



Semnan University

Journal of Modeling in Engineering

Journal homepage: <https://modelling.semnan.ac.ir/>

ISSN: 2783-2538



Research Article

Statistical Data Analysis of Legatum Prosperity Index and Presentation of Prediction Models from Data Mining Process

Ali Naghash Asadi ^{a,*}, Amin Mastourinejad ^a, Romina Aghili ^a

^a Fouman Faculty of Engineering, College of Engineering, University of Tehran, Tehran, Iran

PAPER INFO

Paper history:

Received: 2024-08-02

Revised: 2025-06-03

Accepted: 2025-08-25

Keywords:

Statistical data analysis;
Legatum prosperity index;
Prediction models;
Neural network;
Regression.

ABSTRACT

Data mining is a process to discover patterns from a dataset and it can play an important role in understanding the multifaceted aspects of development on a global scale. Policymakers and researchers can formulate more effective strategies to promote sustainable development by identifying key indexes of prosperity, their interrelationships, and the impact of each index on other indexes. In this paper, first, 12 indexes of prosperity provided by Legatum institute are introduced. Then the values of these indexes are collected for 167 countries and statistical analyzes are presented. In the following, a heat map of the mutual influence of indexes is presented, which shows the influence of each index on other indexes. Finally, prediction models are presented for each index using regression and neural network methods, and they will be evaluated based on different measures. The values of the evaluation measures prove that the results from the prediction models were up to 98% similar to the actual results in the best conditions. Therefore, prediction models can determine how much other indexes will change with the improvement of one index. Based on this, policymakers and researchers can plan and prioritize to improve the status of indexes.

DOI: <https://doi.org/10.22075/jme.2025.34932.2714>

© 2026 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license.(<https://creativecommons.org/licenses/by/4.0/>)

* Corresponding author.

E-mail address: naghashasadi@ut.ac.ir

How to cite this article:

A Naghash Asadi, A Mastourinejad and R Aghili, "Statistical Data Analysis of Legatum Prosperity Index and Presentation of Prediction Models from Data Mining Process," Journal of Modeling in Engineering, 24 85 (2026): 67-84, doi: 10.22075/jme.2025.34932.2714

تحلیل داده آماری از شاخص رفاه لگاتوم و ارائه مدل های پیش بینی از فرآیند داده کاوی

علی نقاش اسدی^{۱*}، امین مستوری نژاد^۱، رومینا عقیلی^۱

اطلاعات مقاله	چکیده
دریافت مقاله: ۱۴۰۳/۰۵/۱۲ بازنگری مقاله: ۱۴۰۴/۰۳/۱۳ پذیرش مقاله: ۱۴۰۴/۰۶/۰۳	داده کاوی، فرآیندی برای کشف الگوها از مجموعه ای از داده ها بوده و می تواند نقش مهمی در درک جنبه های چندوجهی توسعه و شکوفایی در مقیاس جهانی ایفا کند. با شناسایی عوامل کلیدی تعیین کننده رفاه و روابط متقابل آنها، سیاست گذاران و محققان می توانند استراتژی های مؤثرتری را برای تقویت توسعه پایدار تدوین کنند و میزان تاثیر هر شاخص بر شاخص های دیگر را دریافت کنند. در این مقاله، ابتدا ۱۲ شاخص توسعه و رفاه ارائه شده توسط موسسه لگاتوم معرفی می شود. سپس مقادیر این شاخص ها برای ۱۶۷ کشور جمع آوری شده و تحلیل های آماری از آنها ارائه می شود. در ادامه، نقشه حرارتی از تاثیر متقابل شاخص ها نسبت به هم ارائه می شود که نشان دهنده میزان تاثیرگذاری هر شاخص بر شاخص های دیگر است. در انتها با استفاده از روش های رگرسیون و شبکه عصبی، برای هر شاخص، مدل های پیش بینی ارائه می شود و بر اساس معیارهای ارزیابی مختلف، مورد ارزیابی قرار خواهند گرفت. مقادیر معیارهای ارزیابی نشان می دهند که نتایج حاصل از مدل های پیش بینی در بهترین شرایط تا ۹۸ درصد با نتایج واقعی مشابهت داشتند؛ بنابراین با استفاده از مدل های پیش بینی ارائه شده می توان دریافت که با بهبود یک شاخص در یک کشور، سایر شاخص ها به چه میزانی تغییر خواهند کرد و بر همین اساس، سیاست گذاران و محققان می توانند برای بهبود وضعیت شاخص ها، برنامه ریزی و اولویت بندی کنند.

واژگان کلیدی:

تحلیل داده آماری،
شاخص رفاه لگاتوم،
مدل های پیش بینی،
شبکه عصبی،
رگرسیون.DOI: <https://doi.org/10.22075/jme.2025.34932.2714>

© 2026 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

۱- مقدمه

امروزه، سازمان ها و ادارات هر کشور، حجم زیادی از داده ها را در زمینه های مختلف دارا هستند. علاوه بر آن، سازمان های مختلفی مانند سازمان ملل متحد، بانک جهانی و سایر سازمان های بین المللی وجود دارند که داده های مختلفی از شاخص های توسعه، رفاه و شکوفایی کشورها ارائه می کنند. ولی با این حال، سیاست گذاران و مسئولین کشورها، با فقر دانش در تصمیم گیری روبرو هستند [۱]. در صورتی که می توان از این داده ها برای تحلیل آماری استفاده کرده و با استفاده از فرآیندهای داده کاوی، مدل ها و الگوهای پیش بینی از آنها استخراج کرد که این مدل ها، امکان

تصمیم سازی درست را فراهم می کنند. یافته های این فرآیند، به درک عمیق تر روندها و نابرابری های توسعه جهانی کمک می کند. فرآیند داده کاوی به عنوان راه حل این مسئله، شامل مجموعه ای از روش ها است که به منظور کشف الگوهای پنهان در میان داده ها استفاده می شود [۲]. در حوزه هایی مانند علوم کامپیوتر، اقتصاد، پزشکی، مهندسی و غیره، از فرآیند داده کاوی در پژوهش های مختلف استفاده شده است. استفاده از فرآیند داده کاوی در حوزه توسعه و رفاه جهانی می تواند نقش بسیار مهمی ایفا کند. شاخص های بسیاری برای ارزیابی توسعه و رفاه کشورها ارائه شده است که از مهم ترین آنها می توان به

* پست الکترونیک نویسنده مسئول: naghashasadi@ut.ac.ir

۱. دانشکده فنی فومن، دانشکده گان فنی دانشگاه تهران، تهران، ایران

استناد به این مقاله:

علی نقاش اسدی، امین مستوری نژاد و رومینا عقیلی، "تحلیل داده آماری از شاخص رفاه لگاتوم و ارائه مدل های پیش بینی از فرآیند داده کاوی"، مدل سازی در مهندسی،

doi: 10.22075/jme.2025.34932.2714، ۸۴-۶۷: (۱۴۰۵) ۸۵ ۲۴

(۱) ایمنی امنیت^۴: این شاخص شامل معیارهای مختلفی است که به بررسی میزان امنیت فردی و اجتماعی افراد در یک کشور می‌پردازد. این معیارها شامل بررسی میزان وقوع جرم‌های خشونت‌آمیز و غیرخشونت‌آمیز در یک جامعه، سطح درگیری‌های داخلی و خارجی، تعداد و شدت حوادث تروریستی، میزان خشونت‌های سیاسی و اجتماعی، کارایی نیروهای پلیس و سایر نیروهای انتظامی، دسترسی به خدمات اورژانس و واکنش سریع به حوادث، قوانین و مقررات مرتبط با ایمنی عمومی و میزان اجرای آنها، آمادگی کشور در مواجهه با بلایای طبیعی، مدیریت بحران و واکنش سریع به حوادث غیرطبیعی است.

(۲) آزادی مردم^۵: این شاخص به بررسی جوانب مختلفی از زندگی افراد می‌پردازد که مرتبط با حقوق و آزادی‌های اساسی آنها است. از معیارهای اصلی که در این شاخص بررسی می‌شوند می‌توان به توانایی افراد برای بیان نظرات و دیدگاه‌های خود بدون ترس از مجازات، دسترسی آزاد به اطلاعات و منابع خبری، عدم سانسور رسانه‌ها و اینترنت، حق تشکیل و عضویت در سازمان‌ها، انجمن‌ها و اتحادیه‌ها، امکان برگزاری تجمعات و اعتراضات مسالمت‌آمیز بدون محدودیت، حق رای و شرکت در انتخابات آزاد و منصفانه، دسترسی برابر به فرآیندهای سیاسی و امکان کاندیدا شدن در انتخابات، حق انتخاب و پیروی از هر دین یا عقیده‌ای، عدم تبعیض بر اساس مذهب یا باورهای شخصی، امکان جابجایی آزادانه در داخل کشور و سفر به خارج از کشور، حق انتخاب محل اقامت و تغییر آن، حفاظت از حریم خصوصی و اطلاعات شخصی، امنیت در برابر تجسس و تجاوز به حریم شخصی، برابری جنسیتی و عدم تبعیض بر اساس جنسیت، حقوق و آزادی‌های اقلیت‌های قومی، نژادی، مذهبی و جنسی و عدالت و حقوق قانونی اشاره کرد.

(۳) حکومت‌داری^۵: این شاخص شامل معیارهای مختلفی است که به بررسی کارایی، شفافیت، پاسخگویی، و کیفیت کلی نهادهای دولتی و نحوه عملکرد دولت‌ها می‌پردازد. حکمرانی خوب می‌تواند موجب جذب سرمایه‌گذاری و رشد اقتصادی، بهبود خدمات عمومی و افزایش رضایت عمومی و همچنین موجب ثبات سیاسی و کاهش ناآرامی‌های اجتماعی شود.

شاخص رفاه لگاتوم^۲ اشاره کرد [۳]. شاخص رفاه لگاتوم هر ساله کشورها را بر اساس ۱۲ عامل رتبه‌بندی می‌کند. داده‌های ارائه‌شده توسط این موسسه، نقاط قوت و ضعف کشورها را مشخص کرده و می‌تواند راهنمای مناسبی برای بالا بردن توسعه و رفاه کشورها باشد. با وجود اطلاعات ارزشمندی که این موسسه ارائه می‌کند، همچنان اطلاعات بسیار ارزشمند دیگری نیز در میان داده‌های آن وجود دارند که با انجام فرآیند داده‌کاوی می‌توان آنها را استخراج کرد. برای مثال، باید بدانیم که از میان شاخص‌های معرفی‌شده، کدام شاخص اهمیت بالاتری دارد؟ تغییر هر شاخص به چه میزان بر شاخص‌های دیگر تاثیر دارد؟ تاثیرگذارترین شاخص مؤثر بر شاخص‌های دیگر کدام است؟ کشورها در کدام شاخص‌ها ضعف دارند و با توجه به اهمیت متفاوت شاخص‌ها، بهبود کدام شاخص را باید در اولویت قرار دهند؟ پاسخ به این سوالات نیازمند تحلیل داده آماری و انجام فرآیند داده‌کاوی به منظور استخراج مدل‌های پیش‌بینی است. پس از استخراج مدل‌های پیش‌بینی می‌توان به تمامی پرسش‌های فوق پاسخ داد و در نتیجه، سیاست‌گذاران و مسئولین کشورها می‌توانند از آن برای تصمیم‌سازی، برنامه‌ریزی و اولویت‌بندی کارها استفاده کنند.

در بخش ۲، در مورد مفاهیم پایه استفاده‌شده در مقاله شامل شاخص رفاه لگاتوم، روش‌های داده‌کاوی و معیارهای ارزیابی، توضیحاتی ارائه می‌شود. سپس در بخش ۳، کارهای مرتبط معرفی شده و با کارهای انجام‌شده در این مقاله مقایسه می‌شوند. در ادامه و در بخش ۴، تحلیل داده آماری از داده‌های جمع‌آوری‌شده ارائه شده و سپس نقشه حرارتی از تاثیر متقابل شاخص‌ها بر یکدیگر و همچنین مدل‌های پیش‌بینی برای هر شاخص ارائه می‌شود. علاوه بر این، نتایج بدست آمده از مدل‌ها با نتایج واقعی بر اساس معیارهای ارزیابی مختلف مقایسه می‌شوند. در انتها و در بخش ۵، نتیجه‌گیری و کارهای آینده ارائه می‌شود.

۲- مفاهیم پایه

۲-۱- عوامل شاخص رفاه لگاتوم

عوامل معرفی‌شده برای شاخص رفاه لگاتوم شامل ۱۲ مورد است [۳] [۴]:

⁴ Personal Freedom
⁵ Governance

² Legatum Prosperity Index
³ Safety Security

شامل معیارهای مختلفی از جمله دسترسی به بازارها، زیرساخت‌های حمل‌ونقل، زیرساخت‌های فناوری اطلاعات و ارتباطات، زیرساخت‌های انرژی و زیرساخت‌های آب و فاضلاب است.

۸) کیفیت اقتصادی^{۱۰}: این شاخص به مجموعه‌ای از معیارها و فاکتورهای اقتصادی اشاره دارد که نشان‌دهنده توانایی یک اقتصاد در ایجاد رشد پایدار، اشتغال، و بهبود سطح زندگی مردم است. از معیارهای اصلی این شاخص می‌توان به نرخ رشد تولید ناخالص داخلی، سرانه تولید ناخالص داخلی، میزان وابستگی به صنایع و بخش‌های مختلف اقتصادی، نرخ تورم، بدهی عمومی، ثبات مالی و بانکی، بهره‌وری نیروی کار، سرمایه‌گذاری در فناوری و نوآوری، نرخ بیکاری، کیفیت اشتغال، میزان درآمد متوسط افراد، سطح فقر و میزان توسعه و کارایی بازارهای مالی، کالاها و خدمات اشاره کرد.

۹) شرایط زندگی^{۱۱}: این شاخص شامل مجموعه‌ای از عوامل اجتماعی، اقتصادی و محیطی است که به طور مستقیم بر رفاه و آسایش افراد تأثیر می‌گذارند. این شاخص می‌تواند دیدگاه‌های جامعی برای تصمیم‌گیری‌های اجتماعی و برنامه‌ریزی‌های بلندمدت، شناسایی نقاط ضعف و قوت و اتخاذ تدابیر لازم برای بهبود کیفیت زندگی ارائه دهد و همچنین بهبود شرایط زندگی در کشورها و مناطق مختلف موجب جذب سرمایه‌گذاری داخلی و خارجی، افزایش روابط بهتر میان افراد و اعتماد و رضایت جامعه از کشور شود.

۱۰) سلامت^{۱۲}: شاخص سلامت شامل معیارهایی از جمله میانگین سال‌هایی که انتظار می‌رود یک انسان با توجه به نرخ مرگ و میر جاری زندگی کند، بیماری‌های مسری و غیرمسری، دسترسی به خدمات بهداشتی، کیفیت خدمات بهداشتی و تغذیه و سبک زندگی است.

۱۱) آموزش^{۱۳}: شاخص آموزش شامل معیارهایی از جمله دسترسی به آموزش در جامعه، کیفیت آموزش، نرخ باسوادی، نرخ فارغ‌التحصیلی و سرانه هزینه‌های آموزشی است. افزایش کیفیت آموزش دارای مزایای بسیاری است که می‌تواند به بهبود کیفیت زندگی فردی و جمعی کمک کند.

۴) سرمایه اجتماعی^۶: این شاخص به بررسی و ارزیابی میزان روابط اجتماعی و شبکه‌های اجتماعی افراد و گروه‌ها می‌پردازد که به عنوان سرمایه‌ای غیرمادی می‌تواند تأثیر قابل توجهی بر توسعه اجتماعی و اقتصادی داشته باشد. سرمایه اجتماعی بالا می‌تواند موجب تسهیل همکاری‌ها و افزایش کارایی اقتصادی، روابط اجتماعی قوی و اعتماد عمومی بالا، بهبود کیفیت زندگی افراد و همچنین موجب افزایش ثبات اجتماعی و کاهش ناآرامی‌های اجتماعی شود.

۵) محیط سرمایه‌گذاری^۷: شاخص محیط سرمایه‌گذاری شامل معیارهایی است که به ارزیابی جذابیت و پایداری محیط اقتصادی و نهادی یک کشور برای جذب سرمایه‌گذاری‌های داخلی و خارجی می‌پردازد. این معیارها شامل نرخ رشد اقتصادی، نرخ تورم، میزان بدهی دولت، کیفیت زیرساخت‌های فیزیکی همچون وضعیت جاده‌ها، پل‌ها، بنادر، فرودگاه‌ها، و شبکه‌های حمل و نقل، زیرساخت‌های فناوری اطلاعات و ارتباطات، دسترسی به بانک‌ها و مؤسسات مالی، نرخ بهره و شرایط اعتباری، بازار سرمایه، اندازه و رشد بازار، تقاضای مصرف‌کننده و دسترسی به بازارهای خارجی است.

۶) شرایط کسب‌وکار^۸: این شاخص به مجموعه‌ای از عوامل اقتصادی، نهادی و زیرساختی اشاره دارد که می‌تواند بر توانایی شرکت‌ها برای تأسیس، رشد و بهره‌وری تأثیرگذار باشند. از جمله معیارهای مؤثر در این شاخص می‌توان به تعداد مراحل و زمان مورد نیاز برای ثبت و راه‌اندازی یک کسب‌وکار، هزینه‌های مربوط به شروع کسب‌وکار و ثبت شرکت، قوانین و مقررات مربوط به شروع فعالیت‌های تجاری، دسترسی به وام‌ها و تسهیلات مالی برای کسب‌وکارها، کیفیت و میزان دسترسی به نیروی کار ماهر، قوانین و مقررات کار و استخدام، هزینه‌های عملیاتی شامل هزینه‌های انرژی، اجاره، مواد اولیه و نیروی کار، مالیات‌ها و تعرفه‌های دولتی، میزان رقابت در بازار و حمایت از رقابت سالم و حمایت‌های دولتی و خصوصی از فعالیت‌های نوآورانه اشاره کرد.

۷) دسترسی به بازار و زیرساخت‌ها^۹: این شاخص برای ارزیابی ظرفیت کشورها در ایجاد محیطی مناسب برای تجارت، حمل‌ونقل، و ارتباطات اهمیت دارد. این شاخص

¹⁰ Economic Quality

¹¹ Living Conditions

¹² Health

¹³ Education

⁶ Social Capital

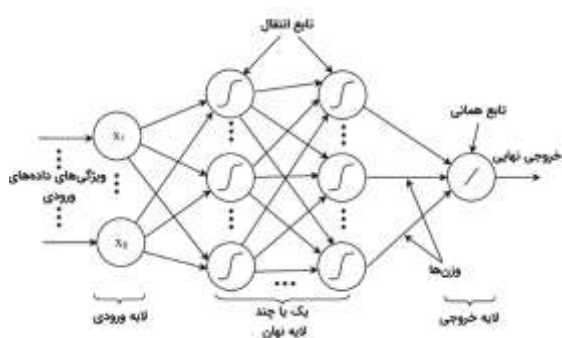
⁷ Investment Environment

⁸ Enterprise Conditions

⁹ Infrastructure & Market Access

ورودی‌های نورون‌های لایه بعد استفاده می‌شوند. با جمع مقادیر ورودی‌ها به همراه مقدار ثابت بایاس^{۲۰}، خروجی یک نورون بدست می‌آید. مطابق با رابطه (۲)، اگر نورون i دارای m ورودی (I_1 الی I_m) باشد، خروجی آن از مجموع حاصل ضرب مقدار هر ورودی در وزن آن (W_i) بدست می‌آید که در انتها با مقدار ثابت بایاس، جمع می‌شود. در نهایت نورون نهایی در لایه خروجی، خروجی شبکه عصبی را ایجاد می‌کند [۷].

$$O_i = \sum_{i=1}^m w_i I_i + b \quad (2)$$



شکل ۱- مثالی از شبکه عصبی

۳-۲- معیارهای ارزیابی

برای ارزیابی مدل‌های پیش‌بینی بدست آمده از فرآیند داده‌کاوی، معیارهای مختلفی وجود دارد. در این مقاله از معیارهای ضریب همبستگی^{۲۱} (CC) با رابطه (۳)، میانگین خطای مطلق^{۲۲} (MAE) با رابطه (۴)، و خطای جذر میانگین مربعات^{۲۳} (RMSE) با رابطه (۵) برای مقایسه مقادیر خروجی بدست آمده از مدل ($\hat{y}_i$) و مقادیر واقعی (y_i) استفاده می‌شود. در معیار ضریب همبستگی، هر چه خروجی به مقدار یک نزدیک‌تر باشد، و در معیارهای میانگین خطای مطلق و خطای جذر میانگین مربعات، هر چه خروجی به مقدار صفر نزدیک‌تر باشد به معنی فاصله کم بین نتایج واقعی و پیش‌بینی شده است.

$$CC = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (4)$$

(۱۲) محیط زیست^{۱۴}: از معیارهای این شاخص می‌توان به کیفیت هوا که خود شامل میزان غلظت آلاینده‌ها در هوا است، کیفیت آب که خود شامل میزان آلودگی منابع آب سطحی و زیرزمینی با مواد شیمیایی، فلزات سنگین، و میکروب‌ها است، حفاظت از تنوع زیستی، استفاده از منابع طبیعی، مدیریت پسماند و سطح آلودگی صوتی در مناطق مختلف اشاره کرد.

۲-۲- روش‌های داده‌کاوی

روش‌های مختلف داده‌کاوی را می‌توان به دو دسته داده‌کاوی پیش‌بینی‌کننده^{۱۵} و داده‌کاوی توصیفی^{۱۶} تقسیم‌بندی کرد [۵] [۶]. در داده‌کاوی پیش‌بینی‌کننده، با بررسی داده‌های گذشته می‌توان رویدادهای آینده را پیش‌بینی کرد. به عبارت دیگر در داده‌کاوی پیش‌بینی‌کننده، بر اساس پارامترهای ورودی، پارامتر هدف پیش‌بینی شده و با مقدار واقعی مقایسه می‌شود. ولی در داده‌کاوی توصیفی، پارامتر هدفی وجود ندارد و صرفاً نمونه‌ها بر اساس میزان شباهت‌شان به یکدیگر، دسته‌بندی می‌شوند. بنابراین با توجه به هدف مقاله، باید از داده‌کاوی پیش‌بینی‌کننده استفاده شود. از مهم‌ترین روش‌های داده‌کاوی پیش‌بینی‌کننده، رگرسیون^{۱۷} و شبکه‌های عصبی مصنوعی^{۱۸} است. مطابق با رابطه (۱)، روش رگرسیون یک روش آماری است که برای یافتن رابطه بین پارامتر خروجی (Y)، با یک یا چند پارامتر ورودی (X_i) استفاده می‌شود. به عبارت دیگر، با فرض وجود n پارامتر ورودی (X_1 الی X_n) خواهیم مقدار پارامتر خروجی Y را محاسبه کنیم. برای این منظور صرفاً باید ضرایب پارامترهای ورودی (a_i) و مقدار ثابت (e) محاسبه شوند.

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n + e \quad (1)$$

مطابق با شکل (۱)، شبکه‌های عصبی مصنوعی برای حل مسائل پیچیده در یادگیری ماشین استفاده می‌شوند. این روش از یک لایه ورودی، یک یا چند لایه پنهان، و یک لایه خروجی تشکیل شده است. هر لایه شامل تعدادی عناصر پردازشی به نام نورون^{۱۹} است که به نورون‌های لایه‌های قبلی و بعدی متصل می‌شوند. در این روش، خروجی نورون‌های لایه قبل با وزنی ضرب شده و نتایج آنها به عنوان

¹⁹ Neurons

²⁰ Bias

²¹ Correlation Coefficient

²² Mean Absolute Error

²³ Root Mean Squared Error

¹⁴ Natural Environment

¹⁵ Predictive Data Mining

¹⁶ Descriptive Data Mining

¹⁷ Regression

¹⁸ Artificial Neural Network

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (5)$$

۳- کارهای مرتبط

تحلیل داده آماری و ارائه مدل‌های پیش‌بینی در حوزه توسعه و رفاه، اقتصاد و مسائل مالی در مقالات متعددی انجام شده است. معمولاً در این مقالات، ابتدا تعدادی عوامل مؤثر یا شاخص اقتصادی و غیراقتصادی انتخاب شده و در ادامه پس از جمع‌آوری داده‌ها و ارائه تحلیل‌های آماری، مدل‌های پیش‌بینی از آنها ارائه می‌شود. برای مثال در [۷]، داده‌های شش شاخص مالی کشور ایران جمع‌آوری شده و بعد از پاک‌سازی و نرمال‌سازی، با استفاده از روش‌های رگرسیون و پرسپترون چندلایه مورد تحلیل قرار گرفته و مدل‌های پیش‌بینی از آنها استخراج شده است. در [۸]، از داده‌های برای بررسی شش حوزه مالی مختلف شامل، رتبه‌بندی اعتباری، پیش‌بینی پیش‌فرض وام، پول‌شویی، پیش‌بینی سهام و پیش‌بینی سری‌های مالی استفاده شده است. در [۹]، با در نظر گرفتن عامل زمان در معاملات مالی، یک چارچوب داده‌کاوی برای کاهش هزینه و کسب سود در بازارهای مالی ارائه شده است. در [۱۰]، برای تحلیل داده‌های سری زمانی و بهبود تحلیل بصری اقتصاد منطقه‌ای، از الگوریتم‌های یادگیری ماشین^{۲۴} استفاده شده

است. در [۱۱] با استفاده از فرآیند داده‌کاوی، رشد تولید ناخالص داخلی که معیاری برای تولید اقتصادی و تعیین سطح استاندارد زندگی مردم است، در ایالتی از کشور هند پیش‌بینی شد. در [۱۲]، به منظور اندازه‌گیری شاخص توسعه اقتصاد دیجیتال در کشور چین از فرآیند داده‌کاوی استفاده شد. در [۱۳]، از فرآیند داده‌کاوی برای پیش‌بینی تلاطم بازار به عنوان یک عامل مؤثر در تعیین ریسک سرمایه‌گذاری استفاده شده است. در [۱۴]، یک روش چهار مرحله‌ای شامل خوشه‌بندی و داده‌کاوی برای شناسایی مسیرهای حیاتی به سوی رفاه چندبعدی کشورها ارائه شده است. در [۱۵]، تفاوت‌های شاخص رفاه کشورها در مناطق جغرافیایی آنها است و از آزمون فریدمن^{۲۵} برای بررسی اهمیت تفاوت‌ها استفاده شده است. در جدول ۱، کارهای مرتبط با این مقاله مقایسه شده است. همان طور که مشاهده می‌شود، در اکثر مقالات، به بررسی شاخص‌های توسعه و رفاه پرداخته‌اند. در محدود مقالاتی که شاخص‌های توسعه و رفاه را مورد بررسی قرار داده‌اند، با وجود استفاده از فرآیندهای تحلیل داده و داده‌کاوی، مدل‌های پیش‌بینی ارائه نشده است. در برخی مقالات نیز با وجود ارائه مدل‌های پیش‌بینی مبتنی بر داده‌کاوی، شاخص‌ها و عوامل مورد بررسی آنها با شاخص‌ها و عوامل بررسی‌شده در این مقاله متفاوت بوده است.

جدول ۱- مقایسه بین کارهای مرتبط

مقاله	عوامل بررسی شده	تحلیل داده	ارائه مدل‌های پیش‌بینی
[۷]	طلا، دلار، یورو، سکه، نفت	✓	✓
[۸]	رتبه‌بندی اعتباری، پیش‌بینی پیش‌فرض وام، پول‌شویی، پیش‌بینی سهام و پیش‌بینی سری‌های مالی	✓	✓
[۹]	نرخ بهره، نرخ ارز، شاخص بازار سهام، کالا، سایر شاخص‌های اقتصادی در کره جنوبی، سایر شاخص‌های اقتصادی در ایالات متحده	✓	×
[۱۰]	تحلیل داده‌های سری زمانی و بهبود تحلیل بصری اقتصاد منطقه‌ای	✓	✓
[۱۱]	تولید ناخالص داخلی، سطح استاندارد زندگی، رشد اقتصادی، کشور هند	✓	✓
[۱۲]	رشد توسعه اقتصاد دیجیتال در چین	✓	✓
[۱۳]	نفت، طلا، تلاطم بازار بورس، تولید ناخالص داخلی، سود بانکی، حجم پول در گردش، نرخ تورم	✓	✓
[۱۴]	۱۲ شاخص توسعه و رفاه جهانی	✓	×
[۱۵]	۱۲ شاخص توسعه و رفاه جهانی	×	×
مقاله ما	۱۲ شاخص توسعه و رفاه جهانی	✓	✓

²⁵ Friedman Test

²⁴ Machine Learning

چهار کشور هم رده با کشور ایران از سال ۲۰۱۸ الی ۲۰۲۳ مطابق با شکل (۵) است. همان طور که مشاهده می‌شود، در سال‌های اخیر رتبه توسعه و رفاه در کشور ایران مشابه با کشورهای هم رده، روند افزایشی (کاهش توسعه و رفاه) داشته است.

