

Research Article



DOI: 10.22034/farayandno.2026.2070757.2015



This journal is an open access journal licensed under an Attribution-Non Commercial 4.0 International Licenses (CC BY-NC 4.0).

Prediction of Dew Point Pressure of Gas Condensates Using Machine Learning Methods

Hamid Sangbor¹, Mohammad Molazadeh Ahmadi¹, Alireza Baghban^{2,3*}

¹ Student, Amirkabir University of Technology, Mahshahr campus, Mahshahr, Iran

² Professor, Amirkabir University of Technology, Mahshahr campus, Mahshahr, Iran

³ National Iranian South Oilfields Company (NISOC), Ahvaz, Iran

Received: 3 Sep 2025

Accepted: 28 Dec 2025

Abstract

Accurate prediction of the dewpoint pressure in gas condensate reservoirs holds significant technical and economic importance, particularly in determining fluid characteristics, reservoir performance calculations, planning the development of gas condensate reservoirs, and the optimal design of production systems. In this context, artificial intelligence-based methods have emerged as innovative and effective tools in various engineering fields. Eight different models, including linear regression, Lasso regression, Ridge regression, Elastic Net, artificial neural networks, decision trees, k-nearest neighbors, and random forest, were utilized for predicting the dew point pressure. The results indicate that the random forest model outperformed the other models, achieving an average relative error of 3.3% and a correlation coefficient of 0.84. This study represents an innovation by applying artificial intelligence methods to predict dew point pressure, demonstrating that with precise analysis of gas condensate data and the use of advanced models, higher accuracy and efficiency in industrial predictions can be achieved.

Keyword: Gas Condensates, Data Mining, Dew Point Pressure, Artificial Intelligence Models

* alireza_baghban@alumni.ut.ac.ir

Please Cite This Article Using:

Sangbor, H., Molazadeh Ahmadi, M., Baghban, A., "Prediction of Dew Point Pressure of Gas Condensates Using Machine Learning Methods", Journal of Farayandno – Vol. 20 – No. 92, pp. 93-107, In Persian, (2026).



DOI: 10.22034/farayandno.2026.2070757.2015



This journal is an open access journal licensed under an Attribution-Non Commercial 4.0 International Licenses (CC BY-NC 4.0).

پیش‌بینی فشار نقطه شب‌نم میعانات گازی با استفاده از روش‌های یادگیری ماشین

حمید سنگ بر^{۱*}، محمد ملازاده احمدی^۲، علیرضا باغبان^۳

^۱ دانشجوی، دانشگاه صنعتی امیر کبیر واحد ماهشهر، ماهشهر، ایران

^۲ استاد، دانشگاه صنعتی امیر کبیر واحد ماهشهر، ماهشهر، ایران

^۳ شرکت ملی مناطق نفتخیز جنوب، اهواز، ایران

دریافت: ۱۴۰۴/۰۶/۱۲ پذیرش: ۱۴۰۴/۱۰/۰۷

چکیده

پیش‌بینی دقیق و به‌موقع فشار نقطه شب‌نم در مخازن گاز میعانی از اهمیت فنی و اقتصادی بالایی برخوردار است و به‌ویژه در تعیین ویژگی‌های سیال، محاسبات عملکرد مخزن، برنامه‌ریزی توسعه مخازن و طراحی بهینه سیستم‌های تولید نقش کلیدی ایفا می‌کند. از هشت مدل متفاوت، شامل رگرسیون خطی، رگرسیون لاسو، رگرسیون ریج، الاستیک نت، شبکه‌های عصبی مصنوعی، درخت تصمیم‌گیری، ترکیب نزدیک‌ترین همسایگی و جنگل تصادفی، برای پیش‌بینی فشار نقطه شب‌نم استفاده گردید. نتایج نشان می‌دهد که مدل جنگل تصادفی با میانگین درصد خطای نسبی ۳/۳٪ و ضریب همبستگی ۰/۸۴ بهترین عملکرد را در مقایسه با سایر مدل‌ها داشته و به‌طور موفقیت‌آمیزی از تمامی مدل‌های دیگر پیشی گرفته است. این مطالعه با استفاده از روش‌های هوش مصنوعی، نوآوری در پیش‌بینی فشار نقطه شب‌نم را نشان می‌دهد و تأکید می‌کند که تحلیل دقیق داده‌های میعانات گازی و مدل‌های پیشرفته، دقت و کارایی بالاتری در پیش‌بینی‌های صنعتی فراهم می‌آورد.

کلمات کلیدی: میعانات گازی، داده کاوی، فشار نقطه شب‌نم، مدل‌های هوش مصنوعی

* alireza_baghban@alumni.ut.ac.ir

۱- مقدمه

خصوصیات دقیق سیال مخزن، به ویژه داده‌های فشار-حجم-دما، نقش مهمی در محاسبات مهندسی مخزن مانند برآورد ذخایر و برنامه‌ریزی برای سناریوهای بهبود بازیابی نفت ایفا می‌کند. مخازن گاز میعانی دارای دو فشار نقطه شبنم هستند که فشار پایین‌تر معمولاً زیر فشار اتمسفر است و معمولاً مورد توجه قرار نمی‌گیرد. تعیین فشار نقطه شبنم بالایی، که به عنوان نقطه شبنم رتروگراد نیز شناخته می‌شود، مورد توجه است و نقش حیاتی در مدیریت مخزن دارد. نقطه شبنم رتروگراد زمانی اتفاق می‌افتد که فشار کاهش می‌یابد [۱]. غلظت کم هیدروکربن‌های سنگین‌تر در مخازن گاز میعانی در مقایسه با مخازن نفتی، نقطه بحرانی را به سمت چپ-پایین نمودار فشار-دما منتقل کرده و منجر به یک نمودار فشار-دما نسبتاً کوچک‌تر می‌شود. در مخازن گاز میعانی، دمای بحرانی و کریکوندنترم به ترتیب پایین‌تر و بالاتر از دمای مخزن هستند. در ابتدای تولید، فشار مخزن بالاتر از فشار نقطه شبنم است، بنابراین مخزن رفتار تک فازی دارد. در این شرایط، میعانات ارزشمند می‌توانند در سطح جدا شوند. با گذشت زمان، فشار مخزن کاهش می‌یابد و منطقه دو فازی حاصل می‌شود. در این شرایط، مقدار قابل توجهی گاز در تعادل با مقدار ناچیزی مایع قرار دارد. با ادامه کاهش فشار، مایعات میعانی به صورت مایع آزاد در مخزن تشکیل می‌شوند؛ اما معمولاً نفوذپذیری مؤثری برای تولید مایع وجود ندارد [۲]. در مخازن گاز میعانی، نسبت گاز به مایع در محدوده ۳/۲ تا ۱۵۰ میلیون فوت مکعب در بشکه استاندارد است [۲]. در چنین مخازنی، بهره‌وری اغلب با کاهش فشار نزدیک به چاه زیر فشار نقطه شبنم کاهش می‌یابد. انسداد جزئی میعانات، انتقال گاز در اطراف چاه را کاهش داده و نفوذپذیری مؤثر گاز را کاهش می‌دهد. در نتیجه، گاز تولید شده دارای کسری‌های کم‌ارزش‌تری است زیرا قبلاً در مخزن و نزدیک چاه جدا شده‌اند. بنابراین، پیش‌بینی دقیق و به‌موقع فشار نقطه شبنم برای مشخص‌سازی سیال، محاسبات عملکرد مخزن، برنامه‌ریزی توسعه مخازن گاز میعانی و طراحی و بهینه‌سازی سیستم‌های تولید بسیار مهم است. اگرچه تعیین تجربی فشار نقطه شبنم دقیق‌ترین و قابل اعتمادترین نتایج را ارائه می‌دهد، اما این فرآیند هزینه‌بر و زمان‌بر است و گاهی با خطاهایی همراه است. هامود و همکاران مشکلات مختلفی را که با روش‌های تجربی همراه است، مانند هزینه‌ها و دشواری‌های تحلیل‌های آزمایشگاهی، عدم توانایی در گرفتن نمونه نماینده، خطاهای مرتبط با تحلیل‌های آزمایشگاهی و کمبود حجم نمونه برای انجام تحلیل‌های کامل، ذکر کرده‌اند [۳]. در نتیجه، اندازه‌گیری‌های آزمایشگاهی همیشه در دسترس نیستند. این واقعیت توسعه روش‌های ساده اما به اندازه کافی دقیق را تشویق می‌کند. روش‌های مبتنی بر معادلات حالت جایگزین‌های شناخته‌شده‌ای هستند. کاربرد روش‌های مبتنی بر معادلات حالت نیازمند مشخص‌سازی دقیق کسر هپتان به علاوه (+C7) است. ال‌شارکوی و همکاران [۴] اشاره کرده است که حتی یک تحلیل کامل نمونه سیال نمی‌تواند تضمین کند که روش‌های مبتنی بر معادلات حالت پیش‌بینی دقیقی از فشار نقطه شبنم ارائه دهند. برخی از روش‌ها بر اساس تخمین نسبت تعادل K-values حاصل از همبستگی‌ها هستند که با آزمون و خطا و عدم قطعیت‌ها به ویژه در دماها و فشارهای بالاتر همراه است. همبستگی‌های مختلف، روش‌های مبتنی بر هوش مصنوعی، روش‌های مبتنی بر معادلات حالت و حتی روش‌های گرافیکی برای پیش‌بینی فشار نقطه شبنم گازهای میعانی پیشنهاد شده‌اند. گنزالس و همکاران [۵] از یک مدل شبکه عصبی برای پیش‌بینی فشار نقطه شبنم مخازن گاز رتروگراد استفاده کردند. در این راستا، آنها از ۸۰۲ داده تجربی تخلیه حجم ثابت استفاده کردند. خطای مطلق متوسط ۸/۷۴٪ گزارش شد. در سال ۲۰۰۷، جلیلی و همکاران [۶] از ۱۱۱ نمونه داده برای توسعه شبکه‌های

عصبی مصنوعی مختلف برای پیش‌بینی فشار نقطه شب‌نم استفاده کردند. آنها نتیجه گرفتند که الگوریتم آموزش لوونبرگ-مارکوارت بهترین نتایج را ارائه می‌دهد. در سال ۲۰۱۱، الدهام و المرهون [۲] مدل‌های مختلفی از رگرسیون غیرخطی چندگانه، مدل رگرسیون غیرپارامتریک (تکنیک انتظارات شرطی متناوب) و شبکه عصبی مصنوعی با استفاده از تعداد کل ۱۱۳ نمونه داده به دست آمده از آزمایش‌های انبساط جرم ثابت جمع‌آوری شده از میدان‌های خاورمیانه توسعه دادند. طی این گزارش مدل شبکه عصبی مصنوعی بهترین نتایج را ارائه داد. گادوین و همکاران [۷] یک همبستگی جدید برای پیش‌بینی فشار نقطه شب‌نم گازهای میعانی با استفاده از ۲۵۹ نمونه داده پیشنهاد کرد. در مقایسه با همبستگی‌های موجود، این مطالعه برتری همبستگی توسعه یافته را ادعا کرد. الزهابی و همکاران [۸] همبستگی بر اساس داده‌های تحلیل سیال درون چاهی پیشنهاد کردند. آنها از رگرسیون خطی چندگانه استفاده کردند و از ۶۶۷ نمونه داده برای توسعه مدل خود استفاده کردند. نتایج این مطالعه خطای حدود ۲٪ را نشان داد. هدف این مطالعه انجام یک مرور و تحلیل بر داده‌ها و مدل‌های پیش‌بینی‌کننده فشار نقطه شب‌نم گازهای میعانی و همچنین ارزیابی عملکرد مدل‌های مختلف می‌باشد. لازم به ذکر است تنها داده‌های مربوط به مخازن گاز میعانی هدف این مطالعه هستند. پاکسازی داده‌ها انجام خواهد شد. کاربرد ۸ مدل مورد بررسی خواهد شد. در مدل‌های شبکه عصبی مصنوعی، الگوریتم درخت تصمیم و جنگل تصادفی با تنظیم هیپرپارامترها نسبت به اخذ بهترین نتایج پیش‌بینی اقدام خواهد شد. دقت مدل‌ها هم از نظر آماری و هم گرافیکی ارزیابی خواهند شد. در نهایت نتایج مورد مقایسه و همچنین نتیجه‌گیری‌ها و توصیه‌ها ارائه خواهند شد. تحقیق حاضر به‌عنوان یک نوآوری، به استفاده از روش‌های هوش مصنوعی در پیش‌بینی فشار نقطه شب‌نم پرداخته و نشان می‌دهد که با تحلیل دقیق داده‌های گاز میعانی و به‌کارگیری مدل‌های پیشرفته، می‌توان به دقت و کارایی بالاتری در پیش‌بینی‌های صنعتی دست یافت. همچنین، نتایج این مطالعه می‌تواند به عنوان مبنایی برای تحقیقات آینده در زمینه بهینه‌سازی مدل‌های پیش‌بینی و بهبود فرآیندهای صنعتی در صنعت گاز میعانی مورد استفاده قرار گیرد.

۲- تئوری مدل‌سازی

در این مطالعه با توجه به داده‌های جمع‌آوری شده از روش‌های مختلف مدل‌سازی استفاده شده است. جزئیات روش‌های مدل‌سازی (مدل رگرسیون خطی، لاسو، ریج، الاستیک نت، نزدیکترین همسایگی، شبکه عصبی مصنوعی، درخت تصمیم و جنگل تصادفی) به شرح زیر می‌باشد [۹-۱۲].

• مدل رگرسیون خطی:

رگرسیون خطی یک مدل پیش‌بینی‌کننده است. برای مسائل رگرسیون (پیش‌بینی مقادیر پیوسته) استفاده شده که رابطه خطی بین متغیرهای مستقل X و متغیر وابسته Y را مدل‌سازی می‌کند. از مزایای این روش می‌توان به سادگی و قابلیت تفسیر با سرعت بالا در آموزش و پیش‌بینی و پایه‌ای برای مدل‌های پیچیده‌تر اشاره نمود. همچنین معایب آن حساسیت به داده‌های پرت است که در صورت نقض فرضیات (مثل غیرخطی بودن)، عملکرد ضعیفی دارد.

• مدل رگرسیونی نزدیکترین همسایگی:

یک مدل غیرپارامتری و پیش‌بینی کننده برای مسائل رگرسیون است که بر اساس همسایگی داده‌های آموزشی پیش‌بینی انجام می‌دهد. مدل مقدار مورد نظر متغیر وابسته Y را با میانگین‌گیری از مقادیر همسایه نزدیک در فضای ویژگی‌ها X تخمین می‌زند. از مزایای روش فوق می‌توان به ساده و تفسیرپذیر، بدون فرض خطی بودن (برخلاف رگرسیون خطی) و تنظیم آسان بوده و همچنین از معایب آن حساس به مقیاس ویژگی‌ها (نیاز به نرمال‌سازی)، محاسبات سنگین برای داده‌های بزرگ (چون تمام داده‌ها ذخیره می‌شوند) و کارایی ضعیف در ابعاد بالا می‌باشد.

• مدل رگرسیون لاسو:

یک مدل خطی تنظیم‌شده برای مسائل رگرسیون است که با اضافه کردن جریمه $L1$ به تابع هزینه، برخی از ضرایب را دقیقاً صفر می‌کند و بنابراین انتخاب ویژگی را انجام می‌دهد. از ویژگی‌های این مدل می‌توان به انتخاب ویژگی، مقابله با همخطی و جلوگیری از بیش‌برازش اشاره نمود.

• مدل رگرسیون ریج:

یک مدل خطی تنظیم‌شده برای مسائل رگرسیون است که با اضافه کردن جریمه $L2$ به تابع هزینه، از بیش‌برازش جلوگیری می‌کند.

• مدل الاستیک نت:

ترکیبی هوشمند از $L1$ لاسو و ریج $L2$ برای رگرسیون خطی تنظیم‌شده که هم انتخاب ویژگی و هم کاهش بیش‌برازش را انجام می‌دهد.

• مدل شبکه عصبی مصنوعی:

یک سیستم یادگیری عمیق انعطاف‌پذیر که برای پیش‌بینی مقادیر پیوسته با تقلید از ساختار مغز انسان طراحی شده است. ساختار کلیدی آن شامل سه لایه ورودی (تعداد نرون)، پنهان و خروجی می‌باشد. مزایای کلیدی مدل فوق یادگیری روابط پیچیده، مقیاس‌پذیری، انعطاف معماری و یادگیری انتها به انتها بوده و چالش‌های آن نیاز به داده (چند هزار نمونه)، پیدا کردن هایپر پارامترها، منابع محاسباتی و تفسیرپذیری می‌باشند.

• مدل درخت تصمیم‌گیری:

مدل درخت تصمیم‌گیری یک روش یادگیری ماشین نظارت شده است که با تقسیم داده‌ها به زیرمجموعه‌های کوچکتر بر اساس معیارهای مختلف (مانند آنتروپی یا جینی)، الگوهای موجود در داده‌ها را استخراج می‌کند. این مدل به صورت درختی از سؤالات بله/خیر ساخته می‌شود و برای هر دو مسئله طبقه‌بندی و رگرسیون قابل استفاده است. سادگی تفسیر، عدم نیاز به نرمال‌سازی داده‌ها و کارایی مناسب روی داده‌های پیچیده از مزایای آن هستند.

• مدل جنگل تصادفی:

مدل جنگل تصادفی یک روش یادگیری ماشین مبتنی بر ادغام چندین درخت تصمیم است که با ترکیب پیش‌بینی‌های آنها، دقت و کاهش واریانس را بهبود می‌بخشد. این مدل از انتخاب تصادفی ویژگی‌ها و نمونه‌های

داده برای ساخت هر درخت استفاده می‌کند تا از بیش‌برازش جلوگیری کند. جنگل تصادفی برای طبقه‌بندی و رگرسیون مناسب است و به دلیل کارایی بالا و مقاومت در برابر نویز، پرکاربرد است.

۳- روش مدل‌سازی

۳-۱- داده‌های استفاده شده:

یک مرور جامع برای جمع‌آوری تمام داده‌های موجود در مورد فشار نقطه شب‌نم گازهای میعانی انجام شد. در این راستا، ۷۲۱ نمونه داده از ۲۲ مطالعه منتشر شده از سال ۱۹۴۲ تا ۲۰۱۸ جمع‌آوری گردیده است [۱۳]. اکثریت داده‌ها از مطالعات منتشر شده در دهه ۱۹۹۰ و ۲۰۰۰ جمع‌آوری شده‌اند. محدوده فشار شب‌نم داده‌های جمع‌آوری شده ۱۴۰۵ - ۱۱۸۳۰ پام می‌باشد. قبل از شروع مدل‌سازی، داده‌ها باید از نظر خطاها و ناسازگاری‌ها بررسی شوند. در این مطالعه تعداد ۷۲۱ داده مورد بررسی قرار گرفته و تعداد ۲۰ داده موجود نبوده که حذف شده و تعداد ۷۰۱ نمونه داده برای مدل‌سازی باقی ماندند.

۳-۲- روش ساخت مدل:

در این پژوهش، برای پیش‌بینی فشار نقطه شب‌نم میعانات گازی، از هشت مدل مختلف یادگیری ماشین استفاده شد. این مدل‌ها شامل روش‌های خطی، منظم‌شده، مبتنی بر درخت، و شبکه‌های عصبی بودند تا بتوانند طیف وسیعی از روابط خطی و غیرخطی بین متغیرهای مستقل و وابسته را پوشش دهند. مراحل ساخت مدل به شرح زیر بود:

۳-۲-۱- پیش پردازش داده‌ها

پایگاه داده شامل ۷۲۱ نمونه پس از پاک‌سازی و حذف ۲۰ نمونه ناقص یا ناسازگار، به ۷۰۱ نمونه کاهش یافت. این داده‌ها شامل ۱۰ متغیر مستقل (ترکیبات گاز، جرم مخصوص، وزن مولکولی، دما) و یک متغیر وابسته (فشار نقطه شب‌نم) بودند. در ابتدا داده‌ها به دو مجموعه آموزش (۸۰٪) و آزمون (۲۰٪) تقسیم شدند.

۳-۲-۲- انتخاب و آموزش مدل‌ها

در فرآیند انتخاب و آموزش مدل‌ها، هشت الگوریتم مختلف یادگیری ماشین به صورت سیستماتیک پیاده‌سازی و آموزش داده شدند. مدل رگرسیون خطی به عنوان پایه‌ای برای سنجش عملکرد مدل‌های پیچیده‌تر مورد استفاده قرار گرفت که رابطه خطی بین متغیرها را مدل‌سازی می‌کند. برای مقابله با چالش همخطی در داده‌ها، از مدل رگرسیون ریبج با اعمال جریمه L2 بهره گرفته شد که با محدود کردن ضرایب مدل، از بیش‌برازش جلوگیری می‌کند. همچنین مدل رگرسیون لاسو با اعمال جریمه L1، علاوه بر کنترل بیش‌برازش، امکان انتخاب ویژگی‌های مهم را با صفر کردن ضرایب کم‌اهمیت فراهم آورد. مدل الاستیک‌نت نیز با ترکیب هوشمندانه دو جریمه L1 و L2، تعادل مناسبی بین انتخاب ویژگی و تنظیم مدل ایجاد کرد. در کنار این مدل‌های خطی، از الگوریتم نزدیک‌ترین همسایه به عنوان یک روش غیرپارامتری مبتنی بر شباهت استفاده شد که بر اساس همسایگی داده‌ها اقدام به پیش‌بینی می‌کند. شبکه عصبی مصنوعی با یک لایه پنهان برای یادگیری روابط غیرخطی پیچیده بین متغیرها به کار گرفته شد. همچنین مدل درخت تصمیم برای مدل‌سازی روابط غیرخطی و شناسایی تعاملات بین ویژگی‌ها مورد استفاده قرار گرفت. در نهایت، الگوریتم جنگل تصادفی با ترکیب چندین درخت تصمیم و استفاده از رویه تجمیع، دقت پیش‌بینی را افزایش داده و واریانس را کاهش داد.

۳-۲-۳- بهینه‌سازی پارامترهای مدل (هیپرپارامترها)

برای هر مدل، پارامترهای کلیدی (مانند تعداد درختان در جنگل، عمق درخت، نرخ یادگیری در شبکه عصبی، تعداد همسایگان) با استفاده از روش جستجوی شبکه و اعتبارسنجی متقابل بهینه شدند.

۳-۲-۴- ارزیابی مدل‌ها

برای ارزیابی جامع عملکرد مدل‌های توسعه‌یافته، از سه معیار آماری کلیدی استفاده شد. ضریب همبستگی به عنوان شاخص اصلی برای سنجش میزان انطباق مدل با داده‌ها به کار رفت که درصد تغییرات تبیین‌شده توسط مدل را نشان می‌دهد. مقادیر نزدیک به یک برای این شاخص حاکی از توانایی بالای مدل در پیش‌بینی دقیق متغیر وابسته است. میانگین درصد خطای نسبی به عنوان معیاری برای اندازه‌گیری دقت پیش‌بینی‌ها مورد استفاده قرار گرفت که میانگین درصد انحراف پیش‌بینی‌ها از مقادیر واقعی را محاسبه می‌کند. این شاخص به ویژه از آن جهت حائز اهمیت است که قابلیت تفسیرپذیری بالایی داشته و امکان مقایسه مستقیم بین مدل‌های مختلف را فراهم می‌آورد. همچنین میانگین مربعات خطا به عنوان معیاری حساس به خطاهای بزرگ به کار گرفته شد که با توان دوم رساندن تفاوت بین مقادیر پیش‌بینی‌شده و واقعی، بر اهمیت خطاهای بزرگتر تأکید می‌کند. این سه معیار در کنار یکدیگر، ارزیابی جامعی از جنبه‌های مختلف عملکرد مدل‌ها شامل دقت، توانایی تبیین تغییرات و حساسیت به خطاهای بزرگ ارائه می‌دهد.

۴- نتایج و بحث

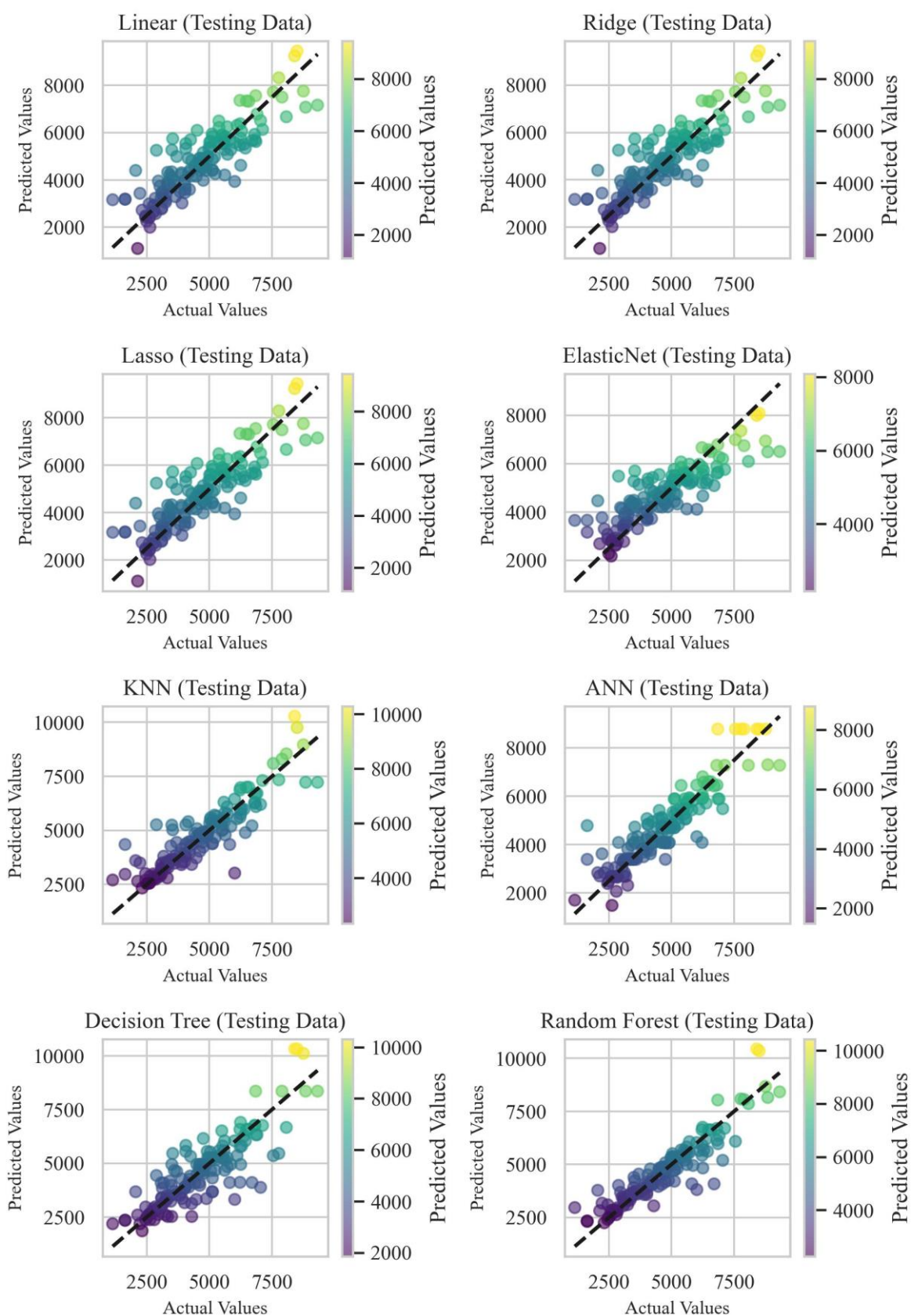
در این مطالعه، به منظور ارزیابی دقت مدل‌های مختلف یادگیری ماشین و آماری در پیش‌بینی متغیر هدف، از معیارهای درصد میانگین خطای نسبی، ضریب رگرسیون (R^2) و میانگین مربعات خطا استفاده شد. نتایج عددی حاصل از اجرای مدل‌ها در دو گروه آموزش و تست در جدول ۱ ارائه شده است.

همان‌طور که در جدول ۱ مشاهده می‌شود، مدل‌های رگرسیونی خطی ساده، لاسو، ریب و الاستیک نت همگی عملکرد مشابهی دارند و با R^2 برابر با ۰/۷۵، نشان‌دهنده این هستند که این مدل‌ها قادر به توضیح تنها ۷۵ درصد از واریانس داده‌ها هستند. این موضوع نشان می‌دهد که این مدل‌ها نمی‌توانند به خوبی روابط پیچیده و غیرخطی موجود در داده‌ها را شناسایی کنند. به عنوان مثال، رگرسیون خطی با میانگین مربعات خطای ۶۷۵۳۸۶ و خطای نسبی ۱۲/۵ درصد، نشان می‌دهد که پیش‌بینی‌های آن قابل قبول نیستند و این مدل نمی‌تواند به خوبی نیازهای پیش‌بینی را برآورده کند. از طرفی مدل شبکه‌های عصبی مصنوعی عملکرد بهتری دارد و با R^2 برابر با ۰/۸۲ و درصد میانگین خطای نسبی ۷/۲، نشان‌دهنده توانایی بالای این مدل در شبیه‌سازی داده‌ها و پیش‌بینی دقیق‌تر فشار نقطه شبنم است. این مدل می‌تواند روابط پیچیده‌تری را در داده‌ها شناسایی کند و به همین دلیل دقت بیشتری در پیش‌بینی‌ها دارد. مدل درخت تصمیم‌گیری نیز به عنوان یکی از مدل‌های با عملکرد بالا شناخته می‌شود که با R^2 آموزش برابر با ۰/۹۸ و درصد میانگین خطای نسبی ۳/۵ برای گروه تست، نشان می‌دهد که این مدل به خوبی می‌تواند الگوهای موجود در داده‌ها را شناسایی کند و پیش‌بینی‌های دقیقی ارائه دهد. مدل نزدیکترین همسایگی (KNN) نیز با R^2 برابر با ۰/۸۰ و درصد میانگین خطای نسبی ۶/۸ برای گروه تست، نشان‌دهنده قابلیت‌های خوب این مدل در پیش‌بینی است، اما هنوز به دقت مدل‌های جنگل تصادفی و درخت تصمیم‌گیری نمی‌رسد. مدل جنگل تصادفی که در اینجا به عنوان بهترین مدل

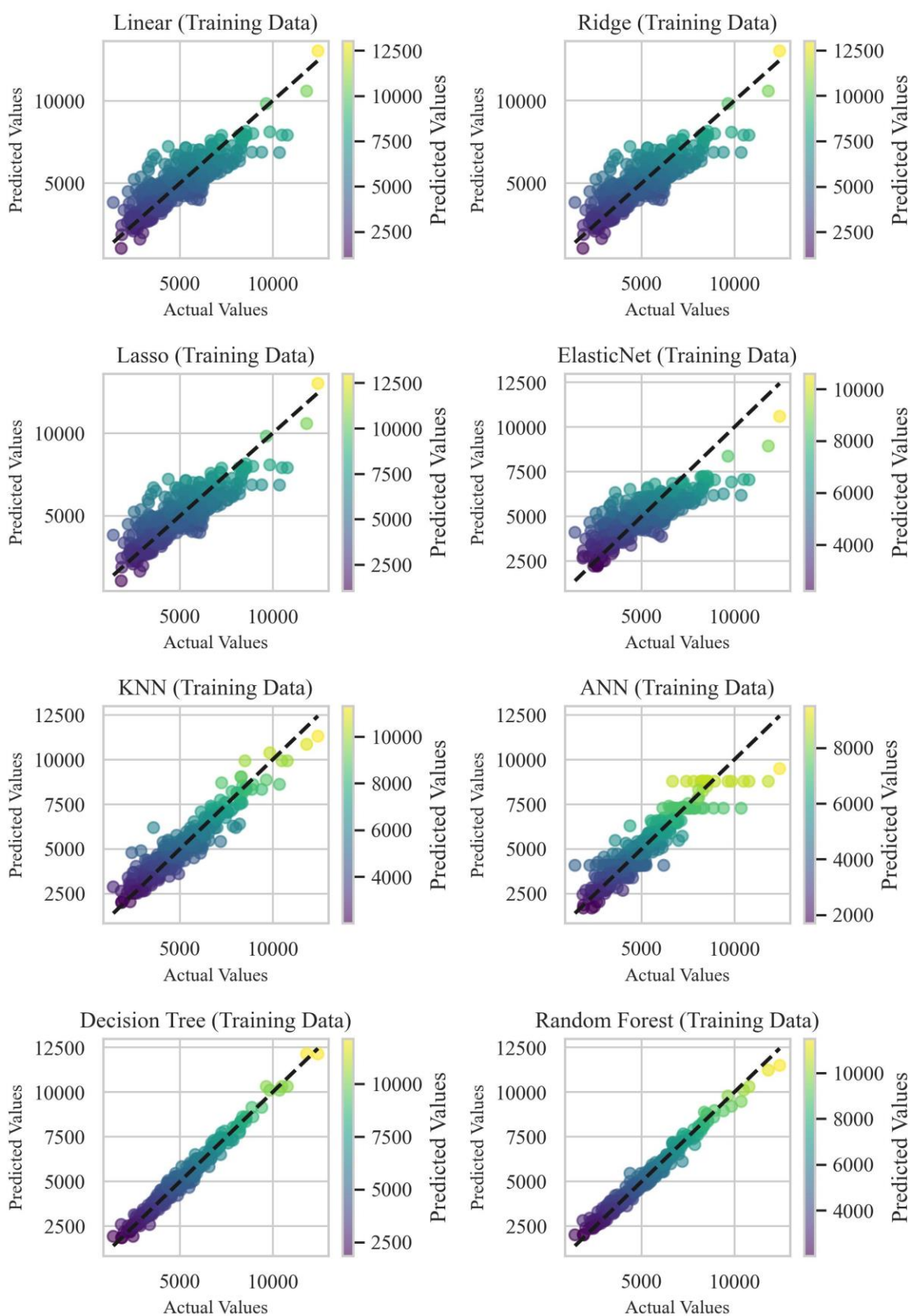
معرفی شده است، با R^2 آموزش برابر با ۰/۹۸ و درصد میانگین خطای نسبی ۳/۲۶ برای گروه تست، به وضوح نشان می‌دهد که می‌تواند به بهترین نحو به پیش‌بینی فشار نقطه شب‌نم پردازد. این مدل به دلیل ساختار درختی و توانایی در شناسایی تعاملات غیرخطی و پیچیدگی‌های موجود در داده‌ها، عملکرد بی‌نظیری دارد.

جدول ۱- ارزیابی آماری مدل‌های مختلف ارائه شده در این مطالعه

مدل	گروه	میانگین مربعات خطا	R^2	درصد میانگین خطای نسبی
رگرسیون خطی	گروه آموزش	۶۷۵۳۸۶	۰/۷۵	۱۳/۵
	گروه تست	۶۶۴۳۱۹	۰/۷۵	۱۲/۵
رگرسیون لاسو	گروه آموزش	۶۷۶۰۶۳	۰/۷۵	۱۳/۴
	گروه تست	۶۶۲۸۹۶	۰/۷۵	۱۲/۵
رگرسیون ریج	گروه آموزش	۶۷۵۳۸۷	۰/۷۵	۱۳/۴
	گروه تست	۶۶۲۸۹۶	۰/۷۵	۱۲/۵
الاستیک نت	گروه تست	۸۲۸۲۵۷	۰/۷۵	۱۳/۴
	گروه آموزش	۷۹۳۳۹۸	۰/۷۵	۱۲/۴
شبکه عصبی مصنوعی	گروه آموزش	۲۹۸۸۵۰	۰/۸۹	۱۱/۴
	گروه تست	۴۹۰۲۲۲	۰/۸۲	۷/۲
نزدیکترین همسایه	گروه آموزش	۲۲۸۳۴۰	۰/۹۱	۱۰.۵
	گروه تست	۵۳۲۵۸۵	۰/۸۰	۶.۸
درخت تصمیم‌گیری	گروه آموزش	۵۱۲۱۱	۰/۹۸	۱۴.۳
	گروه تست	۷۹۲۴۹۴	۰/۷	۳.۵
جنگل تصادفی	گروه آموزش	۵۰۰۵۸	۰/۹۸	۹.۴۶
	گروه تست	۴۱۹۵۱۱	۰/۸۴	۳.۲۶



شکل ۱- نمودار رگرسیون مربوط به داده‌های تست مدل‌های ارائه شده



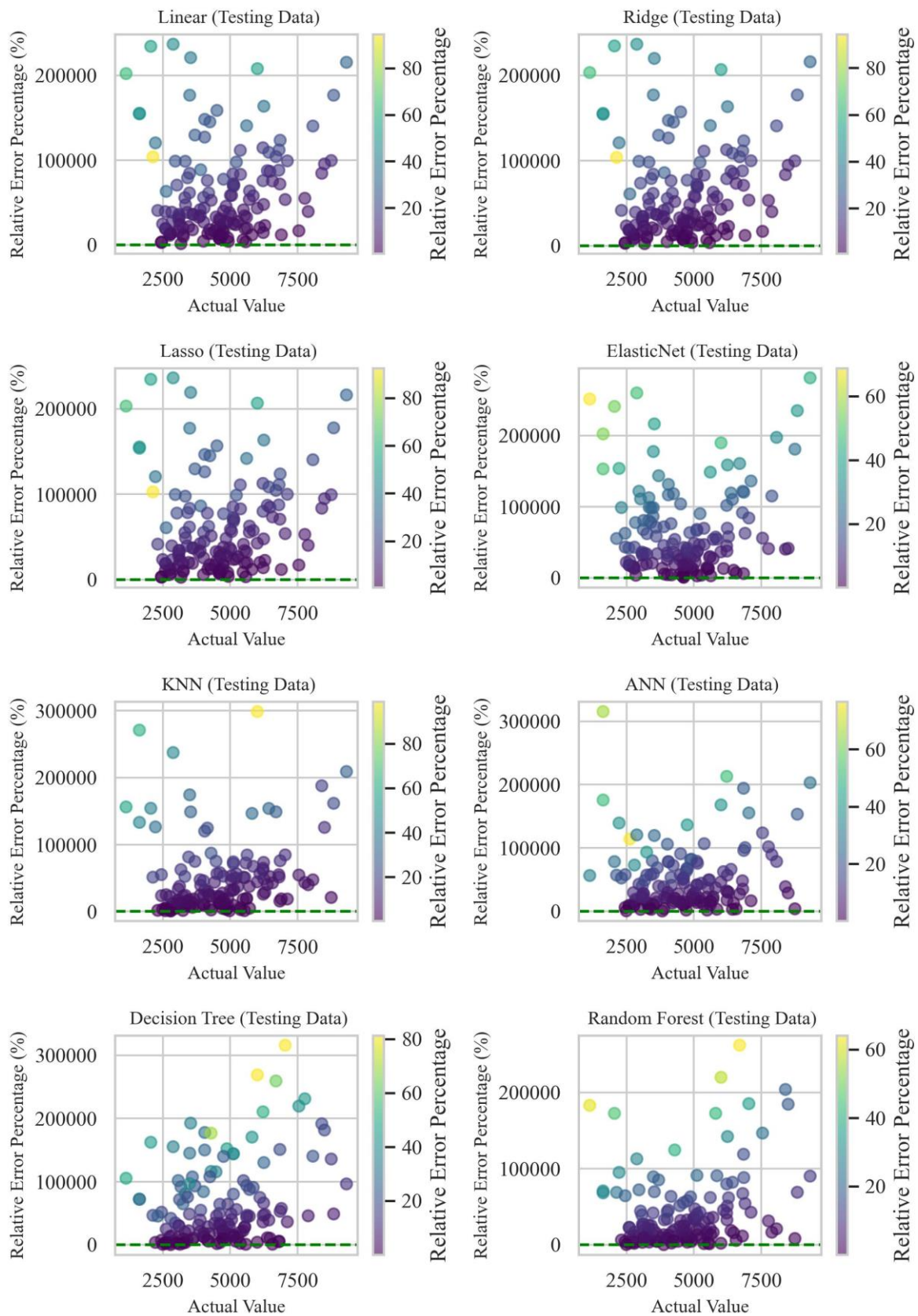
شکل ۲- نمودار رگرسیون مربوط به داده‌های آموزش مدل‌های ارائه شده

برای درک بهتر نحوه عملکرد مدل‌ها، نمودارهای رگرسیون بین مقادیر واقعی و پیش‌بینی شده برای هر مدل در دو گروه آموزش و تست در شکل ۱ و شکل ۲ ارائه شده‌اند. بررسی این نمودارها نشان می‌دهد که مدل‌های درختی (به‌ویژه جنگل تصادفی) تطابق بسیار بهتری بین مقادیر واقعی و پیش‌بینی شده برقرار کرده‌اند، در حالی که مدل‌های آماری کلاسیک پراکندگی بیشتری را نشان می‌دهند. همچنین برای مقایسه بهتر میزان دقت و خطای مدل‌ها، نمودارهای خطای نسبی برای هر مدل در دو مجموعه آموزش و تست در شکل ۳ و شکل ۴ نمایش داده شده‌اند. مدل‌های مبتنی بر درخت به ویژه جنگل تصادفی به دلیل توانایی در شناسایی روابط غیرخطی، عملکرد بهتری داشتند. برخلاف درخت تصمیم که دچار بیش‌برازش شد، جنگل تصادفی با ترکیب چندین درخت، تعادل مناسبی ایجاد کرد. مدل‌های خطی به دلیل سادگی ذاتی نتوانستند روابط پیچیده داده‌ها را شناسایی کنند. در نهایت، جنگل تصادفی به عنوان بهترین مدل برای پیش‌بینی فشار نقطه شبنم میعان‌ات گازی معرفی شد. این نتایج نشان می‌دهند که انتخاب مدل مناسب در پیش‌بینی خواص سیالات می‌تواند تأثیر به‌سزایی در دقت پیش‌بینی‌ها داشته باشد. به نظر می‌رسد که روش‌های یادگیری ماشین، به ویژه مدل‌های پیچیده‌تر، می‌توانند ابزارهای مؤثری برای پیش‌بینی فشار نقطه شبنم و بهینه‌سازی سیستم‌های تولید در صنعت نفت و گاز باشند. این‌ها به مهندسان و محققان در بهبود فرآیندها و مدیریت منابع کمک می‌کند و نشان‌دهنده نیاز به تحقیقات بیشتر در زمینه بهینه‌سازی مدل‌ها و داده‌هاست.

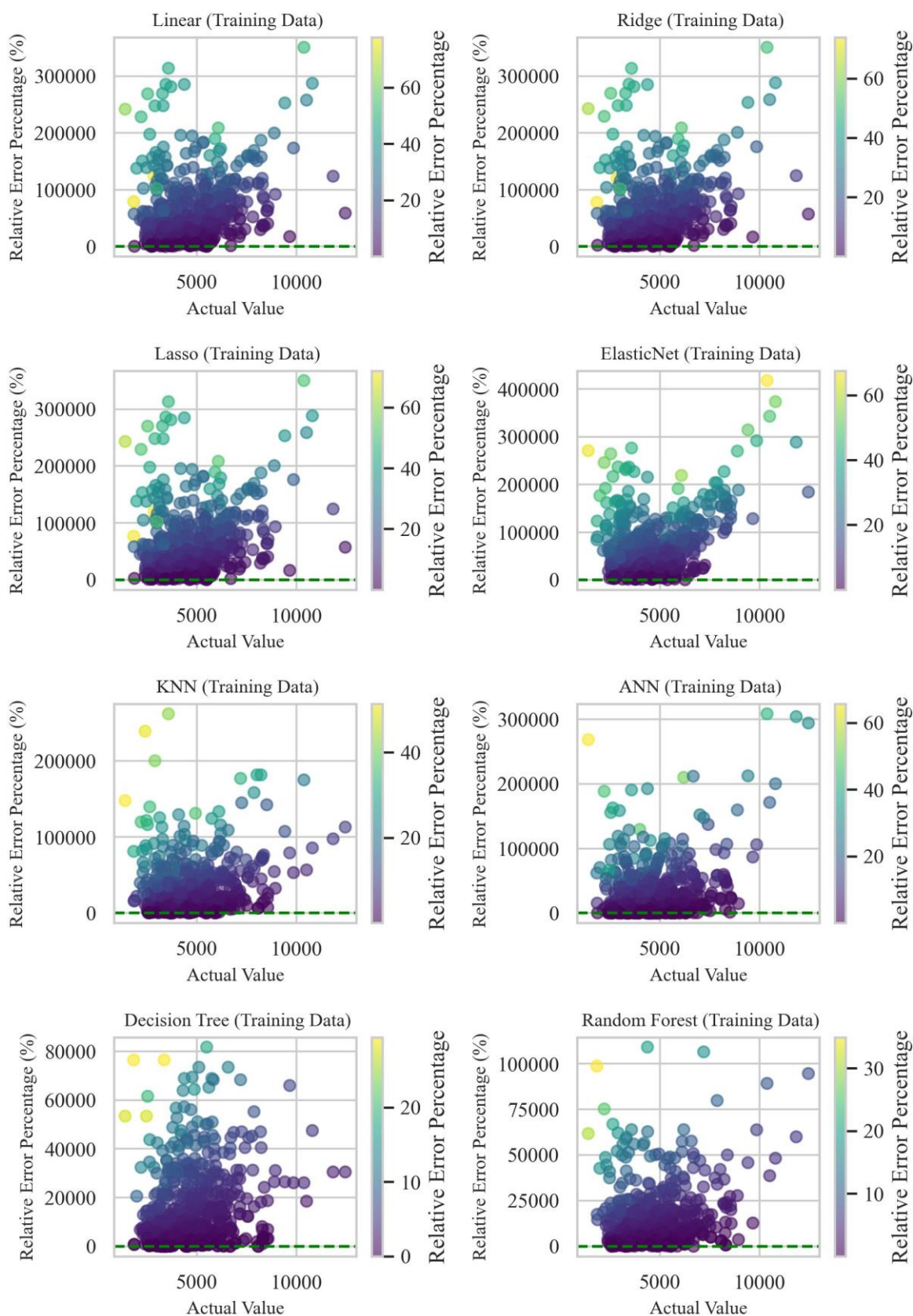
در جدول ۲ مقدار درصد میانگین خطای نسبی مدل‌های ارائه شده در کار حاضر با مدل‌های موجود در مقالات انتشار یافته پیشین مقایسه شده است.

جدول ۲- مقایسه آماری مدل‌های توسعه یافته در این پژوهش با مدل‌های پیشین

مدل	درصد میانگین خطای نسبی	مرجع
رابطه تجربی	۵/۸	[14]
رابطه تجربی	۷/۶۸	[4]
EOS (SRK ¹)	۱۲/۳۴	[4]
EOS (SRK ²)	۸/۶۴	[4]
EOS (PR ¹)	۹/۲۵	[4]
EOS (PR ²)	۸/۶۴	[4]
MLP	۸/۲۳۶	[5]
GP	۴/۲	[15]
PR	۵/۶۵	[15]
انفیس	۵/۱۳	[16]
CSA-LSSVM	۶/۳	[17]
درخت تصمیم	۳/۵	پژوهش حاضر
جنگل تصادفی	۳/۲۶	پژوهش حاضر
نزدیکترین همسایه	۶/۸	پژوهش حاضر
شبکه عصبی مصنوعی	۷/۲	پژوهش حاضر
رگرسیون خطی	۱۲/۵	پژوهش حاضر



شکل ۳- نمودارهای درصد خطای نسبی مربوط به داده‌های تست مدل‌های ارائه شده



شکل ۴- نمودارهای درصد خطای نسبی مربوط به داده های آموزش مدل های ارائه شده

۵- نتیجه گیری

مدل جنگل تصادفی در داده‌های تست کمترین میانگین خطا را در مقایسه با سایر مدل‌ها دارد. این نشان دهنده عملکرد قوی مدل در پیش‌بینی داده‌های جدید است. مقدار ضریب همبستگی مدل فوق در داده‌های آموزش و تست به ترتیب ۰/۹۸ و ۰/۸۴ است که نشان دهنده توانایی بالا در تبیین واریانس داده‌ها است. اگر چه مدلی مانند درخت تصمیم‌گیری در داده‌های آموزشی دارای ضریب همبستگی بالایی (۰/۹۸) بوده اما اختلاف قابل توجهی بین عملکرد آنها در داده‌های آموزش و تست وجود دارد در صورتی که در مدل جنگل تصادفی تعادل بهتری بین داده‌های آموزش و تست نشان می‌دهد. همچنین نمودار هیستوگرام مربوط به درخت تصمیم‌گیری وجود بیش برآزش مربوط به داده‌های آموزش را نشان می‌دهد. لذا با توجه به معیارهای بالا مدل جنگل تصادفی بهترین مدل انتخاب می‌شود، زیرا عملکرد متعادل و قوی در هر دو مجموعه داده دارد و از اورفیتینگ اجتناب می‌کند. در ادامه این تحقیق موارد زیر برای مطالعات آینده پیشنهاد می‌شود:

- در این مطالعه، تنها نمونه‌های داده گزارش شده برای مخازن گاز میعانی برای تحلیل در نظر گرفته شده‌اند. مطالعاتی وجود دارند که داده‌های مربوط به مخلوط‌های گاز سنتزی و مخلوط‌های گاز طبیعی مشتق شده از خطوط لوله را گزارش می‌کنند. داده‌های مطالعات ذکر شده می‌توانند برای توسعه یک پایگاه داده جامع‌تر برای توسعه مدل‌های عمومی‌تر استفاده شوند.

- مقادیر گم‌شده می‌توانند با یک نقطه میانی (میانگین حسابی، میانه یا مد) جایگزین شوند. علاوه بر این، درخت‌های تصمیم تقویت‌شده گرادیان و درخت‌های طبقه‌بندی و رگرسیون می‌توانند زمانی که داده‌ها در حالت پراکنده هستند، استفاده شوند. ارزیابی اثرات استفاده از چنین روش‌هایی برای مطالعات آینده توصیه می‌شود.

۶- منابع

- [1] Daneshfar R, Keivanimehr F, Mohammadi-Khanaposhtani M, Baghban A. A neural computing strategy to estimate dew-point pressure of gas condensate reservoirs. *Pet Sci Technol*; 38:706–12, 2020.
- [2] Al-Dhamen M, Al-Marhoun M. New correlations for dew-point pressure for gas condensate. *SPE Saudi Arab. Sect. young Prof. Tech. Symp.*, pp. SPE-155410, 2011.
- [3] Humoud AA, Al-Marhoun MA. A new correlation for gas-condensate dewpoint pressure prediction. *SPE Middle East Oil gas show Conf.*, pp. SPE-68230, 2001.
- [4] Elsharkawy AM. Characterization of the plus fraction and prediction of the dewpoint pressure for gas condensate reservoirs. *SPE West. Reg. Meet.*, pp. SPE-68776, 2001.
- [5] González A, Barrufet MA, Startzman R. Improved neural-network model predicts dewpoint pressure of retrograde gases. *J Pet Sci Eng*; 37:183–94, 2003.
- [6] Jalali F, Abdy Y, Akbari MK. Using artificial neural network's capability for estimation of gas condensate reservoir's dew point pressure. *Eur. Conf. Exhib. London, UK, Soc. Pet. Eng.*, pp. 10, 2007.
- [7] