



Động lực thu hút vốn đầu tư trực tiếp nước ngoài tại các quốc gia đang phát triển: Bằng chứng thực nghiệm từ tiếp cận học máy

NGUYỄN ANH TÚ *

Trường Đại học Ngân hàng Thành phố Hồ Chí Minh

THÔNG TIN	TÓM TẮT
Ngày nhận: 17/06/2024 Ngày nhận lại: 07/08/2024 Duyệt đăng: 09/08/2024 Mã phân loại JEL: C52; F21; F41; F47. Từ khóa: FDI; Dự báo; Kinh tế lượng; Học máy; Mạng nơ-ron nhân tạo; Rừng ngẫu nhiên. Keywords: FDI; Forecasting; Econometrics; Machine learning; Artificial neural networks; Random forests.	Sử dụng dữ liệu về đầu tư trực tiếp nước ngoài (FDI) của 66 quốc gia đang phát triển giai đoạn 2013–2021, nghiên cứu này triển khai các phương pháp học máy bao gồm: Mạng nơ-ron nhân tạo (Artificial Neural Networks – ANN) và rừng ngẫu nhiên (Random Forest – RF) nhằm so sánh chất lượng dự báo với một phương pháp tiếp cận kinh tế lượng là GMM sai phân (Differenced GMM – DGMM). Trong đó, DGMM phân tích mối liên hệ giữa các nhân tố đầu vào ảnh hưởng tới thu hút FDI, trong khi phương pháp mạng nơ-ron nhân tạo và phương pháp rừng ngẫu nhiên dựa vào các nhân tố có ý nghĩa thống kê để đưa ra dự báo. Kết quả cho thấy các nhân tố độ lớn thị trường, độ mở thương mại, mức độ dồi dào của lao động, sự phát triển của thị trường tài chính là những nhân tố chính trợ lực cho thu hút FDI tại các quốc gia này. Trong khi đó, xét về mặt dự báo, sai số của phương pháp rừng ngẫu nhiên là nhỏ nhất và vượt trội so với các phương pháp so sánh. Các phát hiện của mô hình kinh tế lượng và mô hình học máy cũng được thảo luận trong nghiên cứu. Abstract Utilizing foreign direct investment (FDI) data from 66 developing countries for the period 2013–2021, this study implements machine learning methods, including artificial neural networks (ANN) and random forests (RF), to compare predictive quality with that of an econometric approach, the difference generalized method of moments (DGMM). In this context, DGMM analyzes the relationships between the input factors influencing FDI attraction, while ANN and RF make

* Tác giả liên hệ.

Biên tập viên: Nguyễn Thị Hồng Cẩm.

Email: tuna@hub.edu.vn (Nguyễn Anh Tú).

Trích dẫn bài viết: Nguyễn Anh Tú. (2024). Động lực thu hút vốn đầu tư trực tiếp nước ngoài tại các quốc gia đang phát triển: Bằng chứng thực nghiệm từ tiếp cận học máy. *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 35(5), 119-134.

predictions based on statistically significant factors. The results indicate that market size, trade openness, labor abundance, and financial market development are primary drivers facilitating FDI attraction in these countries. Regarding predictive accuracy, RF exhibits the lowest error and significantly outperforms the selected methods. The findings from both the econometric model and the machine learning models are also discussed in the study.

1. Giới thiệu

Các quốc gia đang phát triển (CQGĐPT) đang tích cực thu hút dòng vốn đầu tư trực tiếp nước ngoài (Foreign Direct Investment – FDI) bởi tính tích cực của dòng vốn này đã được chứng minh tại các nước sở tại. Thật vậy, khi các tập đoàn đa quốc gia (CTĐĐQG) thành lập các công ty con hoặc thuê ngoài một số khâu trong toàn bộ chuỗi cung ứng, đẩy mạnh hoạt động sản xuất của họ sang các nước này nhằm tận dụng các lợi thế khác nhau của các quốc gia nhận FDI. Khi CTĐĐQG đầu tư ra nước ngoài, họ mang theo bằng sáng chế về công nghệ sản xuất và thông tin kinh doanh hoặc bí quyết sản xuất. Điều này kéo theo những lợi ích tiềm năng thúc đẩy các biến số vĩ mô của nước sở tại như: Tăng trưởng kinh tế (Nguyen và cộng sự, 2022), tăng việc làm (Crescenzi và cộng sự, 2022), tăng xuất khẩu (Yao, 2006) và chuyển giao công nghệ hiện đại (Keller, 2021). Ngược lại, để có thể thu hút được FDI thành công, CQGĐPT cũng phải có những cơ sở hấp dẫn để mời gọi đầu tư từ CTĐĐQG. Những lợi thế này có thể bao gồm: Hiệu quả sản xuất (Coughlin & Segev, 2000), cơ sở hạ tầng như: điện năng, giao thông (Head & Ries, 1996), chi phí nhân công thấp (Cheng & Kwan, 2000), hoặc thị trường tiêu thụ rộng lớn (Narula & Dunning, 2000). Trong một nghiên cứu gần đây, Tu (2024) đã kiểm định và phân hạng các nhân tố ảnh hưởng thu hút FDI tại CQGĐPT. Kết quả thực nghiệm cho thấy thị trường, nguồn nhân lực, độ mở nền kinh tế và sự phát triển của thị trường tài chính là những nhân tố chính thúc đẩy sự thu hút FDI tại CQGĐPT. Tuy nhiên, cách tiếp cận của các nghiên cứu này là phương pháp suy luận thống kê, giải thích các mối tương quan có thể có giữa các nhân tố trong mô hình. Trong khi đó, vấn đề dự báo dựa vào các động lực tác động thì phương pháp suy luận thống kê bị hạn chế trong dài hạn do các vấn đề giá trị trung bình thay đổi theo thời gian (Steinbuks, 2019) hay vấn đề dùng sai phân làm cho phương sai sai số tăng tuyến tính với chân trời dự báo (Forecast Horizon) của quá trình này (Clements & Hendry, 2001). Kết quả là các phương pháp tiếp cận khác ngoài kinh tế lượng cho vấn đề dự báo cũng cần được tiến hành nghiên cứu bổ sung. Ngoài ra, dự báo dòng FDI đóng vai trò quan trọng trong việc hoạch định chính sách khi các nguồn lực của CQGĐPT là có hạn để triển khai nhiệm vụ thu hút FDI. Do đó, các mô hình dự báo chính xác đóng góp những gợi ý cho các quốc gia nhằm xây dựng khung chính sách thu hút FDI phù hợp. Trước thực tế như vậy, nghiên cứu này triển khai mô hình học máy nhằm gia tăng chất lượng dự báo FDI vì khả năng phân loại dữ liệu kinh tế của các mô hình học máy tốt hơn các phương pháp kinh tế lượng truyền thống (Shobana & Umamaheswari, 2021). Ngoài ra, dự báo dòng FDI theo phương pháp học máy cho mẫu của CQGĐPT tới nay vẫn chưa được chú ý nhiều trong các nghiên cứu trước.

