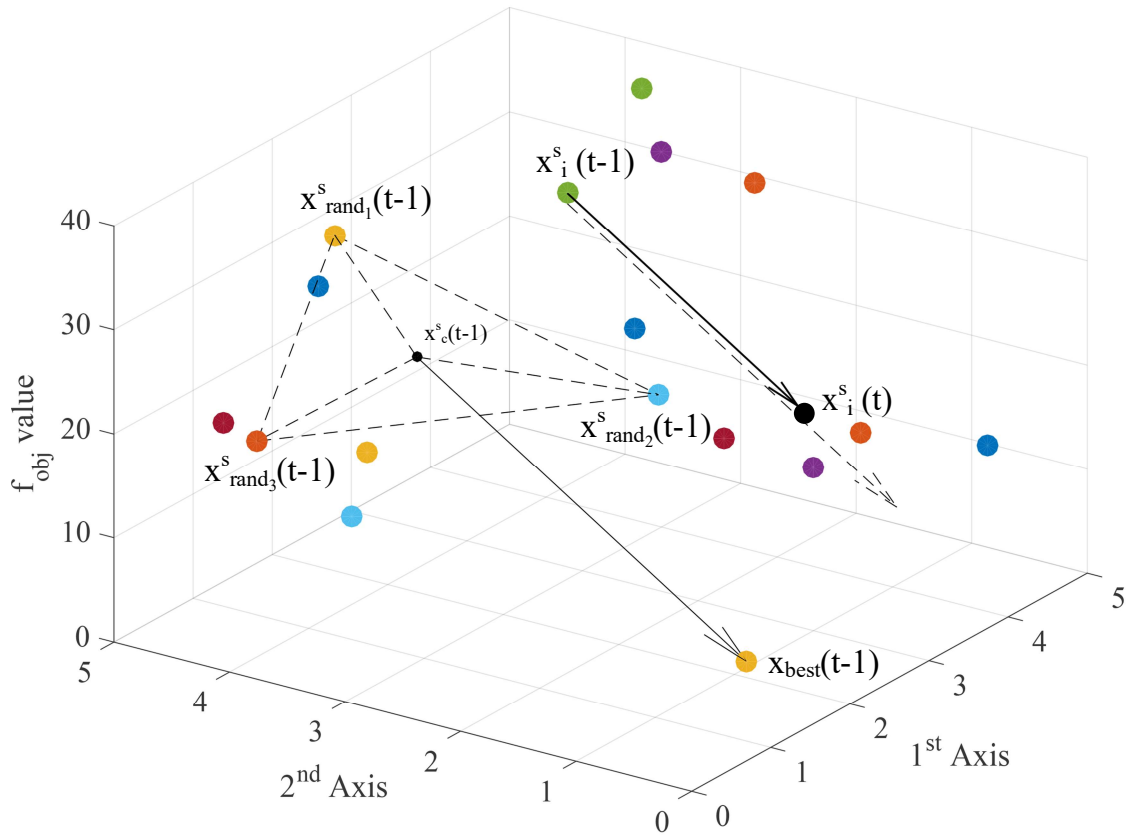
Một trong những nhược điểm của các mô hình cấu trúc cổ điển là thiết kế công việc hầu như hoàn toàn được xác định bởi các biến nhiệm vụ và kỹ thuật và cần phải xem xét các biến xã hội hoặc con người. Vì vậy, việc thiết kế và cấu trúc lại mô hình một

cách phù hợp công việc lý thuyết thông tin theo nghĩa kỹ thuật xã hội ảnh hưởng đến thái độ và đầu ra của con người.

ED có một ưu điểm là giới hạn cấu trúc tổ chức vào quy trình công việc. Một cấu trúc tổ chức mới dự kiến sẽ bị ảnh hưởng bởi cấu trúc của các quy trình công việc khác và quy trình công việc tối ưu hiện tại trong tổ chức, cấu trúc mới được mô phỏng thông qua các phương trình 7 và 8

$$\mathbf{x}_i^s(t) = \mathbf{x}_i^s(t-1) + \text{rand}(-1,1) \times (\mathbf{x}_{\text{best}}(t-1) - \mathbf{x}_c^s(t-1)) \quad (7)$$

$$\mathbf{x}_c^s(t-1) = \frac{\mathbf{x}_{\text{rand}_1}^s(t-1) + \mathbf{x}_{\text{rand}_2}^s(t-1) + \dots + \mathbf{x}_{\text{rand}_m}^s(t-1)}{m} \quad (8)$$



Hình 2.5. Mô phỏng bước cấu trúc.

Việc cân bằng giữa khám phá và khai thác được xem là bước công nghệ ứng dụng quan trọng trong 4 bước của ED

$$\mathbf{x}_i^{\tau}(t) = \mathbf{x}_i^{\tau}(t-1) + \text{rand}^{\alpha}(0,1) \times (\mathbf{x}_{\text{best}}(t-1) - \mathbf{x}_i^{\tau}(t-1)) + \text{rand}^{\beta}(0,1) \times (\mathbf{x}_{\text{best}}(t-1) - \mathbf{x}_{\text{rand}_1}^{\tau}(t-1)) \quad (9)$$

Trong đó $\mathbf{x}_i^s(t)$ là cấu trúc mới, $\mathbf{x}_{\text{best}}(t-1)$ là giải pháp tối ưu hiện tại, $\mathbf{x}_c^s(t-1)$ là trung tâm của các quy trình

Nhân tố con người hay nhân tố nhân sự của công ty là một trong những yếu tố then chốt cho việc vaanh hành công việc. Các tổ chức luôn thúc đẩy văn hóa làm việc có sự tham gia của các cá nhân, thúc đẩy sự sáng tạo và làm việc nhóm của cá nhân thông qua sự tôn trọng của mọi người và các bên liên quan. Văn hóa làm việc như vậy ảnh hưởng đến cam kết và sự tham gia của nhân viên vào tính bền vững. Sự cẩn thận là điều cần thiết cho sự thành công của bất kỳ hệ thống sản xuất hoặc chuỗi cung ứng nào [21]. Giả sử rằng tính năng là một chiều, các phương trình sau đây mô tả cách hoạt động của mọi người được mô phỏng bằng cách chọn ngẫu nhiên tính năng và cập nhật nó

$$\mathbf{x}_{i,d}^p(t) = \mathbf{x}_{i,d}^p(t-1) + \text{rand}(-1,1) \times (\mathbf{x}_{\text{best},d}(t-1) - \mathbf{x}_{c,d}^p(t-1)) \quad (10)$$

$$\mathbf{x}_{c,d}^p(t-1) = \frac{\mathbf{x}_{\text{rand}_1,d}^p(t-1) + \mathbf{x}_{\text{rand}_2}^s(t-1) + \dots + \mathbf{x}_{\text{rand}_m}^s(t-1)}{m} \quad (11)$$

trong đó d là một đặc điểm ngẫu nhiên của mọi người. Đặc điểm này được tính như sau:

$$m = \lceil \text{rand}(0,1) \times n_m \rceil \quad (12)$$

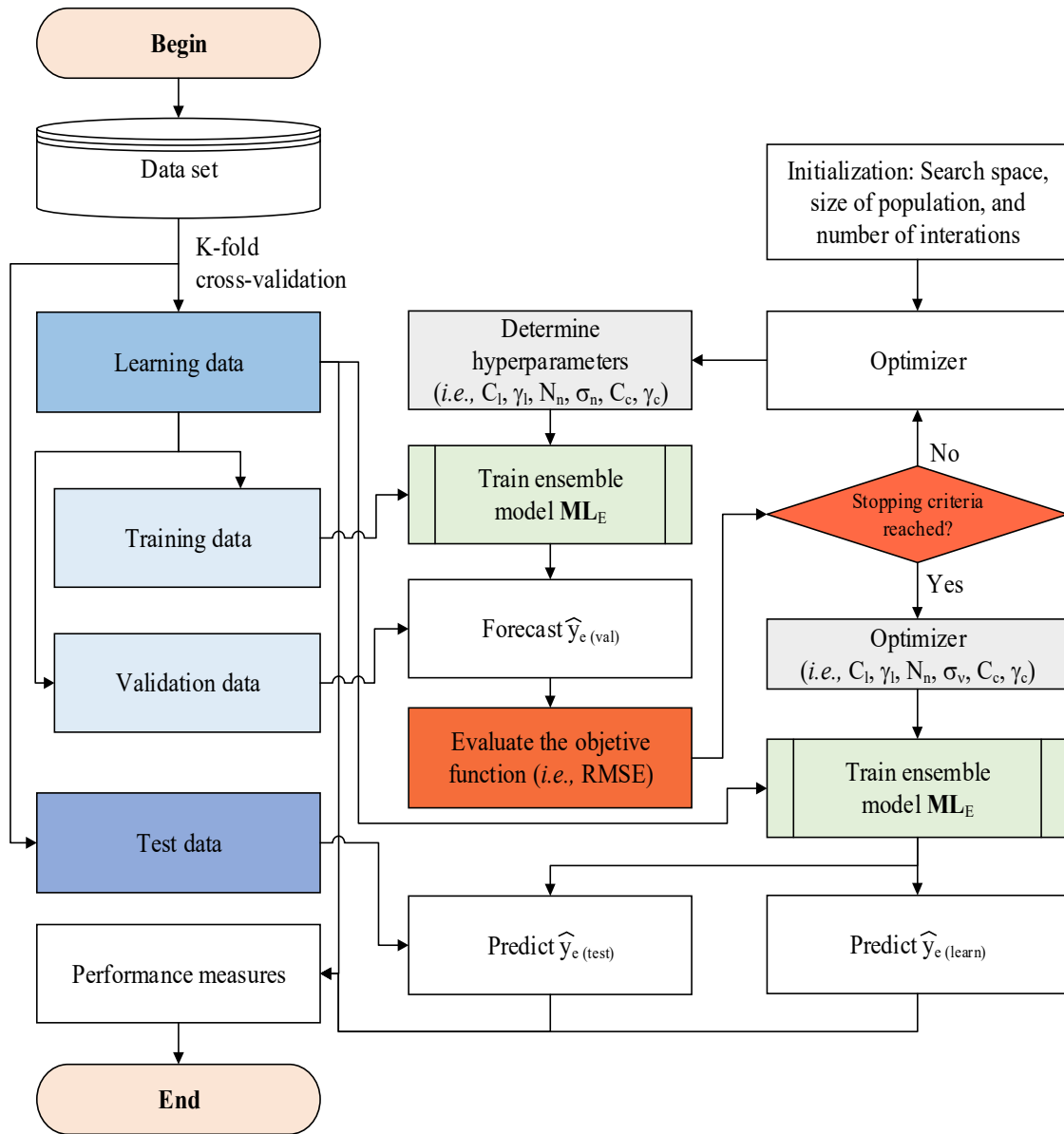
Với m là số lượng người có ảnh hưởng đến cá nhân đang được đề cập, được xác định là mang lại kết quả tối ưu trong thời gian tính toán ngắn và n là số chiều của giải pháp.

Thuật toán ED được đề xuất giả định rằng công ty tập trung vào một bước duy nhất tại một thời điểm. Do đó, chỉ có một trong bốn bước (tức là nhiệm vụ, cấu trúc, công nghệ và con người) xảy ra tại thời điểm t và được kiểm soát bởi cơ chế chuyển đổi hoạt động. Một cơ chế chuyển đổi được thiết kế để hoạt động cho bốn yếu tố chính được giới thiệu là hàm c(t) như trong công thức (13) với Max_{iter} là số lần lặp tối đa.

$$c(t) = \left\lceil 3 \times \left(1 - \frac{t}{\text{Max}_{\text{iter}}}\right) \right\rceil \quad (13)$$

2.4.2. Mô hình dự báo lai ED-SEM

Các mô hình tổng hợp được mô tả và đánh giá bởi Nisbet và cộng sự trong một nghiên cứu chuyên sâu năm 2018 [28]. Trong số đó, mô hình xếp chồng được coi là một phương pháp linh hoạt và mạnh mẽ và được áp dụng rộng rãi trong nghiên cứu liên quan đến phân tích và tối ưu hóa. Các mô hình dựa trên xếp chồng bao gồm nhiều bộ phân loại cơ sở (SEM). Tuy nhiên, việc điều chỉnh các tham số của tất cả các bộ phân loại cơ sở bằng kỹ thuật tối ưu hóa tham số cố định dẫn đến khó khăn trong việc cân bằng hiệu suất của các bộ phân loại cơ sở, ảnh hưởng đến hiệu suất chung của mô hình tổng thể. Do đó, nghiên cứu này đề xuất một mô hình tổng thể xếp chồng lai mới sử dụng trình tối ưu hóa siêu thuật toán để biến các siêu tham số của mô hình tổng thể xếp chồng nhằm nâng cao hiệu suất dự đoán của mô hình tổng thể. Hình 2.6 cho thấy cấu trúc của mô hình tổng thể xếp chồng lai Enterprise Development Optimizer Ensemble Model (ED-SEM).



Hình 2.6. Mô hình dự báo lai ED-SEM

Các siêu tham số của mô hình tổng thể xếp chồng bao gồm C_l, γ_l của ML_1 (LSSVR), N_n, σ_n của ML_2 (RBFNN) và C_c, γ_c của ML_c (LSSVR). Độ chính xác dự đoán của mô hình được đề xuất phụ thuộc rất nhiều vào các siêu tham số này. Việc thiết lập các giá trị $C_l, \gamma_l, N_n, \sigma_n, C_c, \gamma_c$ một cách thích hợp là một bài toán tối ưu hóa, có thể được giải quyết hiệu quả bằng thuật toán ED. Hệ thống được đề xuất tích hợp theo cách này được gọi là mô hình tổng hợp xếp chồng phát triển doanh nghiệp (ED-SEM). Đầu tiên, 90% và 10% dữ liệu được gán cho một tập dữ liệu học và một tập dữ liệu thử nghiệm. Sau đó, 70% dữ liệu học tập được sử dụng để đào tạo ED-

SEM và 30% dữ liệu học tập còn lại được sử dụng để xác thực các siêu tham số của ED-SEM. Hàm mục tiêu của ED-SEM được đề xuất được xác định là

$$f(C_l, \gamma_l, N_n, \sigma_n, C_c, \gamma_c) = \text{MAPE}_{\text{Validation}}^{\text{Training process}} \quad (14)$$

Độ chính xác dự đoán của mô hình phụ thuộc nhiều vào các siêu tham số này. Đặt các giá trị của C , σ một cách thích hợp là một bài toán tối ưu hóa, có thể được giải quyết một cách hiệu quả bằng thuật toán ED. Do đó, hệ thống được đề xuất tích hợp được gọi là hồi quy vector hỗ trợ bình phương nhỏ nhất phát triển doanh nghiệp (ED-SEM). Hình 2.7 cho thấy pseudocode của hệ thống ED-SEM tương ứng.

```

Begin
    Define objective function  $f(\mathbf{x})$ ,  $\mathbf{x}=(x_1, x_2, \dots, x_d)^T$ 
    Set the search space, population size ( $n_{pop}$ ), and maximum iteration ( $Max_{iter}$ )
    Initialize time:  $t = 1$ 
    Initialize population  $\mathbf{x}_i$  ( $i = 1, 2, \dots, n_{pop}$ ) by using Eq. (1)
    Calculate the organization's performance for each  $\mathbf{x}_i$ ,  $f(\mathbf{x}_i)$ 
    Find the organization currently with the best performance ( $X_{best}$ )
    Repeat
        For  $i = 1:n_{pop}$  do
            Calculate  $c(t)$  by using Eq. (9)
            If  $rand < p_1$ , then //  $p_1 = 0.1$ 
                Step "task"
                New task is defined by Eq. (2)
            Else
                Switch  $c(t)$ 
                    Case  $c(t)=1$ 
                        Step "structure"
                        New structure is defined by Eq. (3)
                    Case  $c(t)=2$ 
                        Step "technology"
                        New technology is defined by Eq. (5)
                    Case  $c(t)=3$ 
                        Step "people"
                        New feature of people is defined by Eq. (6)
                End of switch
            End of if
        End of for
        Update the time:  $t = t + 1$ 
    Until the stop criterion is met (e.g.,  $t > Max_{iter}$ )
    Output the best results and visualization (the best organization's performance)
End

```

Hình 1.7. Pseudocode của ED-SEM.

2.5. Đánh giá hiệu suất dự báo

2.5.1. *K-fold Cross Validation*

Trong các bài toán học máy và dự báo thống kê, việc đánh giá hiệu suất mô hình một cách khách quan và đáng tin cậy là một yếu tố then chốt nhằm đảm bảo khả năng tổng quát hóa của mô hình khi áp dụng vào các dữ liệu mới chưa từng thấy. Một trong những kỹ thuật được sử dụng phổ biến và hiệu quả trong đánh giá mô hình là K-fold Cross Validation (kiểm định chéo K phần hay còn được gọi xác thực chéo K lần).

K-fold Cross Validation là một phương pháp kiểm định chéo trong đó tập dữ liệu gốc được chia đều thành K phần (folds) không giao nhau. Trong quá trình huấn luyện và kiểm tra, mô hình sẽ được lặp lại K lần; tại mỗi lần lặp, một fold được sử dụng làm tập kiểm tra, trong khi K-1 fold còn lại được sử dụng để huấn luyện. Kết quả đánh giá từ K lần kiểm tra sẽ được tổng hợp (thường là lấy trung bình) để cung cấp một ước lượng ổn định và toàn diện hơn về hiệu suất mô hình.

Phương pháp này khắc phục được hạn chế của phương pháp kiểm định đơn giản, vốn chỉ chia dữ liệu một lần duy nhất thành hai tập train/test, dễ dẫn đến kết quả thiếu ổn định do ngẫu nhiên. Nhờ sử dụng toàn bộ dữ liệu cả cho huấn luyện và kiểm tra theo cách xoay vòng, K-fold Cross Validation giảm thiểu sai số do cách chia dữ liệu, đồng thời tận dụng tối đa thông tin từ tập dữ liệu gốc, đặc biệt hiệu quả trong các trường hợp dữ liệu hạn chế về số lượng [30].

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Interaction 1	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train
Interaction 2	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train
Interaction 3	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train
Interaction 4	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train
Interaction 5	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train
Interaction 6	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train
Interaction 7	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train
Interaction 8	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train
Interaction 9	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train
Interaction 10	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test

Hình 2.8. K-fold Cross Validation

Tóm lại, K-fold Cross Validation là một công cụ mạnh mẽ, linh hoạt và đáng tin cậy trong việc đánh giá hiệu suất mô hình học máy. Việc ứng dụng kỹ thuật này không chỉ giúp giảm nguy cơ đánh giá sai hiệu suất mà còn giúp cải thiện tính khách quan trong việc lựa chọn mô hình phù hợp cho bài toán dự báo. Trong nghiên cứu này, để đảm bảo tính khách quan, hệ số K=10 sẽ được sử dụng để xác thực chéo các tập dữ liệu.

2.5.2. Các chỉ số hiệu suất đánh giá mô hình

Hiệu suất dự báo hay độ chính xác dự báo của một mô hình được đánh giá thông qua sử dụng tập dữ liệu kiểm chứng (test data). Các chỉ tiêu được sử dụng để đánh giá độ chính xác dự báo gồm: căn bậc hai của sai số bình phương trung bình (root mean square error – RMSE), sai số tuyệt đối trung bình (mean absolute error – MAE), phần trăm sai số tuyệt đối trung bình (mean absolute percentage error – MAPE) và hệ số tương quan (correlation coefficient – R).

Các chỉ số này được sử dụng để đo lường độ tin cậy của giá trị dự đoán và giá trị thực tế. R cao hơn cho thấy sự tương đồng lớn hơn giữa kết quả dự đoán và giá trị thực tế. Tương tự, các giá trị RMSE, MAE và MAPE nhỏ chỉ ra rằng kết quả dự đoán là đáng

tin cậy.

Công thức tính toán của các chỉ số thể hiện bên dưới.

$$R = \frac{n \sum y \cdot y' - (\sum y)(\sum y')}{\sqrt{n(\sum y^2) - (\sum y)^2} \sqrt{n(\sum y'^2) - (\sum y')^2}} \quad (15)$$

trong đó, y là giá trị thực tế; y' là giá trị dự đoán, n là số lượng dữ liệu dự báo.

MAPE là một đại lượng thống kê dùng để đo lường phần trăm sai số tương đối mà một mô hình dự đoán có thể mắc phải. Khi so sánh độ chính xác dự đoán của các mô hình khác nhau, MAPE là chỉ số hữu ích vì nó không bị ảnh hưởng bởi kích thước mẫu và đơn vị của giá trị dự đoán.

MAE là đại lượng phản ánh sai số giữa giá trị thực tế và giá trị dự đoán mà không quan tâm đó là sai số vượt quá hay sai số thiếu hụt.

RMSE được xác định bằng cách lấy giá trị căn bậc hai của tổng bình phương của giá trị thực tế và giá trị dự đoán. RMSE đo lường sự khác biệt giữa giá trị thực tế với giá trị dự đoán bằng mô hình.

Trong quá trình đánh giá hiệu suất của các mô hình dự báo, đặc biệt là trong các lĩnh vực dự báo phức tạp, chỉ số Standardized Index (SI) là một thước đo thống kê không thứ nguyên thường được sử dụng nhằm đánh giá mức độ phân tán tương đối giữa giá trị mô phỏng và giá trị quan sát. SI càng nhỏ thì mức cải thiện của mô hình càng tốt so với mô hình gốc. SI được sử dụng rất rộng rãi trong việc đánh giá khách quan hiệu suất mô hình dự báo.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y - y'}{y} \right| \quad (16)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - y'| \quad (17)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y' - y)^2} \quad (18)$$

$$SI = \frac{1}{m} \sum_{i=1}^m \left(\frac{P_i - P_{\min,i}}{P_{\max,i} - P_{\min,i}} \right) \quad (19)$$

2.6. Phần mềm lập trình

Mô hình dự báo ED-SEM được phát triển trong môi trường MATLAB, một phần mềm toán học nổi tiếng được phát triển bởi MathWorks. MATLAB cho phép người dùng xây dựng ma trận, vẽ các hàm toán học, thực hiện các phép tính, liên kết các chương trình máy tính được viết bằng nhiều ngôn ngữ lập trình khác nhau, phân tích dữ liệu, phát triển thuật toán và tạo mô hình và ứng dụng.

MATLAB là ngôn ngữ bậc cao, tích hợp khả năng tính toán, hình ảnh hóa, lập trình trong một môi trường dễ sử dụng, ở đó vấn đề và giải pháp được trình bày trong cùng một lời chú thích toán học. Thường MATLAB được dùng cho:

- Toán và điện toán
- Phát triển thuật toán
- Dựng mô hình, giả lập, tạo nguyên mẫu
- Phân tích, khám phá hình ảnh hóa dữ liệu
- Đồ họa khoa học và kỹ thuật
- Phát triển ứng dụng, có cả xây dựng giao diện đồ họa người dùng
Graphic User Interface