در شکل (۶)، نحوه تغییر رتبه ۱۲ شاخص توسعه و رفاه در کشور ایران از سال ۲۰۱۸ الی ۲۰۲۳ نشان داده شده است. همان طور که مشاهده می‌شود، کشور ایران به ترتیب در شاخص‌ها سلامت، آموزش و شرایط زندگی نسبت به شاخص‌های دیگر در رتبه‌های بهتری قرار گرفته است و حتی در سال‌های اخیر در شاخص سلامت، رتبه کشور ایران روند کاهشی (بهرتر شدن وضعیت) داشته است. ولی به ترتیب در شاخص‌های آزادی مردم، شرایط کسب‌وکار و محیط زیست، نسبت به شاخص‌های دیگر در رتبه‌های بدتری قرار گرفته است.

۴- تحلیل داده آماری و ارائه مدل‌های پیش‌بینی

۴-۱- داده‌های جمع‌آوری شده و تحلیل داده آماری

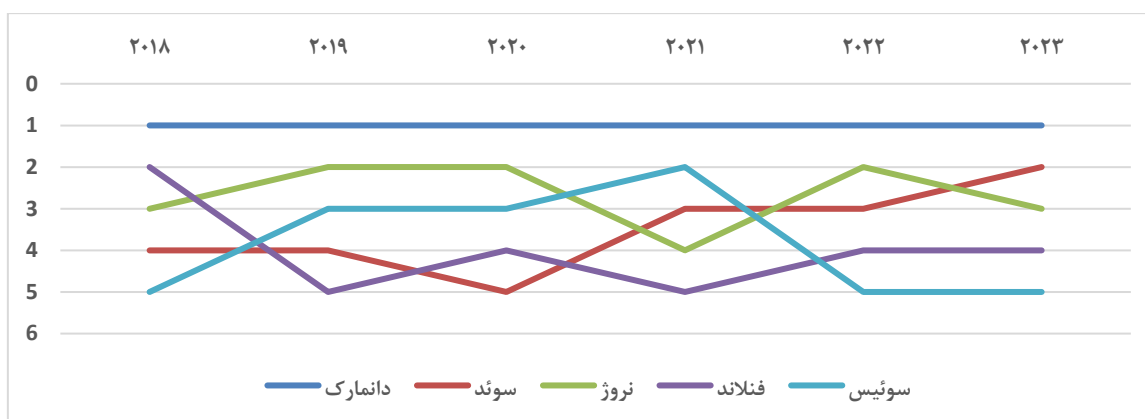
در بخش ۱.۲، ۱۲ شاخص توسعه و رفاه جهانی معرفی شدند. مطابق با این شاخص‌ها، مرجع [۳] هر ساله اقدام به نمره‌دهی و رتبه‌بندی ۱۶۷ کشور براساس این شاخص‌ها می‌کند. این مرجع، مقادیر هر شاخص را عددی بین صفر تا ۱۰۰ در نظر می‌گیرد و با محاسبه میانگین مجموع مقادیر شاخص‌ها، معدل نهایی و رتبه هر کشور را اعلام می‌کند. شکل (۲)، معدل نهایی و مقادیر شاخص‌های توسعه و رفاه هشت کشور اول را در سال ۲۰۲۳ نشان می‌دهد [۳]. شکل (۳)، معدل نهایی و مقادیر شاخص‌های توسعه و رفاه کشور ایران را که در رتبه ۱۲۶ قرار دارد را نشان می‌دهد [۳]. مطابق با اطلاعات ارائه شده از مرجع [۳]، نحوه تغییر رتبه معدل نهایی پنج کشور اول از سال ۲۰۱۸ الی ۲۰۲۳ مطابق با شکل (۴) است. همچنین نحوه تغییر رتبه معدل نهایی

SCORE	COUNTRY													
ADJUST PILLAR WEIGHTING ?			x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	▼ x1	
84.55	Denmark	+	92.59	94.09	89.45	82.56	82.42	79.64	78.79	76.81	95.77	81.07	87.48	73.94
83.67	Sweden	+	90.97	91.90	86.41	78.29	82.81	75.54	79.67	76.18	95.33	82.28	85.92	78.74
83.59	Norway	+	93.30	94.10	89.66	79.03	82.24	75.95	75.87	77.25	94.70	82.98	85.68	72.37
83.47	Finland	+	89.56	91.96	90.41	77.27	84.32	77.25	78.77	70.28	94.46	81.19	88.38	77.99
83.42	Switzerland	+	95.66	87.50	87.67	69.14	80.81	83.84	78.65	79.71	94.66	82.11	87.72	73.60
82.32	Netherlands	+	91.19	90.06	87.34	74.03	84.11	79.09	80.82	74.34	95.86	82.05	86.43	62.49
81.83	Luxembourg	+	96.32	89.20	86.31	66.60	78.91	80.72	80.03	76.93	94.56	81.59	78.79	71.98
81.02	Iceland	+	91.64	88.74	83.30	77.75	79.20	72.86	76.07	69.92	93.82	82.72	85.19	71.01

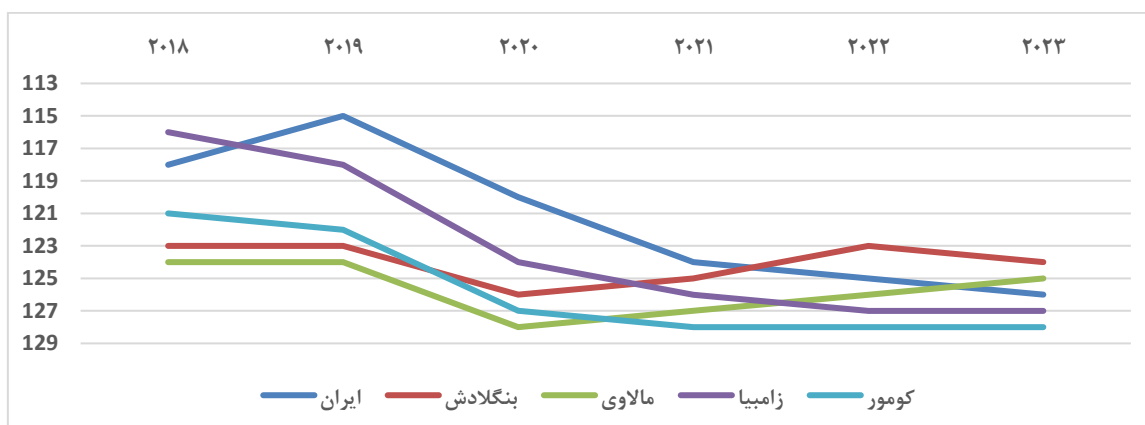
شکل ۲- معدل نهایی و مقادیر شاخص‌های توسعه و رفاه هشت کشور اول در سال ۲۰۲۳

SCORE	COUNTRY													
ADJUST PILLAR WEIGHTING ?			x1	x1	x1	x1	x1	x1	x1	x1	x1	x1	x1	
47.87	Bangladesh	+	50.62	39.87	36.25	50.95	38.18	52.44	47.76	45.09	57.09	67.39	44.19	44.58
47.83	Malawi	+	65.96	61.61	48.38	40.12	42.41	52.04	36.06	37.65	37.18	61.01	33.93	57.58
47.71	Iran	+	55.57	37.09	31.83	48.96	39.30	32.60	50.34	38.93	76.11	74.93	63.75	43.11
47.14	Zambia	+	66.05	48.29	42.03	50.63	43.20	52.92	35.98	30.93	40.82	57.20	39.04	58.60
46.93	Comoros	+	74.80	53.40	29.30	54.53	36.43	43.31	35.86	33.14	48.06	59.60	38.22	56.54

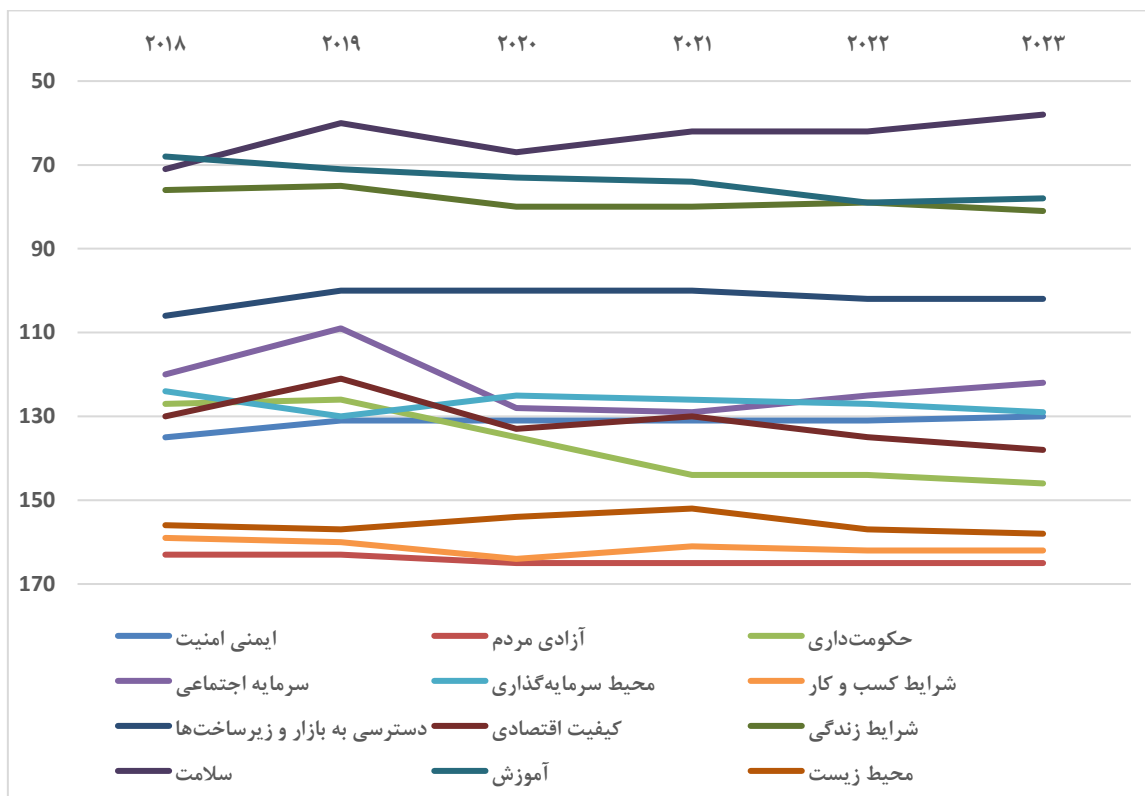
شکل ۳- معدل نهایی و مقادیر شاخص‌های توسعه و رفاه کشور ایران در سال ۲۰۲۳



شکل ۴- رتبه معدل نهایی پنج کشور اول از سال ۲۰۱۸ الی ۲۰۲۳



شکل ۵- رتبه معدل نهایی چهار کشور هم رده با کشور ایران از سال ۲۰۱۸ الی ۲۰۲۳



شکل ۶- رتبه ۱۲ شاخص توسعه و رفاه در کشور ایران از سال ۲۰۱۸ الی ۲۰۲۳

۴-۲- نقشه حرارتی

در شکل (۷) با استفاده از کد پایتون، میزان تاثیر متقابل ۱۲ شاخص توسعه و رفاه و معدل نهایی نسبت به یکدیگر با استفاده از ماتریس ضریب همبستگی پیرسون به عنوان نقشه حرارتی ارائه شده است. با مشاهده این شکل می توان دریافت کرد که شاخص محیط سرمایه گذاری با ضریب ۰.۹۵ بیشترین تاثیر را در معدل نهایی و همچنین بیشترین تاثیر را روی سایر شاخص ها دارد. همچنین شاخص محیط زیست، کمترین تاثیر را روی سایر شاخص ها و معدل نهایی دارد. سایر اطلاعات قابل برداشت در جدول ۲ ارائه شده است. به عنوان مثال از ردیف ۱ جدول ۲ می توان برداشت کرد که از میان شاخص های توسعه و رفاه، شاخص حکومت داری با ضریب ۰.۷۶، بیشترین تاثیر را بر شاخص ایمنی و امنیت داشته است.

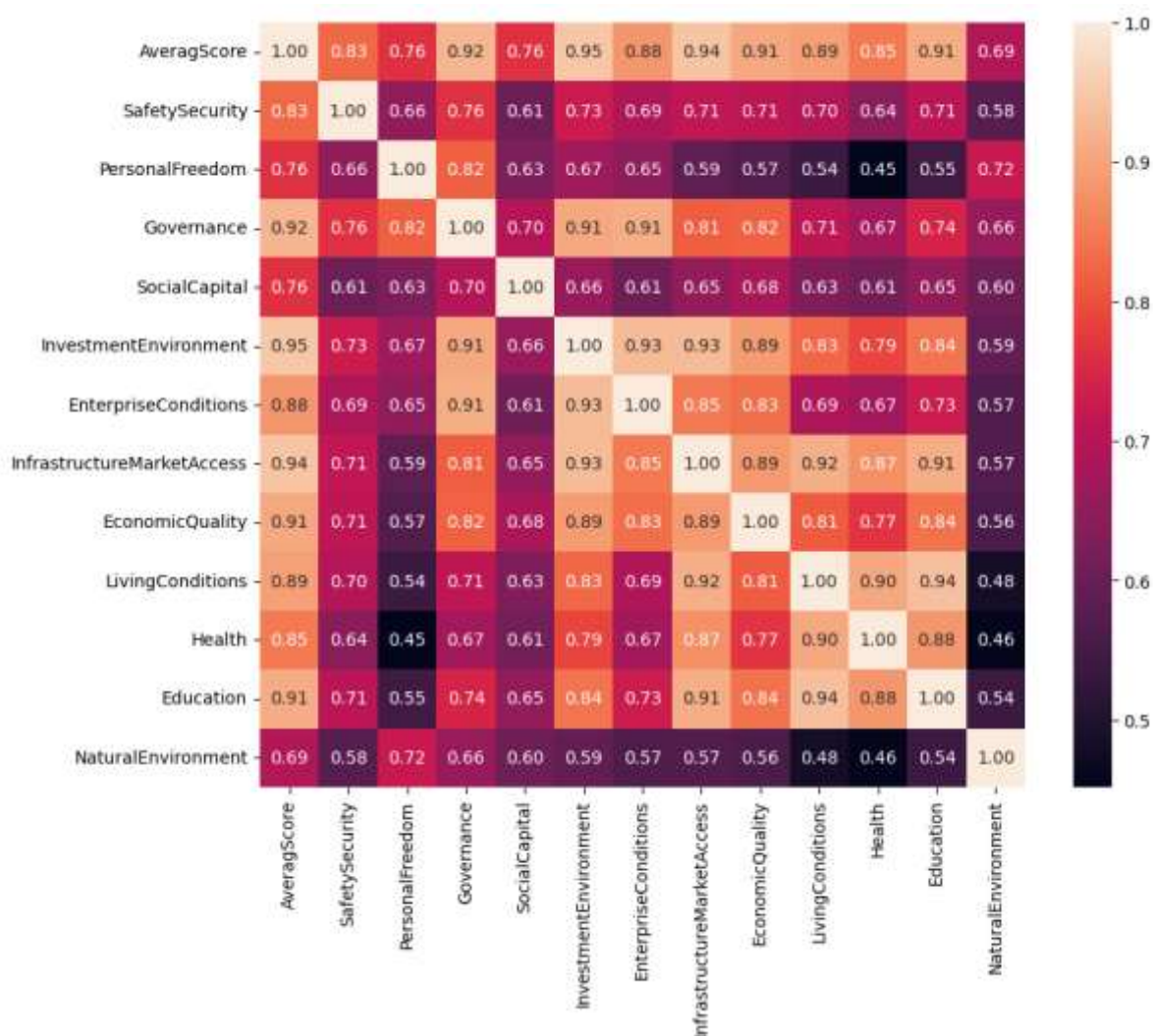
۴-۳- مدل های پیش بینی

در این بخش، با استفاده از روش های رگرسیون خطی و شبکه عصبی، مدل های پیش بینی برای هر یک از شاخص ها ارائه شده و نتایج حاصل از آنها با نتایج واقعی مقایسه و ارزیابی می شود. به عبارت دیگر، با استفاده از مدل های پیش بینی قصد داریم، مقدار هر شاخص را از روی شاخص های دیگر محاسبه کنیم. برای این منظور از ابزار

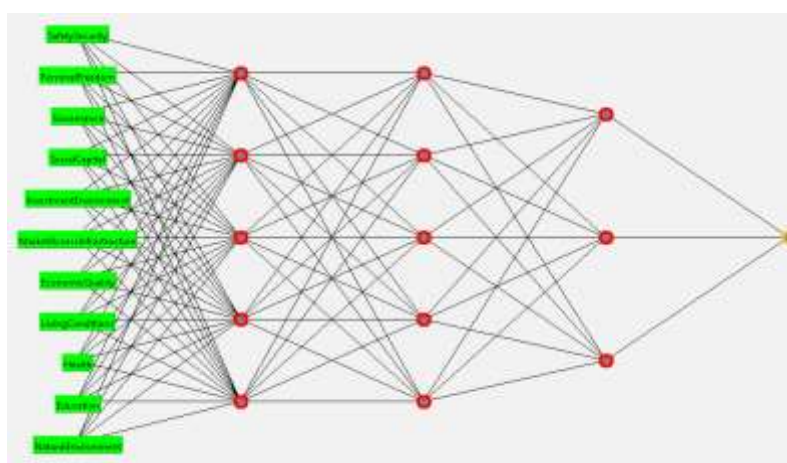
Weka استفاده شده است [۱۶]. برای استخراج مدل های پیش بینی، ۸۰ درصد از داده ها برای آموزش مدل ها، و ۲۰ درصد دیگر برای آزمون آنها مورد استفاده قرار گرفت. در روش رگرسیون، شاخصی که بیشترین ضریب را دارد، به عنوان تاثیرگذارترین شاخص برای شاخص هدف در نظر گرفته شده است. در بررسی های انجام شده برای طراحی شبکه عصبی مصنوعی، شبکه ای با پنج لایه بهترین عملکرد را ارائه کرد؛ به طوری که در لایه اول، ۱۱ شاخص به جز شاخص هدف و در لایه آخر، شاخص هدف قرار می گیرد. مطابق با شکل (۸)، در لایه های پنهان به ترتیب، ۵، ۳ و ۵ نورون قرار می گیرند. در ادامه، مدل های پیش بینی و نتایج ارزیابی هر شاخص ارائه شده است. همان طور که مشاهده می شود، بر اساس معیار CC، دقت مدل های پیش بینی ارائه شده در پایین ترین سطح، ۷۵ درصد و در بالاتری سطح، ۹۸ درصد بوده است. برای مثال مطابق با جدول ۶، مدل های رگرسیون و شبکه عصبی ارائه شده برای شاخص سرمایه اجتماعی، بر اساس معیار CC، به ترتیب ۷۵ و ۸۳ درصد دقت داشتند. همچنین مطابق با جدول ۸، مدل های رگرسیون و شبکه عصبی ارائه شده برای شاخص شرایط کسب و کار، بر اساس معیار CC، به ترتیب ۹۵ و ۹۶ درصد دقت داشتند.

جدول ۲- تاثیرگذارترین شاخص های توسعه و رفاه برای هر شاخص

ردیف	نام شاخص	تاثیرگذارترین شاخص (به جز معدل نهایی)	میزان تاثیر (ضریب همبستگی پیرسون)
۱	ایمنی امنیت	حکومت داری	۰.۷۶
۲	آزادی مردم	حکومت داری	۰.۸۲
۳	حکومت داری	محیط سرمایه گذاری شرایط کسب و کار	۰.۹۱
۴	سرمایه اجتماعی	حکومت داری	۰.۷۰
۵	محیط سرمایه گذاری	شرایط کسب و کار دسترسی به بازار و زیرساخت ها	۰.۹۳
۶	شرایط کسب و کار	محیط سرمایه گذاری	۰.۹۳
۷	دسترسی به بازار و زیرساخت ها	محیط سرمایه گذاری	۰.۹۳
۸	کیفیت اقتصادی	محیط سرمایه گذاری دسترسی به بازار و زیرساخت ها	۰.۸۹
۹	شرایط زندگی	آموزش	۰.۹۴
۱۰	سلامت	شرایط زندگی	۰.۹۰
۱۱	آموزش	شرایط زندگی	۰.۹۴
۱۲	محیط زیست	آزادی مردم	۰.۷۲



شکل ۷- نمودار نقشه حرارتی ۱۲ شاخص توسعه و رفاه و معدل نهایی



شکل ۸- نمودار مدل پیش‌بینی شبکه عصبی مصنوعی برای تمامی شاخص‌ها

که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها بالا است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان

□ شاخص ایمنی امنیت

در جدول ۳، اطلاعات آماری از شاخص ایمنی امنیت، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور

طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص ایمنی امنیت، حکومت‌داری بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص آزادی مردم

در جدول ۴، اطلاعات آماری از شاخص آزادی مردم، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان‌طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص آزادی مردم، حکومت‌داری بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص حکومت‌داری

در جدول ۵، اطلاعات آماری از شاخص حکومت‌داری، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان‌طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص حکومت‌داری، محیط سرمایه‌گذاری و شرایط کسب‌وکار بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص سرمایه اجتماعی

در جدول ۶، اطلاعات آماری از شاخص سرمایه اجتماعی، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. در نقشه حرارتی پیش‌بینی شد که تاثیرگذارترین شاخص در تعیین شاخص سرمایه اجتماعی، حکومت‌داری است؛ ولی در مدل رگرسیون خطی، کیفیت اقتصادی به عنوان تاثیرگذارترین شاخص اعلام شد.

□ شاخص محیط سرمایه‌گذاری

در جدول ۷، اطلاعات آماری از شاخص محیط سرمایه‌گذاری، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان‌طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص محیط سرمایه‌گذاری، شرایط کسب‌وکار و دسترسی به بازار و زیرساخت‌ها بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص شرایط کسب‌وکار

در جدول ۸، اطلاعات آماری از شاخص شرایط کسب‌وکار، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان‌طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص شرایط کسب‌وکار، محیط سرمایه‌گذاری بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص دسترسی به بازار و زیرساخت‌ها

در جدول ۹، اطلاعات آماری از شاخص دسترسی به بازار و زیرساخت‌ها، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان‌طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بین کشورها تقریباً یکنواخت است. به عبارت دیگر، تعداد کشورهایی که در این شاخص، ضعیف، متوسط یا قوی هستند، یکسان است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان‌طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص دسترسی به بازار و زیرساخت‌ها، محیط سرمایه‌گذاری بود که در مدل رگرسیون خطی نیز،

همین نتیجه بدست آمد.

□ شاخص کیفیت اقتصادی

در جدول ۱۰، اطلاعات آماری از شاخص کیفیت اقتصادی، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها پایین است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص کیفیت اقتصادی، محیط سرمایه‌گذاری و دسترسی به بازار و زیرساخت‌ها بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص شرایط زندگی

در جدول ۱۱، اطلاعات آماری از شاخص شرایط زندگی، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها بالا است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. در نقشه حرارتی پیش‌بینی شد که تاثیرگذارترین شاخص در تعیین شاخص شرایط زندگی، آموزش است؛ ولی در مدل رگرسیون خطی، دسترسی به بازار و زیرساخت‌ها به عنوان تاثیرگذارترین شاخص اعلام شد.

□ شاخص سلامت

در جدول ۱۲، اطلاعات آماری از شاخص سلامت، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این

شاخص در بیشتر کشورها بالا است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص سلامت، شرایط زندگی بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

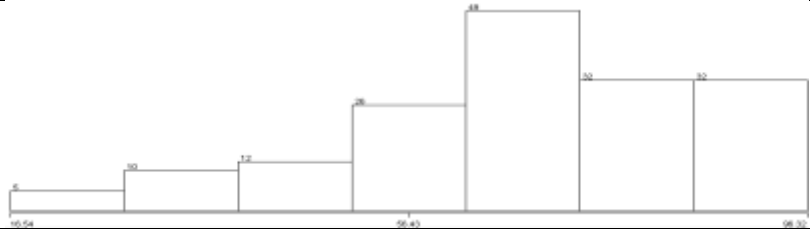
□ شاخص آموزش

در جدول ۱۳، اطلاعات آماری از شاخص آموزش، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها بالا است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. همان طور که در نقشه حرارتی پیش‌بینی شد، تاثیرگذارترین شاخص در تعیین شاخص آموزش، شرایط زندگی بود که در مدل رگرسیون خطی نیز، همین نتیجه بدست آمد.

□ شاخص محیط زیست

در جدول ۱۴، اطلاعات آماری از شاخص محیط زیست، مدل‌های پیش‌بینی رگرسیون خطی و شبکه عصبی از آن، و نتایج معیارهای ارزیابی مدل‌ها ارائه شده است. همان طور که از نمودار فراوانی این شاخص مشاهده می‌شود، مقدار این شاخص در بیشتر کشورها متوسط است. همچنین مدل‌های رگرسیون خطی و شبکه عصبی این شاخص از نظر معیارهای ارزیابی، نتایج قابل قبولی ارائه می‌کنند. در نقشه حرارتی پیش‌بینی شد که تاثیرگذارترین شاخص در تعیین شاخص محیط زیست، آزادی مردم است؛ ولی در مدل رگرسیون خطی، دسترسی به بازار و زیرساخت‌ها به عنوان تاثیرگذارترین شاخص اعلام شد.

جدول ۳- اطلاعات آماری و مدل‌های پیش‌بینی شاخص ایمنی امنیت

کمترین مقدار	۱۶.۵۴	بیشترین مقدار	۹۶.۳۲
نمودار فراوانی			
رابطه رگرسیون	SafetySecurity = 0.5443 * Governance - 0.1686 * InvestmentEnvironment - 0.1733 * InfrastructureMarketAccess + 0.1797 * EconomicQuality + 0.3645 * LivingConditions + 0.2393 * NaturalEnvironment + 10.1439		
تاثیرگذارترین شاخص از نظر مدل رگرسیون			
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۱۱.۲۷۵۵	۸.۸۵۸۹	۰.۷۵۰۹
شبکه عصبی	۹.۵۱۸۶	۷.۱۹۱۷	۰.۸۴۴۸

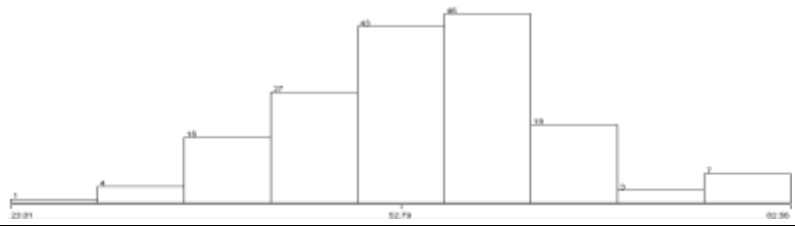
جدول ۴- اطلاعات آماری و مدل‌های پیش‌بینی شاخص آزادی مردم

۹۴.۱	بیشترین مقدار		۱۶.۱۶	کمترین مقدار
				نمودار فراوانی
PersonalFreedom = 1.2043 * Governance + 0.166 * SocialCapital - 0.4779 * EnterpriseConditions - 0.3911 * EconomicQuality + 0.4351 * LivingConditions - 0.5041 * Health - 0.1837 * Education + 0.6652 * NaturalEnvironment + 11.104				رابطه رگرسیون
حکومت‌داری		تاثیرگذارترین شاخص از نظر مدل رگرسیون		
CC	MAE	RMSE	معیارهای ارزیابی	
۰.۸۸۳۷	۸.۳۴۴	۱۰.۰۱۰۵	رگرسیون	
۰.۹۵۷۳	۴.۴۹۱۶	۵.۷۶۲۲	شبکه عصبی	

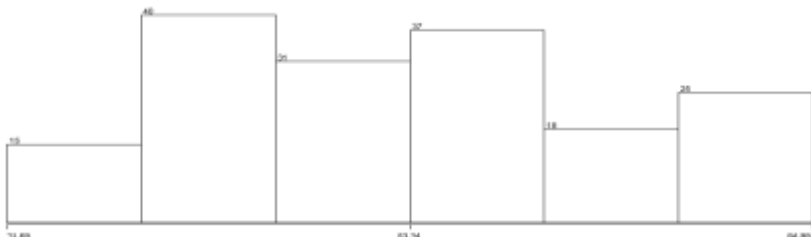
جدول ۵- اطلاعات آماری و مدل‌های پیش‌بینی شاخص حکومت‌داری

کمترین مقدار	۱۳.۰۹	بیشترین مقدار	۹۰.۴۱																
نمودار فراوانی	<table border="1"><caption>Frequency Distribution Data</caption><thead><tr><th>Bin Range</th><th>Frequency</th></tr></thead><tbody><tr><td>90.41 - 92.41</td><td>7</td></tr><tr><td>92.41 - 94.41</td><td>23</td></tr><tr><td>94.41 - 96.41</td><td>48</td></tr><tr><td>96.41 - 98.41</td><td>41</td></tr><tr><td>98.41 - 100.41</td><td>17</td></tr><tr><td>100.41 - 102.41</td><td>15</td></tr><tr><td>102.41 - 104.41</td><td>17</td></tr></tbody></table>			Bin Range	Frequency	90.41 - 92.41	7	92.41 - 94.41	23	94.41 - 96.41	48	96.41 - 98.41	41	98.41 - 100.41	17	100.41 - 102.41	15	102.41 - 104.41	17
Bin Range	Frequency																		
90.41 - 92.41	7																		
92.41 - 94.41	23																		
94.41 - 96.41	48																		
96.41 - 98.41	41																		
98.41 - 100.41	17																		
100.41 - 102.41	15																		
102.41 - 104.41	17																		
رابطه رگرسیون	$\text{Governance} = 0.0765 * \text{SafetySecurity} + 0.3002 * \text{PersonalFreedom} + 0.3015 * \text{InvestmentEnvironment} + 0.5769 * \text{EnterpriseConditions} - 0.2277 * \text{InfrastructureMarketAccess} + 0.1216 * \text{EconomicQuality} - 0.0721 * \text{LivingConditions} + 0.1225 * \text{Health} + 0.0713 * \text{Education} - 20.741$																		
تاثیرگذارترین شاخص از نظر مدل رگرسیون		شرایط کسب و کار																	
معیارهای ارزیابی	RMSE	MAE	CC																
رگرسیون	۴.۳۴۰۷	۳.۳۱۹۶	۰.۹۷۱۱																
شبکه عصبی	۴.۲۲۹۵	۳.۳۳۶۹	۰.۹۷۲۱																

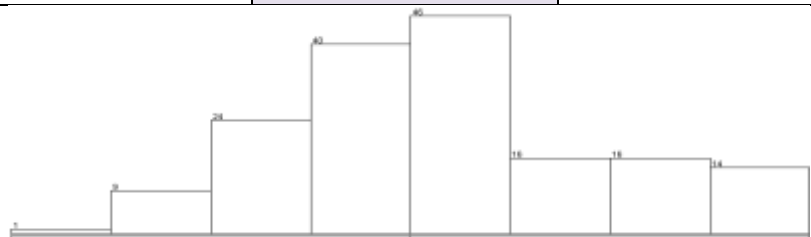
جدول ۶- اطلاعات آماری و مدل‌های پیش‌بینی شاخص سرمایه اجتماعی

کمترین مقدار	۲۳.۰۱	بیشترین مقدار	۸۲.۵۶
نمودار فراوانی			
رابطه رگرسیون	$\text{SocialCapital} = 0.0852 * \text{PersonalFreedom} + 0.1893 * \text{Governance} - 0.1026 * \text{InvestmentEnvironment} - 0.1505 * \text{EnterpriseConditions} + 0.2788 * \text{EconomicQuality} + 0.2065 * \text{Health} + 0.1938 * \text{NaturalEnvironment} + 14.3788$		
تأثیرگذارترین شاخص از نظر مدل رگرسیون	کیفیت اقتصادی		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۶.۹۳۱۹	۵.۱۴۹۲	۰.۷۵۶۹
شبکه عصبی	۶.۰۱۹۶	۵.۰۵۴۵	۰.۸۳۱۷

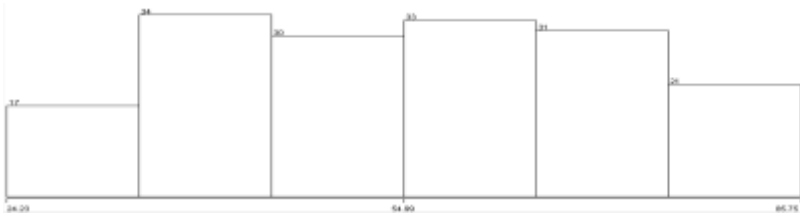
جدول ۷- اطلاعات آماری و مدل‌های پیش‌بینی شاخص محیط سرمایه‌گذاری

کمترین مقدار	۲۱.۶۹	بیشترین مقدار	۸۴.۹۹
نمودار فراوانی			
رابطه رگرسیون	$\text{InvestmentEnvironment} = 0.2096 * \text{Governance} + 0.3938 * \text{EnterpriseConditions} + 0.4387 * \text{InfrastructureMarketAccess} + 0.1079 * \text{EconomicQuality} - 8.8112$		
تأثیرگذارترین شاخص از نظر مدل رگرسیون	دسترسی به بازار و زیرساخت‌ها		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۳.۵۸۱۹	۲.۹۲۲۱	۰.۹۷۹۸
شبکه عصبی	۴.۰۲۳۹	۳.۲۸۸۹	۰.۹۷۶۸

جدول ۸- اطلاعات آماری و مدل‌های پیش‌بینی شاخص شرایط کسب‌وکار

کمترین مقدار	۲۰.۵	بیشترین مقدار	۸۳.۸۴
نمودار فراوانی			
رابطه رگرسیون	$\text{EnterpriseConditions} = -0.0905 * \text{PersonalFreedom} + 0.3871 * \text{Governance} + 0.4006 * \text{InvestmentEnvironment} + 0.2593 * \text{InfrastructureMarketAccess} - 0.2197 * \text{LivingConditions} + 20.2643$		
تأثیرگذارترین شاخص از نظر مدل رگرسیون	محیط سرمایه‌گذاری		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۴.۱۳۷۱	۳.۳۲۵۸	۰.۹۵۲۳
شبکه عصبی	۳.۵۳۶۸	۲.۷۶۲۸	۰.۹۶۸۲

جدول ۹- اطلاعات آماری و مدل های پیش بینی شاخص دسترسی به بازار و زیرساخت ها

۸۵.۷۵	بیشترین مقدار		۲۴.۲۳	کمترین مقدار
				نمودار فراوانی
InfrastructureMarketAccess = -0.1731 * Governance + 0.3315 * InvestmentEnvironment + 0.2816 * EnterpriseConditions + 0.1025 * EconomicQuality + 0.2576 * LivingConditions + 0.1398 * Health + 0.0855 * Education + 0.0988 * NaturalEnvironment - 12.7081				رابطه رگرسیون
محیط سرمایه گذاری		تاثیرگذارترین شاخص از نظر مدل رگرسیون		
CC	MAE	RMSE	معیارهای ارزیابی	
۰.۹۸۰۸	۲.۷۹۷۱	۳.۶۰۱۶	رگرسیون	
۰.۹۷۹۶	۳.۴۱۷	۴.۲۳۲۱	شبکه عصبی	

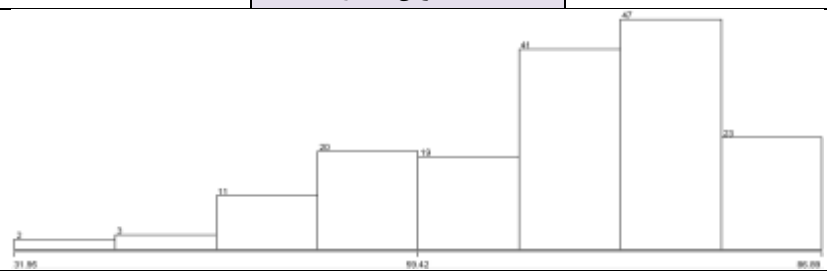
جدول ۱۰- اطلاعات آماری و مدل های پیش بینی شاخص کیفیت اقتصادی

کمترین مقدار	۲۴.۴۶	بیشترین مقدار		۸۰.۱
نمودار فراوانی				
		EconomicQuality = -0.1116 * PersonalFreedom + 0.187 * Governance + 0.1785 * SocialCapital + 0.2451 * InvestmentEnvironment + 0.2433 * InfrastructureMarketAccess + 0.093 * Education + 6.8864		
رابطه رگرسیون		تاثیرگذارترین شاخص از نظر مدل رگرسیون محیط سرمایه گذاری		
معیارهای ارزیابی	RMSE	MAE	CC	
رگرسیون	۵.۹۳۹۲	۴.۶۶۱	۰.۹۲۱۴	
شبکه عصبی	۵.۰۳۹۴	۴.۱۶۴۸	۰.۹۳۱۹	

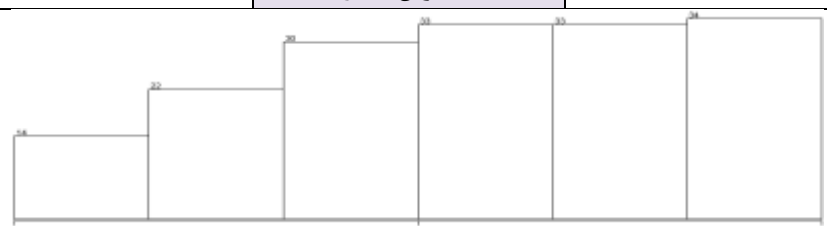
جدول ۱۱- اطلاعات آماری و مدل های پیش بینی شاخص شرایط زندگی

کمترین مقدار	۱۹.۲۱	بیشترین مقدار	۹۵.۸۶
نمودار فراوانی			
رابطه رگرسیون	LivingConditions = 0.0711 * SafetySecurity + 0.1381 * PersonalFreedom - 0.095 * Governance + 0.1363 * InvestmentEnvironment - 0.3756 * EnterpriseConditions + 0.5524 * InfrastructureMarketAccess + 0.4081 * Health + 0.4142 * Education - 0.2339 * NaturalEnvironment + 5.7026		
تاثیرگذارترین شاخص از نظر مدل رگرسیون دسترسی به بازار و زیرساخت‌ها			
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۵.۷۳۸۲	۴.۳۰۰۹	۰.۹۶۶
شبکه عصبی	۵.۸۹۵۷	۴.۸۱۹۲	۰.۹۶۸۶

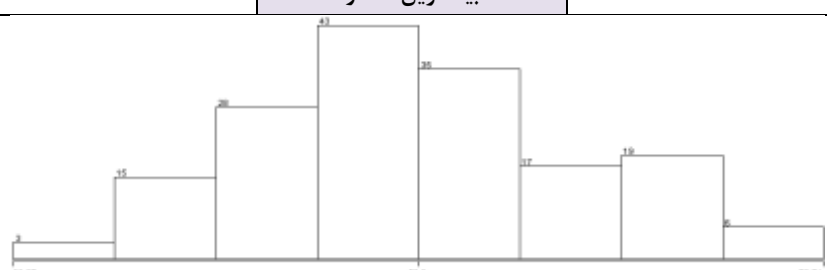
جدول ۱۲- اطلاعات آماری و مدل‌های پیش‌بینی شاخص سلامت

کمترین مقدار	۳۱.۹۵	بیشترین مقدار	۸۶.۸۹
نمودار فراوانی			
رابطه رگرسیون	$\text{Health} = -0.1062 * \text{PersonalFreedom} + 0.0774 * \text{Governance} + 0.1049 * \text{SocialCapital} + 0.1868 * \text{InfrastructureMarketAccess} - 0.0988 * \text{EconomicQuality} + 0.308 * \text{LivingConditions} + 0.0988 * \text{Education} + 32.6064$		
شرایط زندگی	تأثیرگذارترین شاخص از نظر مدل رگرسیون		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۴.۱۱۹۶	۲.۹۴۰۴	۰.۹۵۱
شبکه عصبی	۴.۴۰۲۷	۳.۴۱۶۲	۰.۹۲۹

جدول ۱۳- اطلاعات آماری و مدل‌های پیش‌بینی شاخص آموزش

کمترین مقدار	۱۶.۷۸	بیشترین مقدار	۹۱.۴۴
نمودار فراوانی			
رابطه رگرسیون	$\text{Education} = 0.1981 * \text{InfrastructureMarketAccess} + 0.1657 * \text{EconomicQuality} + 0.5538 * \text{LivingConditions} + 0.2248 * \text{Health} + 0.1223 * \text{NaturalEnvironment} - 21.6444$		
شرایط زندگی	تأثیرگذارترین شاخص از نظر مدل رگرسیون		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۶.۶۷۹۵	۴.۸۹۲۱	۰.۹۵۴۲
شبکه عصبی	۵.۱۸۵۲	۴.۰۶۰۴	۰.۹۶۵۸

جدول ۱۴- اطلاعات آماری و مدل‌های پیش‌بینی شاخص محیط زیست

کمترین مقدار	۳۳.۶۷	بیشترین مقدار	۷۸.۷۴
نمودار فراوانی			
رابطه رگرسیون	$\text{NaturalEnvironment} = 0.2563 * \text{PersonalFreedom} + 0.1588 * \text{SocialCapital} - 0.1253 * \text{EnterpriseConditions} + 0.3019 * \text{InfrastructureMarketAccess} - 0.2711 * \text{LivingConditions} + 0.1445 * \text{Education} + 33.6882$		
دسترسی به بازار و زیرساخت‌ها	تأثیرگذارترین شاخص از نظر مدل رگرسیون		
معیارهای ارزیابی	RMSE	MAE	CC
رگرسیون	۴.۹۹۹۵	۳.۹۸۵۶	۰.۸۰۵۷
شبکه عصبی	۵.۵۱۵۲	۴.۲۸۴۱	۰.۸۱۵۴

۴- نتیجه گیری

در این مقاله، داده‌های ۱۲ شاخص رفاه لگاتوم برای ۱۶۷ کشور در سال ۲۰۲۳ جمع‌آوری و تحلیل‌های آماری از آنها ارائه شد. با بررسی تحلیل‌های آماری دریافت شد که کشور ایران به ترتیب در شاخص‌ها سلامت، آموزش و شرایط زندگی نسبت به شاخص‌های دیگر در رتبه‌های بهتری قرار دارد، ولی به ترتیب در شاخص‌های آزادی مردم، شرایط کسب‌وکار و محیط زیست، نسبت به شاخص‌های دیگر در رتبه‌های بدتری قرار دارد. در ادامه با ارائه نقشه حرارتی، تاثیر متقابل شاخص‌ها نسبت به هم ارائه شد و نتیجه بدست آمده از آن به این شرح بود که به ترتیب شاخص‌های محیط سرمایه‌گذاری و محیط زیست، بیشترین و کمترین تاثیر را بر سایر شاخص‌ها دارند. بنابراین نتیجه گرفتیم که در صورتی که کشورها، شاخص محیط سرمایه‌گذاری را بهبود دهند، تاثیر بیشتر و بهتری روی سایر شاخص‌ها و در نتیجه معدل نهایی آنها خواهد داشت. در انتها با استفاده از روش‌های رگرسیون و شبکه عصبی، برای هر شاخص، مدل‌های پیش‌بینی ارائه شد و بر اساس معیارهای ارزیابی مختلف، مورد ارزیابی قرار گرفتند. مقادیر معیارهای ارزیابی نشان می‌دهند که نتایج حاصل از مدل‌های پیش‌بینی با نتایج واقعی اختلاف کمی داشته‌اند؛ بنابراین دقت مدل‌های بدست آمده اثبات شد. با استفاده از مدل‌های پیش‌بینی، شاخص‌هایی که تاثیر بیشتری بر شاخص هدف داشتند، شناسایی شدند و با استفاده از آنها دریافتیم که با بهبود یک شاخص، سایر شاخص‌ها به چه میزانی تغییر خواهند کرد.

مراجع

- [1] Schuh, Günther, Christina Reuter, Jan-Philipp Prote, Felix Brambring, and Julian Ays. "Increasing data integrity for improving decision making in production planning and control." CIRP Annals 66, no. 1 (2017): 425-428.
- [2] Schuh, Günther, Gunther Reinhart, Jan-Philipp Prote, Frederick Sauermann, Julia Horsthofer, Florian Oppolzer, and Dino Knoll. "Data mining definitions and applications for the management of production complexity." Procedia Cirp 81 (2019): 874-879.
- [3] "Rankings : Legatum Prosperity Index 2023," Legatum Institute, [Online]. Available: <https://www.prosperity.com/rankings>. [Accessed 28 07 2024].
- [4] "2023 Global Country Development & Prosperity Index," Jaggel, [Online]. Available: <https://www.kaggle.com/datasets/tarkunataalt/2023-global-country-development-and-prosperity-index>. [Accessed 30 07 2024].
- [5] Akulwar, Pooja, Sujata Pardeshi, and Arvind Kamble. "Survey on different data mining techniques for prediction." In 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on, pp. 513-519. IEEE, 2018.
- [6] Papakyriakou, Dimitrios, and Ioannis S. Barbounakis. "Data mining methods: A review." International Journal of Computer Applications 183, no. 48 (2022): 5-19.
- [7] Naghash Asadi, A, Z. Mirzaei, and S. Nikookar. "Providing predictive models of financial indicators using regression and neural networks." Computer science 8, no. 3 (1402): 26-38. (in Persian)

به عنوان کار آینده می‌توان داده‌های چند سال اخیر موسسه لگاتوم را جمع‌آوری کرده و مدل‌های پیش‌بینی مبتنی بر زمان ارائه کرد. این موضوع باعث می‌شود که مدل بدست آمده به بازه زمانی خاصی محدود نباشد و در نتیجه اعتماد به نتایج ارائه‌شده توسط آن بیشتر شود. همچنین می‌توان از روش‌های داده‌کاوی دیگر نیز برای استخراج مدل‌های پیش‌بینی استفاده کرد و میزان دقت آنها را نیز ارزیابی کرد.

تعارض منافع

نویسندگان اعلام می‌کنند که در مورد انتشار این مقاله تعارض منافع وجود ندارد.

تاییدیه اخلاقی

نویسندگان متعهد می‌شوند که مطالب این مقاله را در هیچ مجله دیگری به چاپ نرسانده‌اند.

مشارکت‌های نویسندگان

دکتر علی نقاش اسدی: ایده، پیاده‌سازی، نظارت، ارائه نتایج، بررسی نتایج، ویرایش علمی، نهایی‌سازی نسخه
امین مستوری نژاد: ایده، پیاده‌سازی، نسخه ابتدایی مقاله، ارائه نتایج

رومینا عقیلی: ایده، پیاده‌سازی، نسخه ابتدایی مقاله، ارائه نتایج

منابع مالی

برای این پژوهش هیچگونه حمایت مالی دریافت نشده است.

- [8] Taftian. A, and A. Oliaee. "The application of data mining techniques in the financial industry." in The 8th international conference on modern accounting, management and human sciences research in the third millennium, Tehran, Iran, 1400. (in Persian)
- [9] Taftian. A, and A. Teimouri. "Providing a suitable data mining framework for predicting financial markets with emphasis on specialized knowledge." in The 11th International Conference on Accounting, Management and Innovation in Business, Tehran, Iran, 1401. (in Persian)
- [10] Bai, Yang, Min Zhao, Rong Li, and Peizhu Xin. "A new data mining method for time series in visual analysis of regional economy." Information Processing & Management 59, no. 1 (2022): 102741.
- [11] Jain, Rinkesh, and Divakar Singh. "Data Mining and Analysis of Economic Data." International Journal 3, no. 8 (2013).
- [12] Liu, Hongyan. "Research on measurement of digital economy based on Data Mining." In Proceedings of the 7th International Conference on Intelligent Information Processing, pp. 1-5. 2022.
- [13] khodadi. M, A. Yaghoubnejad, and M. Khalili Araghi. "Comparison of estimation of financial market turbulence using regression model and neural network model." Financial Economics 14, no. 52 (1399): 223-240. (in Persian)
- [14] Budsaratagoon, Pornanong, and Boonlert Jitmaneeroj. "Reform priorities for prosperity of nations: The Legatum Index." Journal of Policy Modeling 43, no. 3 (2021): 657-672.
- [15] Puente-López, José Luis, Jenny Paola Lis-Gutiérrez, and Jhonathan Steven Pulido-Flórez. "The Legatum Prosperity Index and non-cooperative tax jurisdictions (2021)." Procedia Computer Science 203 (2022): 514-519.
- [16] Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. "The WEKA data mining software: an update." ACM SIGKDD explorations newsletter 11, no. 1 (2009): 10-18.