Nghiên cứu này đóng góp vào tài liệu hiện nay về thu hút FDI ở CQGĐPT ở hai khía cạnh. *Thứ nhất*, với đóng góp về học thuật, nghiên cứu này cung cấp thêm một góc nhìn dự báo FDI thông qua các nhân tố tác động bằng phương pháp học máy, điều mà theo hiểu biết hạn hẹp của tác giả là còn

nhiều khoảng trống cần thử nghiệm. *Thứ hai*, về mặt thực tiễn, CQGĐPT có thể phân loại mức độ tác động của các nhân tố để có biện pháp phù hợp điều chỉnh dòng FDI thông qua nhân tố được phân tích.

Phần tiếp theo, bài báo sẽ trình bày về các nghiên cứu liên quan. Phần 3 mô tả cách thức chi tiết áp dụng phương pháp kinh tế lượng (Differenced GMM – DGMM) và học máy gồm: Phương pháp mạng nơ-ron nhân tạo (Artificial Neural Networks – ANN) và phương pháp rừng ngẫu nhiên (Random Forest – RF). Phần 4 thảo luận về các phát hiện thực nghiệm từ mô hình đề xuất. Phần 5 rút ra một vài kết luận và các hướng nghiên cứu tiếp theo của chủ đề.

2. Các nghiên cứu liên quan

2.1. Các yếu tố ảnh hưởng tới FDI theo tiếp cận kinh tế lượng

Trọng tâm của việc ứng dụng học máy trong thu hút FDI là sự liên kết của nó với các lý thuyết và mô hình kinh tế đã được thiết lập. Mô hình chiết trung do Dunning (1980) đề xuất cho rằng FDI được thúc đẩy bởi sự kết hợp giữa lợi thế sở hữu, địa điểm và nội bộ hóa. Theo đó, có bốn loại hình FDI tìm kiếm các giá trị khác nhau gồm: (1) Tìm kiếm nguồn lực, (2) tìm kiếm thị trường, (3) tìm kiếm hiệu quả và (4) tìm kiếm tài sản chiến lược. CTĐĐQG có thể tham gia vào nhiều loại hình FDI và mục tiêu của họ có thể thay đổi tùy theo mức độ trưởng thành hoặc mức độ quốc tế hóa (Dunning & Lundan, 2008). Các nghiên cứu thực nghiệm chỉ ra các nhân tố ảnh hưởng tới thu hút vốn FDI đã được nghiên cứu rộng rãi (C. P. Nguyen & Lee, 2021; Popovici và cộng sự, 2021; Teixeira & Nascimento, 2019). Tùy theo góc độ tiếp cận, các nghiên cứu trước thể hiện tính đa dạng trong việc nhận diện các nhân tố ảnh hưởng FDI tùy theo khu vực, thị trường, mức độ phát triển của nước nhận FDI. Theo đó các nhân tố như: Độ lớn thị trường, nguồn lực lao động, độ mở thương mại, cơ sở hạ tầng, sự phát triển của thị trường tài chính thường có ý nghĩa trong việc thu hút FDI. Điều này gợi ý rằng có một sự thống nhất tương đối trong các nghiên cứu về chiều tác động và mức độ đóng góp của các nhân tố này trong việc thu hút dòng FDI. Ngoài ra, các nhân tố mới nổi liên quan tới các yếu tố đổi mới công nghệ (Dorakh, 2020), chất lượng giáo dục (Baltas và cộng sự, 2018) và cơ chế ưu đãi của chính phủ với dòng FDI (Traore và cộng sự, 2022) cũng đã được nghiên cứu.

2.2. Các yếu tố ảnh hưởng tới FDI theo tiếp cận học máy

Kỹ thuật học máy cung cấp phương tiện để vận hành và tinh chỉnh các cấu trúc lý thuyết bằng cách tận dụng các phân tích dự đoán để phân biệt các mô hình và mối quan hệ trong các bộ dữ liệu phức tạp. Hơn nữa, các thuật toán học máy dựa trên các nguyên tắc lý thuyết thống kê nhằm tối ưu hóa hiệu suất mô hình và nâng cao tính khái quát (Vapnik, 1999). Cụ thể, các phương pháp học máy được cho là có khả năng dự báo tốt, đặc biệt với tập dữ liệu nhiều chiều. Ngoài ra, học máy có khả năng khám phá tiềm năng chưa biết và cả các mối quan hệ phi tuyến tính phức tạp trong dữ liệu, mang lại độ chính xác dự đoán nâng cao và hiểu biết sâu sắc hơn về dữ liệu (Chronopoulos và cộng sự, 2023). Do đó, việc tích hợp các kỹ thuật học máy vào phân tích FDI đánh dấu một sự thay đổi đáng kể so với các phương pháp truyền thống. Các ứng dụng của học máy trong nghiên cứu FDI đã sử dụng các kỹ thuật như: Cây quyết định, rừng ngẫu nhiên, máy vector hỗ trợ và mạng thần kinh nhân tạo cho việc nghiên cứu nhân tố ảnh hưởng tới FDI theo từng quốc gia riêng rẽ (Alon và cộng sự, 2022; Arel-Bundock, 2017; Gasanova và cộng sự, 2017; Singh, 2021). Điểm chung của các nghiên cứu này

thể hiện xu hướng rằng học máy thể hiện độ chính xác cao trong nhiệm vụ dự báo. Cụ thể, Arel-Bundock (2017) đặt giả thuyết về khả năng giải thích thực sự của các biến số chính trị trong việc dự báo dòng vốn FDI. Thông qua tập dữ liệu ở cấp độ công ty bao gồm các quyết định đầu tư của các tập đoàn đa quốc gia và áp dụng các kỹ thuật học máy phi tham số (cây quyết định tăng cường, rừng ngẫu nhiên, Ada Boost) và thực hiện các bài kiểm tra ngoài mẫu. Kết quả cho thấy các biến số chính trị không tác động đáng kể tới FDI và không nâng cao khả năng dự đoán của mô hình. Nghiên cứu cũng chỉ ra rằng các mô hình sử dụng các biến như quy mô kinh tế và khoảng cách giữa các quốc gia sẽ dự đoán FDI chính xác hơn. Singh (2021) xác định các yếu tố quyết định chính ảnh hưởng đến FDI của Hungary giai đoạn 2001–2018 bằng cách kiểm tra và so sánh độ chính xác dự báo thông qua ba mô hình: (1) Mô hình tuyến tính tổng quát (General Linear Model – GLM), (2) Mô hình độ dốc tăng cường (Gradient Boosting Machine – GBM), và (3) Rừng ngẫu nhiên (RF). Kết quả cho thấy GBM vượt trội hơn GLM và RF trong việc dự đoán các yếu tố quyết định FDI. Liên quan tới nhân tố ảnh hưởng, giá trị sản xuất trong một khu vực là yếu tố quyết định quan trọng nhất đối với dòng vốn FDI. Các yếu tố quan trọng khác bao gồm năng suất lao động và đô thị hóa. Alon và cộng sự (2022) trình bày các chỉ số cụ thể theo ngành về năng lượng, dịch vụ tài chính, công nghệ thông tin và truyền thông. Về mặt phương pháp, nghiên cứu cho thấy mô hình học máy tự động mang lại lợi thế so với các mô hình tuyến tính truyền thống bằng cách chọn mô hình dự đoán tốt nhất từ nhiều lựa chọn thay thế. Trong đó, nghiên cứu sử dụng quy trình ba giai đoạn bao gồm: (1) Phân tích nhân tố để xác định các tính năng liên quan, (2) mô hình học máy tự động để xác định tầm quan trọng của từng yếu tố và (3) tạo chỉ số hiệu chỉnh. Từ đó, mô hình học máy tự động tận dụng sức mạnh của học máy để cải thiện độ chính xác dự đoán và cung cấp những hiểu biết sâu sắc hơn về các yếu tố quyết định FDI.

Tuy nhiên, có rất ít nghiên cứu về chủ đề FDI theo một nhóm quốc gia hoặc nhóm khu vực. Điều này có thể do tính không đồng nhất tiềm ẩn của từng quốc gia trong sự kết hợp dữ liệu chéo và dữ liệu chuỗi thời gian trong dữ liệu bảng, làm cho việc kiểm soát tính không đồng nhất trở nên khó khăn với các phương pháp học máy (Chronopoulos và cộng sự, 2023). Do đó, còn nhiều khoảng trống nghiên cứu cần bổ sung cho chủ đề thu hút vốn FDI tiếp cận với các phương pháp học máy. Từ đó, nghiên cứu này tập trung vào hai mục tiêu cơ bản.

Một là, nghiên cứu này rút ra các nhân tố ảnh hưởng tới FDI với mẫu của 66 CQGĐPT dựa trên cách tiếp cận kinh tế lượng thông qua DGMM (Arellano & Bond, 1991). Kết quả đầu ra là chỉ ra được các nhân tố thu hút FDI của CQGĐPT. *Hai là*, nghiên cứu này so sánh chất lượng dự báo của các mô hình học máy (ANN, RF) với nhau và so sánh với phương pháp DGMM nhằm rút ra kết luận về tính chính xác dự báo của các phương pháp học máy. Điều này được kỳ vọng làm tăng chất lượng dự báo FDI khi đã xác định được nhóm các nhân tố tác động tới FDI từ mô hình kinh tế lượng. Kết quả của nghiên cứu sẽ xác định một cách định lượng độ lệch của giá trị dự báo với giá trị thật của FDI, từ đó rút ra các khuyến nghị liên quan tới mô hình đề xuất.

3. Phương pháp nghiên cứu và dữ liệu nghiên cứu

3.1. Phương pháp kinh tế lượng

Kế thừa các mô hình lý thuyết chiết trung (Dunning, 1973), lợi thế so sánh (Kojima & Ozawa, 1985), mô hình tăng trưởng nội sinh (Aghion & Howitt, 1998) và mô hình tân cổ điển (Solow, 1999),

mô hình nghiên cứu thực nghiệm đề xuất trong nghiên cứu này được trình bày dưới dạng tổng quát như sau:

$$FDI_{it} = \theta_0 + \theta_1 MAR_{it} + \theta_2 TRA_{it} + \theta_3 LAV_{it} + \theta_4 LQU_{it} + \theta_5 RES_{it} + \theta_6 IFS_{it} + \theta_7 FID_{it} + \theta_8 TEC_{it} + \theta_9 INF_{it} \quad (1)$$

Trong đó, các biến đo lường, cách tính, kỳ vọng dấu và nguồn tham khảo được trình bày tại Bảng 1 như sau:

Bảng 1.

Mô tả biến và kỳ vọng dấu của nghiên cứu

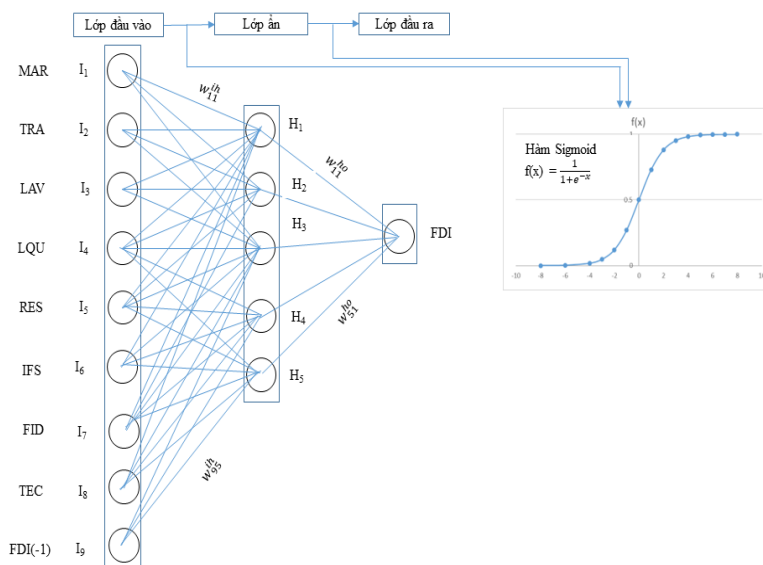
Biến	Tên biến	Cách tính	Nguồn	Kỳ vọng dấu
Đầu tư trực tiếp nước ngoài	FDI	FDI/GDP (%)	Nguyen và cộng sự (2022)	
Độ lớn thị phần	MAR	GDP tính theo hiện giá (logarit)	Asiedu và Lien (2004)	+
Độ mở thương mại	TRA	(Xuất khẩu + Nhập khẩu) / GDP (%)	Popovici và cộng sự (2021)	+
Lao động (theo sự sẵn có)	LAV	Tổng lực lượng lao động từ 15 tuổi trở lên (logarit)	Rodriguez và Pallas (2008)	+
Lao động (chất lượng)	LQU	Tỷ lệ tuyển sinh đại học (%)	Neugebauer và Schindler (2012)	+
Tài nguyên thiên nhiên	RES	Tổng tiền thuê tài nguyên thiên nhiên / GDP (%)	Teixeira và Nascimento (2019)	+
Cơ sở hạ tầng	IFS	Số lượng thuê bao di động (tính trên 100 người) (%)	Popovici và cộng sự (2021)	+
Sự phát triển thị trường tài chính	FID	Chỉ số phát triển tài chính (điểm số)	Nguyen và Lee (2021)	+
Sự phát triển của công nghệ	TEC	Chỉ số đổi mới sáng tạo (điểm số)	Tu (2024)	+
Lạm phát	INF	Chỉ số giá tiêu dùng (%)	Baltas và cộng sự (2018)	-

Để ước lượng phương trình (1) nghiên cứu sử dụng phương pháp GMM để giải quyết các vấn đề nội sinh tiềm ẩn có thể phát sinh trong phương trình (1) (xem thêm phần giải thích lý do chọn GMM trong Tu (2024)). Với cách tiếp cận này, kết quả ước lượng các hệ số của phương trình (1) được kỳ vọng không chệch. Lưu ý rằng nghiên cứu này chỉ sử dụng các biến có ý nghĩa thống kê để dự báo cho cả mô hình kinh tế lượng và các mô hình học máy bởi vì các biến có ý nghĩa thống kê là chỉ dấu tốt cho vấn đề dự báo (Beyene và cộng sự, 2009) cũng như giảm chiều nhằm tăng hiệu suất dự báo (Guyon & Elisseeff, 2003).

3.2. Phương pháp học máy

3.2.1. Mô hình mạng nơ-ron nhân tạo (ANN)

Một ANN cơ bản bao gồm ba lớp: Lớp đầu vào (Input), lớp ẩn (giữa) (Hidden) và lớp đầu ra (Output) (Rosenblatt, 1958). Lớp đầu vào là nơi nhận các dữ liệu đầu vào với mỗi nút là một đầu vào để huấn luyện mô hình. Trong nghiên cứu này, số nút lớp đầu vào là số biến độc lập trong mô hình kinh tế lượng. Lớp ẩn dùng để xử lý thông tin và truyền các kết quả đến các lớp tiếp theo. Số lượng lớp ẩn tùy thuộc vào độ phức tạp của mô hình và kỳ vọng về sự đánh đổi giữa thời gian đào tạo mô hình với độ chính xác của mô hình. Do tính chất của bộ dữ liệu không phức tạp, nghiên cứu này dùng một lớp ẩn và số lượng nơ-ron trong lớp ẩn cũng được thử nghiệm để báo cáo kết quả tốt nhất. Lớp đầu ra sẽ cung cấp kết quả dự báo cuối cùng. Trong nghiên cứu này, đầu ra chỉ có 1 biến là biến phụ thuộc FDI. Hàm kích hoạt giữa các lớp được thử nghiệm với hàm Sigmoid: $f(x) = \frac{1}{1+e^{-x}}$; hàm Tanh: $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, và hàm ReLU: $f(x) = \max(0, x)$.



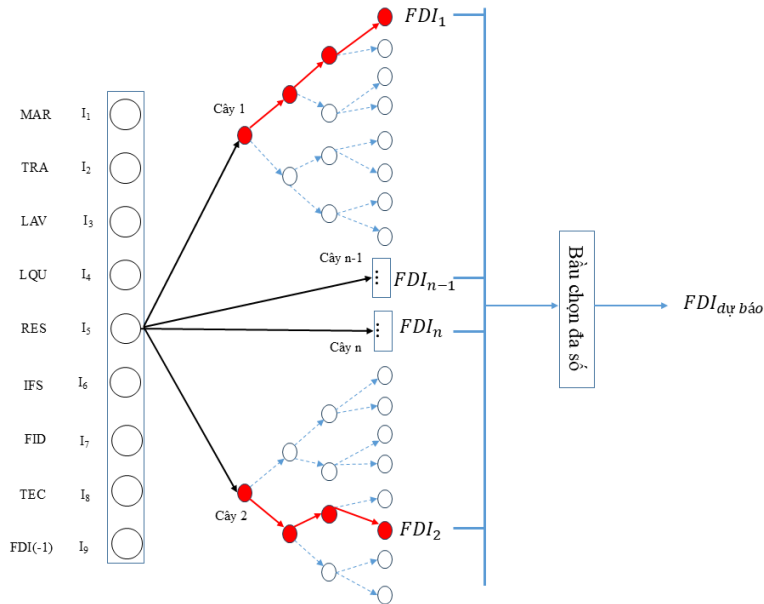
Hình 1. Thiết kế mạng nơ-ron với hàm kích hoạt Sigmoid được thực hiện trong nghiên cứu

3.2.2. Mô hình Rừng ngẫu nhiên (RF)

Rừng ngẫu nhiên là một tập hợp các cây quyết định (Decision Tree) được xây dựng ngẫu nhiên, mỗi cây được huấn luyện trên một tập dữ liệu con và một tập hợp tính năng con được lấy từ một bộ dữ liệu chung. Mỗi cây sẽ đưa ra quyết định (hay dự đoán) và tạo ra các kết quả dựa trên việc bỏ phiếu của đa số quyết định giống nhau từ các cây (Breiman, 2001). Nói cách khác, RF tạo ra các cây quyết định khác nhau trong giai đoạn đào tạo và quyết định đầu ra với đa số phiếu bầu (Liu và cộng sự, 2012).

Về bản chất, RF cải thiện độ chính xác dự đoán tổng thể và độ mạnh mẽ của quy trình phân loại nhị phân bằng cách tận dụng khả năng ra quyết định tập thể của nhiều cây huấn luyện trên các tập hợp con dữ liệu và bộ dự đoán khác nhau. Trong nghiên cứu này, các tham số được điều chỉnh là

“max_depth” và “max_feature”, trong đó, “max_depth” xem xét vấn đề khớp quá mức (Overfitting) và “max_feature” tìm kiếm sự phân chia tốt nhất để cải thiện kết quả (Sameen và cộng sự, 2019). Nghiên cứu trình bày kết quả tốt nhất của tất cả cài đặt trong mô hình RF.



Hình 2. Thiết kế rừng ngẫu nhiên được thực hiện trong nghiên cứu

3.3. Các thông số đo lường chất lượng dự báo

Nghiên cứu này sử dụng hai thông số để đo lường chất lượng dự báo là sai số toàn phương trung bình (Mean Squared Error – MSE) và sai số tuyệt đối trung bình (Mean Absolute Error – MAE) theo công thức như sau:

$$MSE = \frac{1}{n} \sum_{t=1}^n (A_t - P_t)^2 \quad (2)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |A_t - P_t| \quad (3)$$

Trong đó, n: Số quan sát; A_t và P_t lần lượt là giá trị thực và giá trị dự đoán của FDI tại thời điểm t. Cả hai chỉ số MSE và MAE càng nhỏ càng cho thấy chất lượng dự báo tốt. Ngoài ra, MAE còn giảm tác động của các điểm ngoại vi có trọng số cao (Bermejo & Cabestany, 2001).

3.4. Dữ liệu nghiên cứu

Về nguồn dữ liệu, các biến trong mô hình được trích xuất từ World Bank¹. Dữ liệu được thu thập cho 66 quốc gia đang phát triển (các quốc gia trong mẫu dựa vào phân loại thu nhập thấp, trung bình thấp, và trung bình của World Bank) giai đoạn 2013–2021. Mô hình hồi quy DGMM được thực hiện

¹ Truy cập dữ liệu tại: <https://databank.worldbank.org/source/world-development-indicators>

trên phần mềm Eviews 12, trong khi các mô hình ANN và RF được thực hiện dựa trên mã lệnh ngôn ngữ lập trình Python².

Tổng cộng có 594 quan sát được ghi nhận theo dạng bảng cân bằng cho toàn bộ mẫu được sử dụng để tìm kết quả của mô hình hồi quy DGMM. Trong khi đó, với mô hình ANN và RF, nghiên cứu này chia tập huấn luyện và tập thử nghiệm theo nguyên tắc: Tập huấn luyện là các quan sát trong giai đoạn 2013–2019, còn các quan sát trong giai đoạn 2020–2021 được sử dụng cho tập thử nghiệm. Như vậy, tập huấn luyện có 461 quan sát (chiếm 78%) và tập thử nghiệm có 133 quan sát (chiếm 22%). Các mô hình ANN và RF được thiết kế bằng mã ngôn ngữ lập trình Python và thực thi trên môi trường Google Colab. Để các mô hình ANN và RF có sự ổn định trong việc cập nhật giá trị tối ưu, nghiên cứu này cài đặt mỗi thuật toán vòng lặp 50 lần.

4. Kết quả nghiên cứu và thảo luận

4.1. Mô hình kinh tế lượng

4.1.1. Thống kê mô tả

Các biến mô tả trong nghiên cứu này được thể hiện trong Bảng 2. Kết quả cho thấy giá trị FDI/GDP trung bình tại khu vực CQGĐPT giai đoạn 2013–2021 là 3,64%, cho thấy FDI chiếm tỷ trọng tương đối lớn so với quy mô của CQGĐPT. Ngoài ra, độ lệch chuẩn của FDI là 4,09%, cao hơn giá trị trung bình của cả giai đoạn, cho thấy có sự chênh lệch về thu hút FDI tại CQGĐPT.

Bảng 2.

Thống kê mô tả các biến trong mô hình

	FDI	FID	IFS	INF	LAV	LQU	MAR	RES	TEC	TRA
	(%)	(điểm)	(%)	(%)	(Logarit)	(%)	(Logarit)	(%)	(điểm)	(%)
Trung bình	3,64	0,26	108,21	6,40	16,01	33,73	24,86	5,05	29,67	70,58
Trung vị	2,59	0,23	111,68	3,60	15,86	30,04	24,56	2,90	29,35	63,67
Lớn nhất	39,46	0,74	182,58	557,20	20,49	101,42	30,51	35,3	54,80	186,47
Nhỏ nhất	0,01	0,06	26,42	-3,75	12,44	2,08	19,77	0,01	16,70	16,35
Độ lệch	4,09	0,15	31,23	26,48	1,51	22,60	1,70	5,47	6,46	30,73
Số quan sát	594	594	594	594	594	594	594	594	594	594

4.1.2. Hệ số tương quan giữa các biến

Về kiểm định tương quan đối với các biến đưa vào mô hình đề xuất, Bảng 3 cho thấy một cặp biến độc lập có hệ số tương quan cao $r(\text{MAR}, \text{LAV}) = 0,837$, điều này cho thấy có thể có sự trùng lặp giữa MAR và LAB. Nói cách khác, mô hình có nguy cơ xảy ra vấn đề đa cộng tuyến (Gujarati & Porter, 2009). Do đó, nghiên cứu đã kiểm tra kiểm định Wald để khẳng định sự cần thiết của hai biến này trong mô hình đề xuất. Kết quả cho thấy giá trị F (MAR, LAV) = 18,16 ở mức nghĩa thống kê

² Tham khảo chi tiết với dữ liệu của bài báo tại Mendeley: <https://data.mendeley.com/datasets/3yrx8xxccb/2>

1%, ngụ ý rằng có thể bác bỏ giả thuyết H_0 ($a = b = 0$), cho thấy không có sự thừa biến trong mô hình đề xuất.

Bảng 3.

Ma trận tương quan giữa các biến

Hệ số tương quan	FDI	FID	IFS	INF	LAV	LQU	MAR	RES	TEC	TRA
FDI	1									
FID	0,10**	1								
IFS	-0,02	0,46***	1							
INF	-0,04	-0,47	-0,08**	1						
LAV	-0,22**	0,37***	-0,15***	0,01	1					
LQU	0,01	0,39***	0,52***	-0,52	-0,09**	1				
MAR	-0,27***	0,64***	0,14***	-0,03	0,84***	0,27***	1			
RES	0,16***	-0,18***	-0,19***	0,03	0,05	-0,15***	-0,07*	1		
TEC	0,06	0,68***	0,50***	-0,12***	0,16***	0,63***	0,44***	-0,24***	1	
TRA	0,37***	0,13***	0,37***	-0,04	-0,43***	0,29***	-0,30***	-0,05	0,30***	1

Ghi chú: *, **, *** lần lượt tương ứng với các mức ý nghĩa thống kê 10%, 5% và 1%.

4.1.3. Kiểm định nghiệm đơn vị

Bảng 4 cho thấy các biến đưa vào mô hình dừng ở bậc gốc với ít nhất một phương pháp kiểm định. Cụ thể, các biến ít nhất dừng ở hai phương pháp ADF-Fisher và PP-Fisher (Maddala & Wu, 1999). Điều này cung cấp bằng chứng cho thấy các biến trong mô hình có bước đi ngẫu nhiên, ngụ ý rằng các giá trị ước lượng (trung bình, phương sai) ổn định theo thời gian.

Bảng 4.

Kết quả kiểm định nghiệm đơn vị

	Phương pháp Im-Pesaran-Shin		Phương pháp ADF-Fisher		Phương pháp PP-Fisher	
	Hệ số chặn	Hệ số chặn & xu hướng	Hệ số chặn	Hệ số chặn & xu hướng	Hệ số chặn	Hệ số chặn & xu hướng
FDI	-1,12	-0,62	163,03**	165,55**	305,23***	370,34***
FID	0,86	0,63	129,16	131,92	195,07***	330,42***
IFS	0,96	-0,41	128,47	164,91**	210,82***	234,22***
INF	-5,57***	-0,96	270,38***	192,79***	297,25***	277,21***
LAV	0,92	3,14	124,97	91,95	214,96***	201,30***
LQU	0,14	0,76	155,93*	128,31	188,52***	207,46***
MAR	0,13	-1,94**	138,23	220,57***	137,10	167,38**

	Phương pháp Im-Pesaran-Shin		Phương pháp ADF-Fisher		Phương pháp PP-Fisher	
RES	-3,33***	-0,35	222,38***	174,07***	212,82***	98,47
TEC	3,94	1,04	139,02**	120,21	173,40***	364,99***
TRA	-2,75***	-0,75	190,42***	181,67***	258,59***	189,70***

Ghi chú: *, **, *** lần lượt tương ứng với các mức ý nghĩa thống kê 10%, 5% và 1%.

4.1.4. Kết quả hồi quy

Bảng 5 (3 dòng đầu tiên) cho thấy hệ số ước lượng γ của FDI(-1) của 3 phương pháp POLS, DGMM, và FEM lần lượt là 0,75; 0,37; và 0,23. Như vậy, hệ số ước lượng theo phương pháp DGMM nằm trên ngưỡng giới hạn dưới (phương pháp FEM). Do đó, ước lượng theo phương pháp DGMM là phù hợp trong mẫu nghiên cứu.

Các kiểm định tăng cường như kiểm định J Sargan-Hansen và Arellano-Bond (AR) xác nhận độ tin cậy của phương pháp DGMM. Cụ thể, $P(\text{kiểm định } J) = 0,225$, ngụ ý rằng không bác bỏ được giả thuyết H_0 về khả năng các biến công cụ là các biến ngoại sinh chặt. Kết quả của xác suất AR(1) là 0,0058 và AR(2) là 0,6438 cung cấp bằng chứng cho thấy có tự tương quan bậc 1 và không có tự tương quan bậc 2 trong mô hình. Như vậy, các hệ số ước lượng theo DGMM có độ tin cậy.

Kết quả thực nghiệm của các nhân tố ảnh hưởng FDI cho thấy ngoại trừ INF không có ý nghĩa thống kê, các biến còn lại đều có ý nghĩa thống kê và cùng kỳ vọng của nghiên cứu đã nêu tại Bảng 1. Trong đó, các biến quan trọng như: Độ lớn thị trường, nguồn lao động, ngoại thương, tài nguyên, thị trường tài chính phát triển vẫn là những nhân tố chủ chốt ảnh hưởng rõ rệt tới thu hút FDI tại CQGDPT. Các phát hiện thực nghiệm này một lần nữa củng cố cho các lập luận của các lý thuyết liên quan về thu hút FDI như: mô hình chiết trung (Dunning, 1973), mô hình nội sinh (Aghion & Howitt, 1998), mô hình tân cổ điển (Solow, 1999) và các phát hiện thực nghiệm đã được dẫn chứng tại Bảng 1.

Bảng 5.

Kết quả hồi quy theo DGMM

Kiểm định chọn DGMM			
	γ_{POLS}	γ_{D-GMM}	γ_{FEM}
	0,7449	0,3742	0,2265
Hệ số ước lượng phương trình (1) theo DGMM			
Biến độc lập	Hệ số		
FDI(-1)	0,374***		
FID	11,047***		
IFS	0,017***		
INF	-0,001		
LAV	11,047**		
LQU	-0,034**		

Kiểm định chọn DGMM	
MAR	0,296**
RES	0,058***
TEC	0,074***
TRA	0,032***
Các kiểm định tăng cường	
Kiểm định J	32,193
P[Kiểm định J]	0,225
P[Tự tương quan bậc 1]	0,006
P[Tự tương quan bậc 2]	0,644

Ghi chú: **,*** lần lượt tương ứng với các mức ý nghĩa thống kê 5% và 1%.

4.1.5. Kết quả dự báo từ mô hình kinh tế lượng

Từ kết quả tham số ước lượng theo DGMM của phương trình (1) sau khi bỏ biến INF không có ý nghĩa thống kê, nghiên cứu này tính toán giá trị của FDI dự báo. Kết quả cho thấy sai số toàn phương trung bình (MSE) của phương pháp kinh tế lượng có giá trị là 48,5308(%)² và sai số tuyệt đối trung bình (MAE) có giá trị là 5,7348(%).

4.2. Mô hình học máy

4.2.1. Mô hình ANN

Dựa vào các biến đầu vào có ý nghĩa thống kê (ngoại trừ biến INF) tại Bảng 5, nghiên cứu này thiết kế mạng ANN gồm 3 lớp như mô tả tại Hình 1 với 9 biến đầu vào là các biến độc lập của phương trình (1) và 1 biến đầu ra (FDI). Các tham số được thử nghiệm theo các điều kiện được đặt ra để lựa chọn kết quả tốt nhất. Cụ thể, về số nút trong lớp ẩn, bài báo thử nghiệm số nút tối ưu trong lớp ẩn được tính toán theo các công thức tại các nghiên cứu trước (Hecht-Nielsen, 1987; Hunter và cộng sự, 2012; Turban và cộng sự, 2011). Về tham số tốc độ học (Learning Rate), tốc độ học chậm khiến kết quả hội tụ chậm trong khi tốc độ học nhanh sẽ khiến kết quả phân tán (Bengio, 2012; Darken & Moody, 1991). Có rất nhiều các thử nghiệm khác nhau cho tốc độ học, ví dụ: Witten và cộng sự (2017) cho rằng khoảng biến thiên tốt nhất là $[10^{-8}; 1]$, trong khi đó, Vandeginste và cộng sự (1998) cho rằng khoảng này nên là $[0,5; 1]$. Nghiên cứu này thực hiện theo khuyến nghị của Schaul và cộng sự (2013) cho rằng sai số giữa tập huấn luyện và tập kiểm định nhỏ nhất khi tốc độ học biến thiên trong khoảng $[10^{-1}; 10^0]$. Do đó, nghiên cứu này cài đặt tốc độ học ở các mức giá trị 0,01; 0,1; và 0,9 (Nguyễn và cộng sự, 2021). Các hàm kích hoạt được thử nghiệm với hàm Sigmoid, Tanh và ReLU³. Kết quả giá trị FDI dự báo tốt nhất theo ANN được trình bày tại Bảng 6.

Kết quả nhìn chung cho thấy khi tốc độ học là 0,01 thì các giá trị sai số toàn phương trung bình (MSE) và sai số tuyệt đối trung bình (MAE) nhỏ đáng kể, nhưng khi tăng tốc độ học lên 0,1 và 0,9 thì mức cải thiện của MSE và MAE không còn, ngụ ý rằng các tốc độ học trung bình và nhanh làm

³ Tham khảo chi tiết kết quả của ANN tại Mendeley: <https://data.mendeley.com/datasets/3yrx8xxccb/2>

thuật toán hội tụ nhanh hơn. Tuy nhiên, việc huấn luyện nhanh khiến hàm mất mát (Loss Function) dao động ngăn cản mô hình hội tụ đến giải pháp tối ưu, dẫn đến hiệu suất kém (Takase và cộng sự, 2018). Kết quả tối ưu thuộc về cài đặt với tốc độ học ở mức 0,01. Liên quan tới số nút tối ưu trong lớp ẩn, kết quả thực nghiệm chỉ ra số nút trong lớp ẩn bằng 10 như công thức tính của Hunter và cộng sự (2012). Lúc này, các giá trị MAE và MSE của tập thử nghiệm lần lượt là 2,8203(%) và 25,3174(%²). Các giá trị này đều nhỏ hơn tập huấn luyện, ngụ ý rằng chất lượng dự báo của tập thử nghiệm được cải thiện nhờ khả năng học từ tập huấn luyện (không có hiện tượng khớp quá mức).

4.2.2. Mô hình RF

Dựa vào các biến đầu vào có ý nghĩa thống kê (ngoại trừ biến INF) tại Bảng 5, nghiên cứu này thiết kế mạng RF gồm 100 cây như mô tả tại Hình 2 với 9 biến đầu vào là các biến độc lập của phương trình (1) và 1 biến đầu ra (FDI). Các tham số được thử nghiệm theo các điều kiện được đặt ra để lựa chọn kết quả tốt nhất⁴. Kết quả giá trị FDI dự báo tốt nhất theo RF được trình bày tại Bảng 6.

Kết quả nhìn chung cho thấy các giá trị độ sâu của từng cây (max_depth) và số lượng tính năng cần xem xét khi tìm kiếm sự phân chia tốt nhất (max_feature) không thay đổi nhiều khi xem xét thay đổi các giá trị này. Các cài đặt max_depth = 100 và max_feature = log2 cho kết quả MSE và MAE đạt giá trị nhỏ nhất. Cụ thể, bộ cài đặt này mang lại giá trị nhỏ nhất của giá trị MAE và MSE lần lượt là 0,0121 (%) và 0,0006 (%²). Các giá trị này cũng không khác biệt quá lớn với các giá trị nhận được trong các cài đặt khác, ngụ ý rằng mô hình RF đạt tới sự ổn định trả ra kết quả trong thử nghiệm. Ngoài ra, các giá trị MAE và MSE của tập thử nghiệm đều tốt hơn so với các giá trị này của tập huấn luyện, cho thấy mô hình RF không bị hiện tượng khớp quá mức. Điều này mang lại tính vững cho chất lượng mô hình dự báo.

4.3. Tổng hợp kết quả và thảo luận

Các giá trị FDI dự báo tốt nhất của 3 phương pháp được tổng hợp tại Bảng 6. Kết quả cho thấy các phương pháp học máy (ANN và RF) có chất lượng dự báo tốt hơn so với phương pháp kinh tế lượng (Basu & Ferreira, 2020; Nabipour và cộng sự, 2020; Shobana & Umamaheswari, 2021). Nguyên nhân có thể do các phương pháp học máy có thể nhận biết được các mẫu hình phức tạp hơn của dữ liệu so với cách tiếp cận kinh tế lượng (Basu & Ferreira, 2020). Cách tiếp cận dựa trên dữ liệu này phù hợp tốt với các bộ dữ liệu thử nghiệm khác nhau, trong đó không có hướng dẫn lý thuyết nào có sẵn để đề xuất quy trình tạo dữ liệu phù hợp.

Bảng 6.

Tổng hợp kết quả chất lượng dự báo FDI của các phương pháp tiếp cận

	DGMM	ANN	RF
MAE (%)	5,7348	2,8203	0,0121
MSE (% ²)	48,5308	25,3174	0,0006

Kết quả tại Bảng 6 cũng cho thấy RF đạt được chất lượng dự báo trên tập thử nghiệm tốt nhất và vượt trội so với các phương pháp đã trình bày. Điều này cho thấy RF là một phương pháp hữu hiệu áp dụng cho các bài toán hồi quy (Borup và cộng sự, 2023; Yoon, 2021). Kết quả này có thể tới từ

⁴ Tham khảo chi tiết kết quả của RF tại Mendeley: <https://data.mendeley.com/datasets/3yrx8xxccb/2>

nguyên nhân thuật toán RF tận dụng được sức mạnh của khả năng ra quyết định tập thể từ nhiều cây độc lập (trong các thử nghiệm, nghiên cứu này sử dụng 100 cây quyết định để dự báo giá trị FDI).

5. Kết luận và hướng nghiên cứu tiếp theo

Nghiên cứu này nhận diện các nhân tố chính ảnh hưởng tới thu hút FDI của 66 quốc gia đang phát triển giai đoạn 2013–2021. Trong đó, bài báo sử dụng một phương pháp kinh tế lượng DGMM để giải thích các mối quan hệ giữa các biến trong mô hình. Kết quả của phương pháp DGMM cung cấp bằng chứng thống kê cho thấy sự hiện diện của các biến độc lập được lựa chọn ảnh hưởng tới FDI rõ nét. Từ đó, bài báo lập hai mô hình học máy ANN và RF dựa trên các biến có ý nghĩa thống kê để dự báo FDI. Kết quả của các thử nghiệm cho thấy RF có giá trị vượt trội trong việc dự đoán FDI với chất lượng sai số của giá trị MSE và MAE nhỏ nhất trong các nhóm phương pháp.

Đóng góp của nghiên cứu này vào tài liệu là giới thiệu hướng tiếp cận áp dụng các phương pháp học máy nhằm dự báo FDI với mẫu của CQGĐPT. Trong đó, cách tiếp cận là dùng phương pháp kinh tế lượng làm nền tảng thiết lập các biến đầu vào cho các mô hình học máy dựa trên bằng chứng thống kê giữa các biến. Việc này giúp giảm chi phí thu thập dữ liệu cho các biến không có ý nghĩa trong mô hình học máy (Mullainathan & Spiess, 2017). Với cách tiếp cận này, các giá trị FDI dự đoán được kỳ vọng bám sát dòng FDI thực tế do các biến độc lập có tính giải thích cao với dòng FDI.

Tuy nhiên, hạn chế của nghiên cứu này là trong các mô hình ANN và RF, các thuật toán chưa phân biệt được các giá trị chéo của các quốc gia trong mẫu, thay vào đó, các thuật toán học vẫn xét toàn bộ dữ liệu như một chuỗi dữ liệu của một đơn vị chéo. Trong khi đó, từng quốc gia trong mẫu có mức chênh lệch giá trị giữa các biến cao. Điều này làm cho chất lượng dự báo có phần hạn chế do có sự dao động mạnh giữa các điểm dữ liệu quan sát. Các hướng nghiên cứu tiếp có thể tập trung vào việc xử lý tính không đồng nhất này để tăng chất lượng dự báo.

Tài liệu tham khảo

- Aghion, P., Howitt, P. (1998). *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Alon, I., Bretas, V. P., Sclip, A., & Paltrinieri, A. (2022). Greenfield FDI attractiveness index: A machine learning approach. *Competitiveness Review: An International Business Journal*, 32(7), 85-108.
- Arel-Bundock, V. (2017). The political determinants of foreign direct investment: A firm-level analysis. *International Interactions*, 43(3), 424-452.
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297.
- Asiedu, E., & Lien, D. (2004). Capital controls and foreign direct investment. *World Development*, 32(3), 479-490.
- Baltas, N., Tsionas, M. G., & Baltas, K. (2018). Foreign direct investment in OECD countries: A special focus in the case of Greece. *Applied Economics*, 50(52), 5579-5591.
- Basu, R., & Ferreira, J. (2020). Understanding household vehicle ownership in Singapore through a comparison of econometric and machine learning models. *Transportation Research Procedia*, 48, 1674-1693.

- Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In Montavon, G., Orr, G.B., Müller, K.R. (eds), *Neural Networks: Tricks of the Trade (2nd ed.)* (pp. 437-478). Berlin: Springer.
- Bermejo, S., & Cabestany, J. (2001). Oriented principal component analysis for large margin classifiers. *Neural Networks*, 14(10), 1447-1461.
- Beyene, J., Atenafu, E. G., Hamid, J. S., To, T., & Sung, L. (2009). Determining relative importance of variables in developing and validating predictive models. *BMC Medical Research Methodology*, 9(64), 1-10.
- Borup, D., Christensen, B. J., Mühlbach, N. S., & Nielsen, M. S. (2023). Targeting predictors in random forest regression. *International Journal of Forecasting*, 39(2), 841-868.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Cheng, L., & Kwan, Y. K. (2000). The location of foreign direct investment in Chinese regions: Further analysis of labor quality. In Ito, T., Krueger, A. O. (Ed.), *The Role of Foreign Direct Investment in East Asian Economic Development* (pp. 213-238). University of Chicago Press.
- Chronopoulos, I., Chrysikou, K., Kapetanios, G., Mitchell, J., & Raftapostolos, A. (2023). *Deep neural network estimation in panel data models*. FRB of Cleveland Working Paper No. 23-15. Available at SSRN: <https://ssrn.com/abstract=4501438>
- Clements, M. P., & Hendry, D. F. (2001). *Forecasting Non-Stationary Economic Time Series*. The MIT Press.
- Coughlin, C. C., & Segev, E. (2000). Foreign direct investment in China: A spatial econometric study. *The World Economy*, 23(1), 1-23.
- Crescenzi, R., Ganau, Roberto, & Storper, M. (2022). Does foreign investment hurt job creation at home? The geography of outward FDI and employment in the USA. *Journal of Economic Geography*, 22(1), 53-79.
- Darken, C., & Moody, J. (1991). *Towards faster stochastic gradient search*. In Proceedings of the 4th International Conference on Neural Information Processing Systems (NIPS'91). Morgan Kaufmann Publishers Inc. (pp. 1009-1016), San Francisco, CA, USA.
- Dorakh, A. (2020). FDI determinants in the european union: Empirical application. *Scientific Annals of Economics and Business*, 67(2), 219-232.
- Dunning, J. H. (1973). The determinants of international production. *Oxford Economic Papers*, 25(3), 289-336.
- Dunning, J. H. (1980). Toward an eclectic theory of international production: Some empirical tests. *Journal of International Business Studies*, 11, 9-31.
- Dunning, J. H., & Lundan, S. M. (2008). *Multinational Enterprises and the Global Economy*. Edward Elgar Publishing.
- Gasanova, A., Medvedev, A. N., & Komotskiy, E. I. (2017). The assessment of corruption impact on the inflow of foreign direct investment. *AIP Conference Proceedings*, 1836(1). <https://doi.org/10.1063/1.4981951>
- Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometric*. McGraw-Hill Irwin.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.

- Head, K., & Ries, J. (1996). Inter-city competition for foreign investment: Static and dynamic effects of China's incentive areas. *Journal of Urban Economics*, 40(1), 38-60.
- Hecht-Nielsen, R. (1987). *Kolmogorov's mapping neural network existence theorem*. Paper presented at the Proceedings of the international conference on Neural Networks (pp. 11-13), San Diego, CA.
- Hunter, D., Yu, H., Pukish III, M. S., Kolbusz, J., & Wilamowski, B. M. (2012). Selection of proper neural network sizes and architectures – A comparative study. *IEEE Transactions on Industrial Informatics*, 8(2), 228-240.
- Keller, W. (2021). *Knowledge spillovers, trade, and FDI* (NBER Working Paper 28739). Retrieved from the National Bureau of Economic Research: <https://www.nber.org/papers/w28739>
- Kojima, K., & Ozawa, T. (1985). Toward a theory of industrial restructuring and dynamic comparative advantage. *Hitotsubashi Journal of Economics*, 26(2), 135-145.
- Liu, Y., Wang, Y., & Zhang, J. (2012). *New machine learning algorithm: Random forest*. In Liu, B., Ma, M., Chang, J. (eds), Information Computing and Applications. ICICA 2012, Lecture Notes in Computer Science, 7473, Berlin, Heidelberg: Springer.
- Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, 61(S1), 631-652.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.
- Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *Ieee Access*, 8, 150199-150212.
- Narula, R., & Dunning, J. H. (2000). Industrial development, globalization and multinational enterprises: New realities for developing countries. *Oxford Development Studies*, 28(2), 141-167.
- Neugebauer, M., & Schindler, S. (2012). Early transitions and tertiary enrolment: The cumulative impact of primary and secondary effects on entering university in Germany. *Acta Sociologica*, 55(1), 19-36.
- Nguyen, A.-T., Lu, S.-H., & Nguyen, P. T. T. (2021). Validating and forecasting carbon emissions in the framework of the environmental Kuznets curve: The case of Vietnam. *Energies*, 14(11), 3144.
- Nguyen, A.-T., Anwar, S., Alexander, W. R. J., & Lu, S.-H. (2022). Openness to trade, foreign direct investment, and economic growth in Vietnam. *Applied Economics*, 54(29), 3373-3391.
- Nguyen, C. P., & Lee, G. S. (2021). Uncertainty, financial development, and FDI inflows: Global evidence. *Economic Modelling*, 99, 105473.
- Popovici, O. C., Călin, A. C., Ivana, D., & Sorin, D. (2021). FDI determinants revisited: Extensive evidence. *Romanian Journal of Economic Forecasting*, 24(4), 103-123.
- Rodriguez, X. A., & Pallas, J. (2008). Determinants of foreign direct investment in Spain. *Applied Economics*, 40(19), 2443-2450.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386-408.
- Sameen, M. I., Pradhan, B., & Lee, S. (2019). Self-learning random forests model for mapping groundwater yield in data-scarce areas. *Natural Resources Research*, 28, 757-775.

- Schaul, T., Zhang, S., & LeCun, Y. (2013). *No more pesky learning rates*. Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, 2013.
- Shobana, G., & Umamaheswari, K. (2021). *Forecasting by machine learning techniques and econometrics: A review*. Paper presented at the 2021 6th international Conference On Inventive Computation Technologies (ICICT), Coimbatore, India, 20-22 January 2021. <https://ieeexplore.ieee.org/document/9358514>
- Singh, D. (2021). Interpretable machine-learning approach in estimating FDI inflow: Visualization of ML models with LIME and H2O. *TalTech Journal of European Studies*, 11(1), 133-152.
- Solow, R. M. (1999). Neoclassical growth theory. *Handbook of Macroeconomics*, 1, 637-667.
- Steinbuks, J. (2019). Assessing the accuracy of electricity production forecasts in developing countries. *International Journal of Forecasting*, 35(3), 1175-1185.
- Takase, T., Oyama, S., & Kurihara, M. (2018). Effective neural network training with adaptive learning rate based on training loss. *Neural Networks*, 101, 68-78.
- Teixeira, A. A., & Nascimento, S. (2019). FDI in small states and unattractive markets: The case of São Tomé and Príncipe. *Journal of Contemporary African Studies*, 37(4), 351-378.
- Traore, Y., Muchie, M., & Worku, Z. (2022). South African ambitious investment objectives: The current foreign direct investment determinants in South Africa. *African Journal of Science, Technology, Innovation and Development*, 14(1), 31-48.
- Tu, N. A. (2024). Innovation and foreign direct investment attraction in developing countries. *Cogent Economics & Finance*, 12(1), 2312386.
- Turban, E., Sharda, R., & Delen, D. (2011). *Decision Support and Business Intelligence Systems* (9th ed.). Pearson Education India.
- Vandeginste, B. G. M., Massart, D. L., Buydens, L. M. C., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1998). Artificial neural networks. In B. G. M. Vandeginste, D. L. Massart, L. M. C. Buydens, S. De Jong, P. J. Lewi, & J. Smeyers-Verbeke (Eds.), *Data Handling in Science and Technology*, 20 (pp. 649-699). Elsevier.
- Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5), 988-999.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Deep learning. In I. H. Witten, E. Frank, M. A. Hall, & C. J. Pal (Eds.), *Data Mining* (4th ed.) (pp. 417-466). Morgan Kaufmann.
- Yao, S. (2006). On economic growth, FDI and exports in China. *Applied Economics*, 38(3), 339-351.
- Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Computational Economics*, 57(1), 247-265.