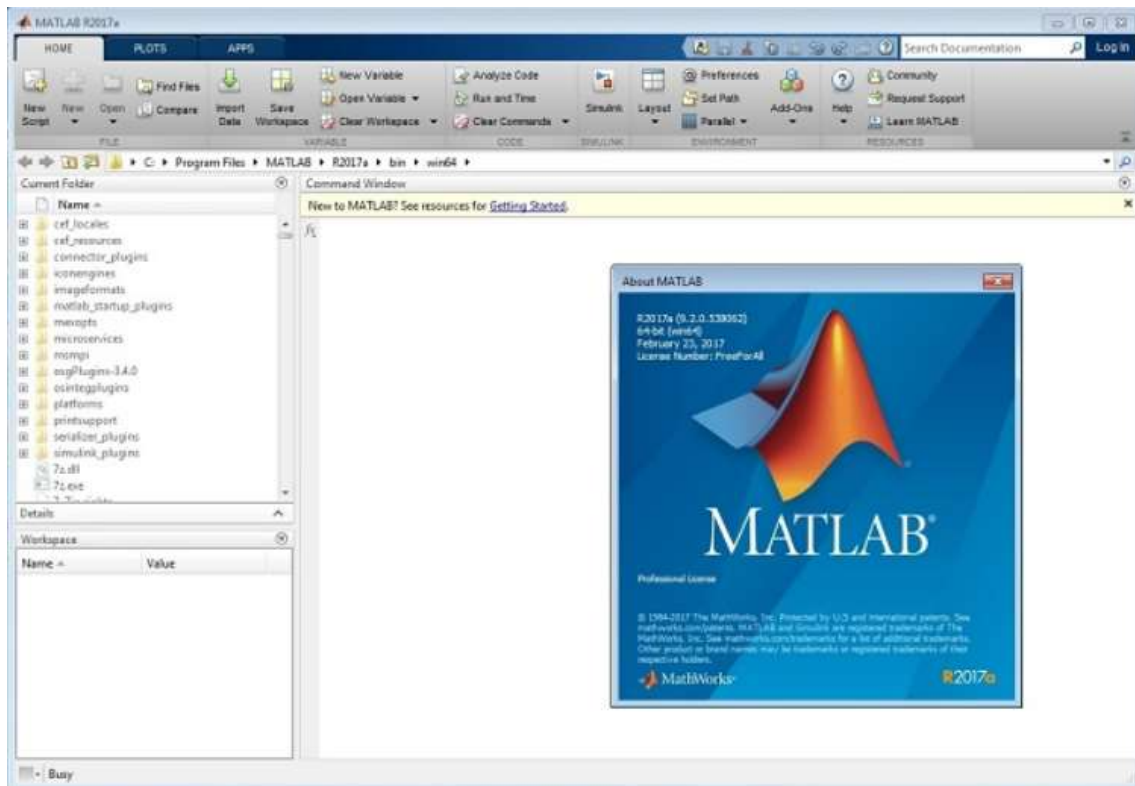
MATLAB là hệ thống tương tác, trong đó các phần tử dữ liệu xếp dưới dạng mảng, không cần chiều hướng, cho phép giải quyết nhiều vấn đề tính toán, đặc biệt là với ma trận và véc-tơ, trong thời gian nhanh chóng, chỉ bằng một phần so với viết phần mềm bằng các ngôn ngữ không tương tác vô hướng như C hay Fortran.

MATLAB là viết tắt của Matrix Laboratory (phòng thí nghiệm ma trận). Ban đầu MATLAB được viết để dễ dàng truy cập phần mềm ma trận do các dự án của LINPACK và EISPACK phát triển. Họ cũng đã mang tới những tính năng mới nhất cho phần mềm trong thế giới điện toán ma trận.

Trong môi trường đào tạo, nó là công cụ hướng dẫn chuẩn mực cho cả các khóa học dẫn nhập và chuyên sâu trong toán học, kỹ thuật và khoa học. Trong ngành, MATLAB

cũng là công cụ được nhiều nghiên cứu, phân tích, phát triển lựa chọn.

MATLAB còn có một bộ các giải pháp hướng tới ứng dụng có tên toolbox. Toolbox rất quan trọng với hầu hết người dùng MATLAB bởi nó cho phép học và áp dụng công nghệ chuyên môn hóa. Toolbox là bộ sưu tập các hàm MATLAB (M-file) mở rộng môi trường MATLAB để giải quyết từng lớp vấn đề. Các lĩnh vực mà toolbox có thể làm việc gồm xử lý tín hiệu, hệ thống kiểm soát, mạng thần kinh, logic mờ, phép biến đổi wavelet, mô phỏng.



Hình 2.9. Giao diện MATLAB

Chính vì vậy việc áp dụng phần mềm MATLAB trong nghiên cứu về kỹ thuật máy học là việc cần thiết. Nghiên cứu này thiết lập mô hình dự báo dựa trên các thuật toán thông minh trong môi trường MATLAB, nhằm hỗ trợ giải quyết các vấn đề về phân tích các siêu dữ liệu và thiết lập mô hình dự báo thông minh.

CHƯƠNG III. THU THẬP DỮ LIỆU VÀ KẾT QUẢ THỰC NGHIỆM

3.1 Thu thập dữ liệu

Ba tập dữ liệu với cường độ nén bê tông (CCS) là biến đầu ra và nhiều biến đầu vào khác nhau được xem xét và trình bày trong bảng 3.1 và bảng 3.2.

Bảng 3.1. Nguồn dữ liệu.

Bộ dữ liệu	Nguồn	Vật liệu bổ sung	Phòng thực nghiệm	Số lượng mẫu
1	Videla et al. [30]	Xi lò cao, muối silic	Chi Lê	200
2	Lam et al. [31]	Tro bay, muối silic	Hồng Kông	144
3	Lim et al. [2]	Tro bay, muối silic, phụ gia siêu dẻo	Hàn Quốc	104

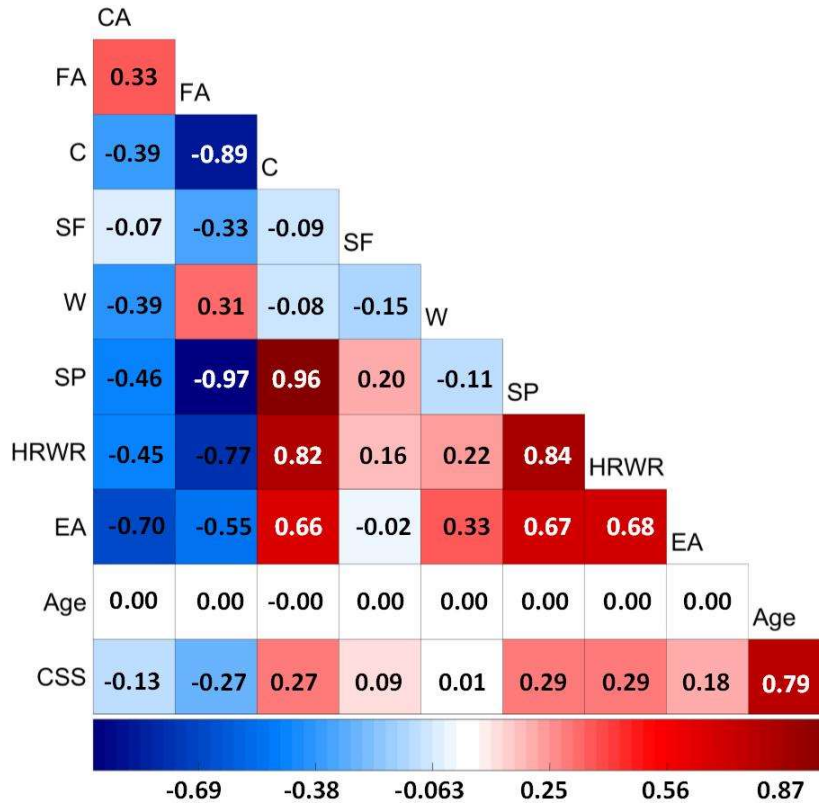
Bảng 3.2. Các thành phần chính của HPC trong bộ dữ liệu.

Thành phần	Đơn vị	Thấp nhất	Trung bình	Cao nhất	Chiều dữ liệu
<i>Bộ dữ liệu 1- Chi Lê [30]</i>					
Cốt liệu thô	kg/m ³	1105.0	1135.73	1173.0	Đầu vào
Cốt liệu mịn	kg/m ³	488.0	602.71	700.0	
Xi măng	kg/m ³	408.0	518.31	659.0	
Muối silic	kg/m ³	0.0	24.57	59.0	
Nước	kg/m ³	160.0	164.74	168.0	

Thành phần	Đơn vị	Thấp nhất	Trung bình	Cao nhất	Chiều dữ liệu
Phụ gia siêu dẻo	kg/m ³	2.2	2.73	3.3	
Phụ gia khử nước	kg/m ³	6.7	9.30	14.5	
Hàm lượng không khí	%	1.3	1.82	2.5	
Tuổi mẫu	ngày	1.0	18.81	56.0	
Độ bền nén	MPa	21.2	67.11	113.7	Đầu ra
<i>Bộ dữ liệu 2- Hồng Kông [31]</i>					
Hàm lượng tro bay	%	0.0	25.00	55.0	Đầu vào
Hàm lượng muối silic	%	0.0	1.88	5.0	
Xi măng	kg/m ³	400.0	436.67	500.0	
Cốt liệu mịn	kg/m ³	536.0	639.38	724.0	
Cốt liệu thô	kg/m ³	1086.0	1125.00	1157.0	
Nước	lit/m ³	150.0	171.98	205.0	
Phụ gia khử nước	lit/m ³	0.0	4.87	13.0	
Tuổi mẫu	ngày	3.0	60.67	180.0	
Độ bền nén	MPa	7.8	56.63	107.8	Đầu ra
<i>Bộ dữ liệu 3- Hàn Quốc [2]</i>					

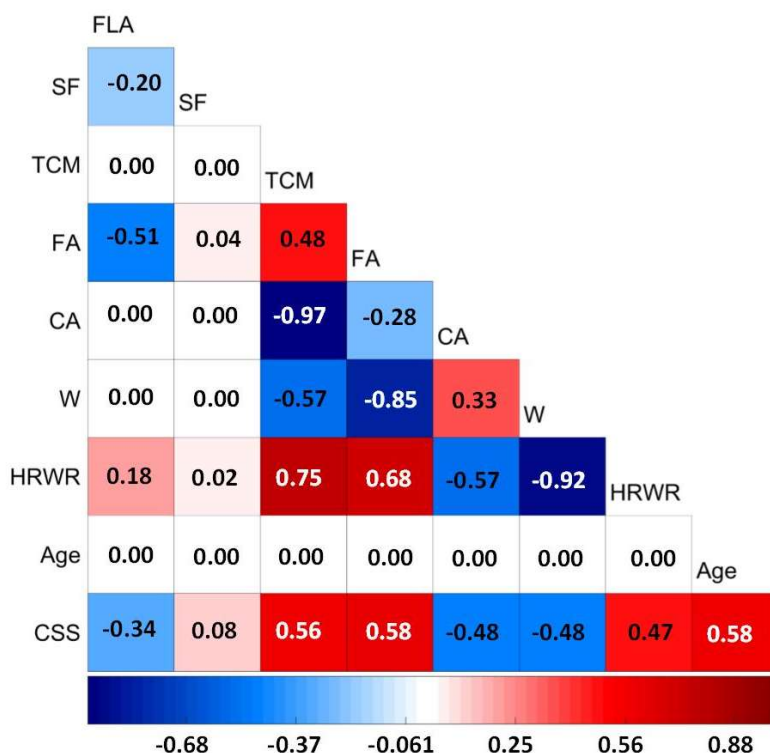
Thành phần	Đơn vị	Thấp nhất	Trung bình	Cao nhất	Chiều dữ liệu
Tỷ lệ nước/chất kết dính	%	30.0	37.60	45.0	Input
Nước	kg/m ³	160.0	170.00	180.0	
Cốt liệu mịn	%	37.0	46.00	53.0	
Tro bay	%	0.0	10.10	20.0	
Tỉ lệ không khí	kg/m ³	0.04	0.05	0.08	
Phụ gia siêu dẻo	kg/m ³	1.89	4.48	8.5	
Độ bền nén	MPa	38.0	52.68	74.0	Output

Tập dữ liệu 1 – Chi Lê: Xi măng sử dụng là loại xi măng xỉ lò cao đông kết nhanh, có thêm muối silic, cốt liệu thô và mịn nghiền từ vật liệu silic, cùng với phụ gia dẻo và phụ gia siêu dẻo giảm nước cường độ cao. Các mẫu bê tông thử nghiệm được phối trộn theo tỷ lệ cốt liệu thô khuyến nghị trong tiêu chuẩn ACI 211.4R-93. Tất cả mẫu được bảo dưỡng trong phòng tiêu chuẩn (nhiệt độ 20 ± 3 °C và độ ẩm tương đối > 90%) cho đến khi thử nghiệm ở các độ tuổi khác nhau. Các thí nghiệm cường độ nén được thực hiện trên hai mẫu trụ có kích thước 150×300 mm theo tiêu chuẩn ASTM C39 vào các ngày 1, 3, 7, 28 và 56. Tập dữ liệu này bao gồm tổng cộng 200 mẫu, được thu thập bởi Videla và Gaedicke. Cường độ nén (CCS) có thể được dự đoán dựa trên 9 biến đầu vào: cốt liệu thô (CA), cốt liệu mịn (FA), xi măng (C), silica fume (SF), nước (W), phụ gia siêu dẻo (SP), phụ gia giảm nước cường độ cao (HRWR), hàm lượng không khí bị giữ lại (EA), và độ tuổi thử nghiệm (Age).



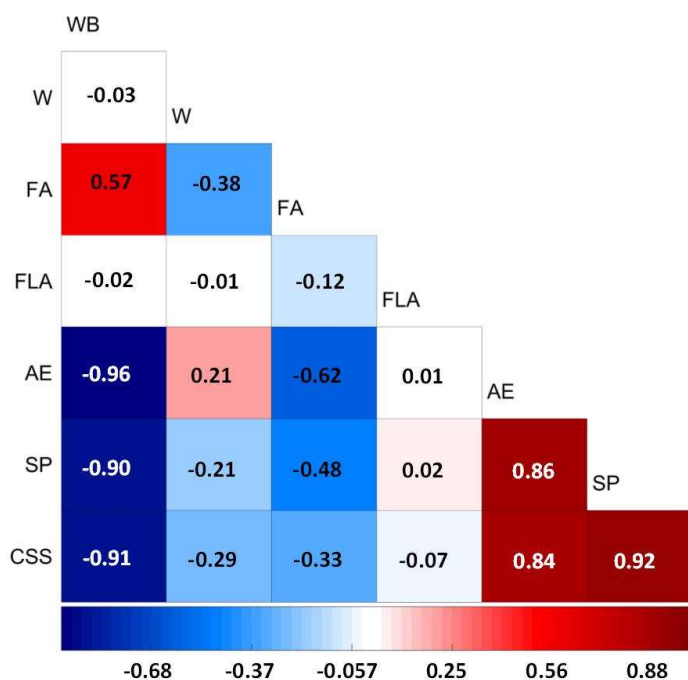
Hình 3.1. Ma trận tương quan tính năng của bộ dữ liệu 1

Tập dữ liệu 2 – Hồng Kông: Các mẫu bê tông được trộn với các tỷ lệ nước trên vật liệu kết dính khác nhau, với hàm lượng tro bay thấp và cao, và có hoặc không có thêm một lượng nhỏ silica fume. Tổng cộng có 144 mẫu khác nhau từ 24 công thức phối trộn, được bảo dưỡng trong 3, 7, 28, 56 hoặc 180 ngày. Trong mỗi nhóm trộn, tỷ lệ thay thế xi măng bằng tro bay (theo khối lượng) dao động từ 0% đến 55%. Một số hỗn hợp có thêm 5% silica fume thay thế. Các vật liệu kết dính bao gồm xi măng Portland tương đương với ASTM Loại I, tro bay canxi thấp tương đương ASTM Class F, và silica fume cô đặc được thương mại hóa tại Hồng Kông. Tập dữ liệu 3 gồm 144 kết quả thử nghiệm, được thu thập bởi Lam và cộng sự, với 8 biến đầu vào: tỷ lệ thay thế tro bay (FLA), tỷ lệ thay thế silica fume (SF), tổng vật liệu kết dính (TCM), cốt liệu mịn (FA), cốt liệu thô (CA), hàm lượng nước (W), phụ gia giảm nước cường độ cao (HRWR), và độ tuổi mẫu (Age). Hình 3.2 cho thấy hầu hết các biến không có mối tương quan mạnh với CCS. Không có ngoại lệ lớn (outlier) nào gây ảnh hưởng đáng kể đến hiệu suất mô hình.



Hình 3.2. Ma trận tương quan tính năng của bộ dữ liệu 2

Tập dữ liệu 3 – Hàn Quốc: Tất cả các vật liệu sử dụng trong thí nghiệm đều được sản xuất tại Hàn Quốc. Cường độ nén dao động từ 40–80 MPa. Tỷ lệ nước trên chất kết dính nằm trong khoảng 0,30 đến 0,45, và hàm lượng tro bay thay thế chiếm từ 0% đến 20% tổng lượng chất kết dính. Hàm lượng phụ gia siêu dẻo và phụ gia tạo bọt khí lần lượt là từ 0–2% và 0,010–0,013% (tính theo phần trăm khối lượng chất khô so với chất kết dính). Xi măng Portland được sử dụng tuân theo tiêu chuẩn ASTM Loại I. Cốt liệu thô là đá granite nghiền (tỷ trọng riêng: 2,7; mô đun độ mịn: 7,2; kích thước hạt lớn nhất: 19 mm). Cốt liệu mịn là cát thạch anh (tỷ trọng riêng: 2,61; mô đun độ mịn: 2,94). Tập dữ liệu 4 được thu thập bởi Lim và cộng sự, gồm 6 biến đầu vào: tỷ lệ nước/chất kết dính (WB), hàm lượng nước (W), cốt liệu mịn (FA), tro bay (FLA), tỷ lệ tạo bọt khí (AE) và phụ gia siêu dẻo (SP). Như thể hiện trong Hình 5c, hai biến AE và SP có mối tương quan tuyến tính dương mạnh với CCS trong tập dữ liệu này.



Hình 3.3. Ma trận tương quan tính năng của bộ dữ liệu 3

Bảng 3.3 Thống kê mẫu dữ liệu.

Bộ dữ liệu 1- Chi Lê - 200 mẫu

STT	Nước (kg/m ³)	Phụ gia khử nước (kg)	Độ sụt (mm)	Phụ gia siêu dẻo (kg)	Hàm lượng không khí (%)	Muội silic (kg)	Ngày	Độ bền nén (Mpa)
1	168	8.9	140	2.5	2	0	1	30.6
2	167	8.8	100	2.5	2.2	15	1	31.2
3	167	7.8	135	2.5	1.9	29	1	29.0
4	168	8.9	160	2.5	1.7	44	1	25.7
5	166	8.6	145	2.7	2.2	0	1	28.1
...	...	...	...	...	...	...	...	...
196	165	10.5	105	2.9	1.8	53	56	98.4

197	162	7.6	135	2.4	1.7	43	56	82.2
198	160	9.3	140	2.6	1.6	31	56	93.6
199	161	8.0	120	2.9	1.7	17	56	85.6
200	161	8	125	2.9	1.8	34	56	89.7

Bộ dữ liệu 2- Hồng Kông - 144 mẫu

ST T	Nước (kg/m ³)	Tỉ lệ phụ gia khử nước W/B (%)	Ngày	Cốt liệu thô (kg)	Cốt liệu mịn (kg)	Hàm lượng muội silic (%)	Hàm lượng tro bay (%)	Độ bền nén (Mpa)
1	150	30	3	1086	724	0	0	64.9
2	150	30	3	1086	700	0	15	52.1
3	150	30	3	1086	683	0	25	48
4	150	30	3	1086	650	0	45	34.1
5	150	30	3	1086	634	0	55	22.3
...	...	...	...	...	...	...	...	...
140	205	50	180	1132	549	0	45	56.6
141	205	50	180	1132	536	0	55	48.3
142	205	50	180	1132	605	5	0	66.3
143	205	50	180	1132	578	5	20	68.0
144	205	50	180	1132	552	5	40	58.4

Bộ dữ liệu 3 – Hàn Quốc – 104 mẫu

ST T	Nước (kg/m ³)	Tỉ lệ phụ gia khử nước	Độ sụt	Phụ siêu dẻo	Không khí (kg/m ³)	Hàm lượng muội	Hàm lượng tro bay	Độ bền nén
---------	------------------------------	------------------------------	-----------	-----------------	--------------------------------------	----------------------	-------------------------	---------------

		W/B (%)	(mm)	(kg/m3)	)	silic (%)	(%)	(Mpa)
1	160	30	215	8	0.069	0	10	74.0
2	160	30	245	8	0.069	0	20	74.0
3	160	30	200	8	0.069	0	0	71.0
4	160	30	204	8	0.069	0	0	70
5	160	30	205	8	0.069	0	20	69
...	...	...	...	...	...	...	...	...
100	180	45	170	2	0.040	0	0	42.0
101	180	45	185	2	0.040	0	20	41.0
102	180	45	175	2	0.040	0	0	43.0
103	180	45	220	2	0.040	0	10	40.0
104	180	45	170	2	0.040	0	20	38.0

3.2 Xử lý dữ liệu và kết quả thực nghiệm

Các tập dữ liệu sẽ lần lượt được thử nghiệm và đánh giá bởi các mô hình đơn và mô hình tổ hợp để chọn ra mô hình tốt nhất và tối ưu hóa các siêu tham số của nó bằng thuật toán ED (bảng 3.4 và 3.5)

Bảng 3.4. Cài đặt tham số cho các mô hình LSSVR, RBFNN, LSSVR (RBFNN, LSSVR) và ED-SEM.

LSSVR model			
No.	Paramter	Purpose	Setting
1	C_1	Regularization parameter	8.604E+10
2	σ_1	RBF parameter	2.048E+03
RBFNN model			

No.	Paramter	Purpose	Setting
1	N_2	Number of hidden neurons	437
2	σ_2	RBF parameter	97.206
SEM model			
No.	Paramter	Purpose	Setting
1	C_1	Regularization parameter	9.706E+11
2	σ_1	RBF parameter	9.572E+01
3	N_2	Number of hidden neurons	400
4	σ_2	RBF parameter	2.427E+02
5	C_c	Regularization parameter	1.419E+11
6	σ_c	RBF parameter	2.109E+02
ED- SEM model			
No.	Paramter	Purpose	Setting
1	Popular size	Population size	25
2	Max iteration	Stopping criteria	20
3	C_1	Regularization parameter	[1E-4 1E+12]
4	σ_1	RBF parameter	[1E-4 1E+12]
5	N_2	Number of hidden neurons	[1 500]
6	σ_2	RBF parameter	[1E-4 1E+12]
7	C_c	Regularization parameter	[1E-4 1E+12]
8	σ_c	RBF parameter	[1E-4 1E+12]

Bảng 3.5. Cài đặt tham số cho các thuật toán tối ưu hóa nổi tiếng để tối ưu hóa SEM

No.	Paramter	Setting
Parameters of machine learning		
1	C_1	[1E-4 1E+12]
2	σ_1	[1E-4 1E+12]
3	N_2	[1 500]
4	σ_2	[1E-4 1E+12]
5	C_c	[1E-4 1E+12]
6	σ_c	[1E-4 1E+12]
Parameters of optimizer algorithms		
Genetic Algorithm (GA)		
1	Number of population	25
2	Maximum Number of Iterations	20
3	Crossover Percentage	0.7
4	Mutation Rate	0.1
Particle Swarm Optimization (PSO)		
1	Population Size (Swarm Size)	25
2	Maximum Number of Iterations	20
3	Inertia Weight	0.5
4	Inertia Weight Damping Ratio	0.99
5	Personal Learning Coefficient	2.05
6	Global Learning Coefficient	2.05
Teaching-Learning-Based Optimization (TLBO)		
1	Population Size	25

2	Maximum number of evaluation	500
Symbiotic Organisms Search (SOS)		
1	Eco size	25
2	Maximum number of evaluation	500
Enterprise Development Optimizer (ED)		
1	Popular size	25
2	Max iteration	20

Bảng 3.6 trình bày hiệu suất của mô hình ED-SEM được đề xuất trong việc dự đoán CCS của HPC. Ngoài ra, nghiên cứu này so sánh kết quả với các phương pháp khác được báo cáo trong nghiên cứu trước đây, như thể hiện trong Bảng 3.7 đến 3.9 cho các tập dữ liệu từ 1 đến 3.

Bảng 3.6. Hiệu suất dự đoán của mô hình đề xuất.

Dữ liệu	Mô hình	MAE (Mpa)	RMSE (MPa)	MAPE (%)	R	SI (Rank)
Bộ 1	RBFNN	255.044	153.25	275.651	0.098	1.000 (8)
	LSSVR	6.296	2.574	8.182	0.909	0.035 (7)
	SEM	6.216	2.399	9.076	0.926	0.031 (6)
	GA-SEM	3.536	0.91	5.316	0.973	0.009 (5)
	PSO-SEM	2.983	0.707	4.62	0.978	0.006 (4)
	TLBO-SEM	2.791	0.716	4.326	0.984	0.004 (2)
	SOS-SEM	2.785	0.699	4.568	0.986	0.004 (2)
	ED-SEM	2.168	0.361	3.521	0.991	0.000 (1)
Bộ 2	RBFNN	172.839	92.482	257.303	-0.009	1.000 (8)
	LSSVR	9.436	3.65	16.009	0.798	0.082 (7)

Dữ liệu	Mô hình	MAE (Mpa)	RMSE (MPa)	MAPE (%)	R	SI (Rank)
	SEM	6.763	2.259	12.588	0.885	0.050 (6)
	GA-SEM	4.1	0.815	7.649	0.96	0.018 (4)
	PSO-SEM	4.233	1.337	7.219	0.949	0.022 (5)
	TLBO-SEM	2.566	0.636	4.771	0.986	0.006 (2)
	SOS-SEM	2.954	0.704	5.791	0.984	0.009 (3)
	ED-SEM	1.532	0.379	2.96	0.996	0.000 (1)
Bộ 3	RBFNN	5.657	4.289	11.446	0.865	1.000 (8)
	LSSVR	1.809	0.668	3.443	0.97	0.180 (6)
	SEM	2.181	0.885	4.194	0.959	0.252 (7)
	GA-SEM	1.055	0.306	2.017	0.99	0.045 (4)
	PSO-SEM	1.028	0.447	1.955	0.989	0.053 (5)
	TLBO-SEM	0.847	0.393	1.617	0.992	0.026 (3)
	SOS-SEM	0.892	0.363	1.718	0.994	0.025 (2)
	ED-SEM	0.714	0.278	1.345	0.995	0.000 (1)

Bảng 3.7. Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 1.

No.	ML technique	Author	MAE (MPa)	RMSE (MPa)	MAPE (%)	R
1	Hyp-Exp	Videla and Gaedicke (2004) [30]	4.438	N/A	N/A	0.980
2	Voting ANNs+CART	Chou and Pham (2013)	3.406	4.466	5.9	0.982

No.	ML technique	Author	MAE (MPa)	RMSE (MPa)	MAPE (%)	R
		[20]				
3	Bagging CART	Chou và cộng sự (2014) [13]	3.82	5.13	6.81	N/A
4	BooST	Anyaocha và cộng sự (2020) [32]	4.70	6.30	N/A	0.90
5	Optimum semi-empirical framework	Nguyen và cộng sự (2021) [14]	5.394	6.426	N/A	0.9621
6	Hybrid ED-SEM	Nghiên cứu này	2.168	0.361	3.521	0.991

Bảng 3.8. Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 2.

No	ML technique	Author	MAE (MPa)	RMSE (MPa)	MAPE (%)	R
1	ANN	Pala và cộng sự (2007) [33]	N/A	0.0154	4.845	0.9995
2	Voting ANNs+SVR	Chou and Pham (2013) [20]	4.242	5.078	1.5	0.993
3	MLP	Chou và cộng sự (2014) [13]	4.28	5.81	10.51	N/A
4	BooST	Anyaocha và cộng sự (2020) [34]	4.70	6.30	N/A	0.90

No	ML technique	Author	MAE (MPa)	RMSE (MPa)	MAPE (%)	R
5	ANN	Sun và cộng sự (2021) [34]	N/A	N/A	N/A	0.9667
6	Polynomial-SVR	Huang và cộng sự (2021) [35]	4.602	5.517	N/A	0.9869
7	Optimum semi- empirical framework	Nguyen và cộng sự (2021) [14]	3.711	5.175	N/A	0.9775
8	Hybrid ED-SEM	Nghiên cứu này	1.532	0.379	2.960	0.996

Bảng 3.9. Kết quả so sánh từ các nghiên cứu đã công bố cho bộ dữ liệu 3.

No	ML technique	Author	MAE (MPa)	RMSE (MPa)	MAPE (%)	R
1	WGP	Tsai and Lin (2011) [36]	N/A	2.180	N/A	0.978
2	Bagging (ANNs+ CART)	Chou and Pham (2013) [20]	1.075	1.426	2.0	0.989
3	Stacking SVM (CART+SVM+LR)	Chou và cộng sự (2014) [13]	1.09	1.51	2.11	N/A
4	MR	Lim và cộng sự (2004) [2]	N/A	N/A	N/A	0.976
5	NaOT	Nguyen và cộng sự (2021) [14]	2.807	2.059	N/A	0.9587
6	Hybrid ED-SEM	Nghiên cứu này	0.714	0.2778	1.345	0.995

Kết quả thực nghiệm được trình bày trong các bảng 3.5 đến 3.8 đã chứng minh một

cách nhất quán rằng mô hình ED-SEM có hiệu suất vượt trội trong việc dự báo độ bền nén của bê tông. Cụ thể, mô hình này được đánh giá trên ba tập dữ liệu đại diện cho các trường hợp khác nhau trong thực tế, và ở tất cả các trường hợp đó, ED-SEM đều thể hiện khả năng dự báo chính xác cao hơn so với các phương pháp truyền thống trước đây.

Ba chỉ số lỗi chính được sử dụng để đánh giá hiệu suất bao gồm: MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) và MAPE (Mean Absolute Percentage Error). Kết quả cho thấy mô hình ED-SEM đạt được giá trị MAE, RMSE và MAPE thấp hơn đáng kể so với các mô hình đối chứng, phản ánh độ sai lệch giữa giá trị dự báo và giá trị thực tế đã được giảm thiểu rõ rệt. Điều này chứng minh rằng mô hình không những dự báo chính xác mà còn có khả năng học tốt các mối quan hệ phức tạp giữa các yếu tố đầu vào và đầu ra.

Một trong những yếu tố nổi bật nhất của mô hình ED-SEM là hệ số tương quan (R – Correlation Coefficient) giữa giá trị dự báo và giá trị thực tế. Trong tất cả các trường hợp được kiểm tra, hệ số R của ED-SEM đều tiến gần đến giá trị 1 – là giá trị cao nhất có thể đạt được đối với một mô hình thống kê tuyến tính và phi tuyến. Điều này chứng tỏ mối quan hệ giữa đầu vào và đầu ra được mô hình nắm bắt một cách gần như hoàn hảo, cho thấy mức độ tin cậy cao của mô hình trong ứng dụng thực tế.

Bên cạnh đó, mô hình còn cho thấy tính ổn định khi được đánh giá trên nhiều tập dữ liệu khác nhau với tính chất đa dạng về thành phần bê tông và điều kiện thử nghiệm. Đây là một điểm mạnh quan trọng trong bối cảnh ngành xây dựng ngày càng đòi hỏi các mô hình có khả năng tổng quát hóa cao.

Từ các phân tích định lượng nêu trên, có thể rút ra kết luận rằng mô hình ED-SEM là một công cụ hữu ích để dự báo độ bền nén của bê tông. Với những đặc điểm kỹ thuật nổi bật, mô hình này hoàn toàn có thể được ứng dụng trong các hệ thống hỗ trợ ra quyết định kỹ thuật tại hiện trường, trong thiết kế kết cấu, hoặc trong công tác kiểm tra chất lượng vật liệu xây dựng. Điều này không chỉ góp phần nâng cao hiệu quả và độ an toàn của công trình, mà còn thúc đẩy việc áp dụng trí tuệ nhân tạo vào ngành xây dựng một cách thực tiễn và bền vững.

KẾT LUẬN VÀ KIẾN NGHỊ

Nghiên cứu này đã đề xuất và kiểm chứng một phương pháp tiếp cận mới dựa trên mô hình ED-SEM (Enterprise Development Optimizer Ensemble Model) nhằm dự báo độ bền nén của bê tông hiệu suất cao (High-Performance Concrete – HPC). Việc lựa chọn HPC làm đối tượng nghiên cứu là hoàn toàn phù hợp với xu hướng phát triển hiện đại của ngành xây dựng, khi các loại bê tông có tính năng cơ học và độ bền vượt trội ngày càng được ứng dụng rộng rãi trong các công trình quy mô lớn và yêu cầu kỹ thuật cao.

Để đảm bảo độ tin cậy và tính tổng quát của mô hình đề xuất, nhóm nghiên cứu đã sử dụng ba bộ dữ liệu HPC được thu thập từ các phòng thí nghiệm khác nhau, phản ánh sự đa dạng về thành phần vật liệu, điều kiện thử nghiệm và kỹ thuật sản xuất. Mô hình ED-SEM được huấn luyện và đánh giá bằng kỹ thuật xác thực chéo 10 lần (10-fold cross validation), giúp giảm thiểu hiện tượng quá khớp (overfitting) và cải thiện khả năng dự báo trong điều kiện thực tế.

Kết quả phân tích cho thấy rằng ED-SEM không chỉ mang lại độ chính xác cao hơn so với các phương pháp truyền thống, mà còn cải thiện đáng kể mối quan hệ giữa các biến đầu vào (như hàm lượng xi măng, nước, cốt liệu, phụ gia,...) và biến đầu ra (cường độ chịu nén).

Khả năng ứng dụng của mô hình ED-SEM là rất tiềm năng, không chỉ trong môi trường nghiên cứu mà còn trong thực tiễn công nghiệp xây dựng. Nhờ hiệu suất dự báo cao, tính linh hoạt, và khả năng thích ứng với các bộ dữ liệu khác nhau, phương pháp này có thể hỗ trợ kỹ sư và nhà thầu đưa ra các quyết định chính xác và nhanh chóng liên quan đến thiết kế, kiểm tra chất lượng và tối ưu hóa quy trình sản xuất bê tông.

Tóm lại, mô hình ED-SEM được chứng minh là một giải pháp tiên tiến, ổn định và hiệu quả, phù hợp để thay thế hoặc bổ sung cho các phương pháp hiện có trong dự báo cường độ chịu nén của bê tông hiệu suất cao.

Những kết quả trên cho thấy tiềm năng phát triển của mô hình ED-SEM là khá tốt để

mở rộng nghiên cứu sang các loại bê tông khác như bê tông siêu tính năng (Ultra-High Performance Concrete – UHPC), bê tông tự lèn (Self-Compacting Concrete – SCC), bê tông có vật liệu tái chế hoặc bê tông thân thiện môi trường hoặc thậm chí có thể thử nghiệm trên các lĩnh vực khác nhau ngoài ngành xây dựng. Điều này giúp đánh giá tính tổng quát hóa và độ linh hoạt của mô hình trong nhiều ngữ cảnh kỹ thuật khác nhau.

Kết hợp ED-SEM với các công nghệ hiện đại như IoT, cảm biến thông minh và hệ thống điều khiển tự động, nhằm tạo ra các giải pháp giám sát chất lượng vật liệu xây dựng theo thời gian thực. Điều này sẽ góp phần nâng cao chất lượng và độ an toàn của công trình, đồng thời tiết kiệm thời gian và nhân lực.

Phát triển giao diện ứng dụng người dùng thân thiện cho phép các kỹ sư, nhà thầu hoặc chuyên gia vật liệu dễ dàng nhập dữ liệu đầu vào và nhận được kết quả dự báo nhanh chóng mà không cần kiến thức sâu về mô hình học máy. Việc này giúp tăng cường khả năng tiếp cận của các mô hình dự báo nói chung và ED-SEM nói riêng cho những người không chuyên, giúp phát triển các lĩnh vực khác ngoài xây dựng.

TÀI LIỆU THAM KHẢO

1. Parichatprecha, R., P.J.C. Nimityongskul, and B. Materials, *Analysis of durability of high performance concrete using artificial neural networks*. 2009. **23**(2): p. 910-917.
2. Lim, C.-H., et al., *Genetic algorithm in mix proportioning of high-performance concrete*. 2004. **34**(3): p. 409-420.
3. Ni, H.-G., J.-Z.J.C. Wang, and C. Research, *Prediction of compressive strength of concrete by neural networks*. 2000. **30**(8): p. 1245-1250.
4. Schmidt, M. and E.J.A.S.P. Fehling, *Ultra-high-performance concrete: research, development and application in Europe*. 2005. **228**(1): p. 51-78.
5. Voo, Y., et al., *Ultra-high performance concrete—technology for present and future*. 2017: p. 12-14.
6. Maher, K. and Y.J.A.C.B.M. Voo, Summer, *Taking ultra-high performance concrete to new height—the Malaysian experience*. 2016. **2016**: p. 36-38.
7. Schmidt, M. *Sustainable building with ultra-high performance concrete (UHPC)—Coordinated research program in Germany*. in *Proceedings of Hipermat 2012 3rd International Symposium on UHPC and Nanotechnology for High Performance Construction Materials*, Ed., Schmidt, M., Fehling, E., Glotzbach, C., Fröhlich, S., and Piotrowski, S., Kassel University Press, Kassel, Germany. 2012.
8. Chou, J.-S., et al., *Bio-inspired optimization of weighted-feature machine learning for strength property prediction of fiber-reinforced soil*. 2021. **180**: p. 115042.
9. Rabczuk, T. and T.J.I.j.o.f. Belytschko, *Application of particle methods to static fracture of reinforced concrete structures*. 2006. **137**: p. 19-49.
10. Rabczuk, T., et al., *A numerical model for reinforced concrete structures*. 2005. **42**(5-6): p. 1327-1354.
11. Drzymala, T., et al., *Effects of high temperature on the properties of high performance concrete (HPC)*. 2017. **172**: p. 256-263.
12. Zhao, Y., et al., *Experimental study on shrinkage of HPC containing fly ash and ground granulated blast-furnace slag*. 2017. **155**: p. 145-153.
13. Chou, J.-S., et al., *Machine learning in concrete strength simulations: Multi-nation data analytics*. 2014. **73**: p. 771-780.
14. Nguyen, N.-H., et al., *Heuristic algorithm-based semi-empirical formulas for estimating the compressive strength of the normal and high performance concrete*. 2021. **304**: p. 124467.
15. Song, H., et al., *Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms*. 2021. **308**: p. 125021.

16. Bui, D.-K., et al., *A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete*. 2018. **180**: p. 320-333.
17. Zhang, W., D. Yang, and S.J.E.S.w.A. Zhang, *A new hybrid ensemble model with voting-based outlier detection and balanced sampling for credit scoring*. 2021. **174**: p. 114744.
18. Nazari, A. and J.G.J.C.I. Sanjayan, *Modelling of compressive strength of geopolymer paste, mortar and concrete by optimized support vector machine*. 2015. **41**(9): p. 12164-12177.
19. Arabameri, A., et al., *Modeling groundwater potential using novel GIS-based machine-learning ensemble techniques*. 2021. **36**: p. 100848.
20. Chou, J.-S., A.-D.J.C. Pham, and B. Materials, *Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength*. 2013. **49**: p. 554-563.
21. Truong, D.-N. and J.-S.J.E.S. Chou, *Metaheuristic algorithm inspired by enterprise development for global optimization and structural engineering problems with frequency constraints*. 2024. **318**: p. 118679.
22. Chou, J.S., A.D.J.C.A.C. Pham, and I. Engineering, *Smart artificial firefly colony algorithm-based support vector regression for enhanced forecasting in civil engineering*. 2015. **30**(9): p. 715-732.
23. Chou, J.-S., W.K. Chong, and D.-K.J.J.o.C.i.C.E. Bui, *Nature-inspired metaheuristic regression system: programming and implementation for civil engineering applications*. 2016. **30**(5): p. 04016007.
24. Chou, J.-S. and N.-T.J.A.i.c. Ngo, *Smart grid data analytics framework for increasing energy savings in residential buildings*. 2016. **72**: p. 247-257.
25. Cortes, C. and V.J.M.l. Vapnik, *Support-vector networks*. 1995. **20**: p. 273-297.
26. Lowe, D. and D.J.C.s. Broomhead, *Multivariable functional interpolation and adaptive networks*. 1988. **2**(3): p. 321-355.
27. Suykens, J.A., et al., *Weighted least squares support vector machines: robustness and sparse approximation*. 2002. **48**(1-4): p. 85-105.
28. Nisbet, R., J. Elder, and G.D. Miner, *Handbook of statistical analysis and data mining applications*. 2009: Academic press.
29. Kohavi, R. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. in *Ijcai*. 1995. Montreal, Canada.
30. Videla, C. and C.J.M.J. Gaedicke, *Modeling Portland blast-furnace slag cement high-performance concrete*. 2004. **101**(5): p. 365-375.
31. Lam, L., et al., *Effect of fly ash and silica fume on compressive and fracture behaviors of concrete*. 1998. **28**(2): p. 271-283.

32. Anyaoha, U., et al., *Soft computing in estimating the compressive strength for high-performance concrete via concrete composition appraisal*. 2020. **257**: p. 119472.
33. Pala, M., et al., *Appraisal of long-term effects of fly ash and silica fume on compressive strength of concrete by neural networks*. 2007. **21**(2): p. 384-394.
34. Sun, L., et al., *Applying a meta-heuristic algorithm to predict and optimize compressive strength of concrete samples*. 2021. **37**: p. 1133-1145.
35. Huang, J., Y. Sun, and J.J.E.w.C. Zhang, *Reduction of computational error by optimizing SVR kernel coefficients to simulate concrete compressive strength through the use of a human learning optimization algorithm*. 2022: p. 1-18.
36. Tsai, H.-C. and Y.-H.J.E.w.C. Lin, *Predicting high-strength concrete parameters using weighted genetic programming*. 2011. **27**: p. 347-355.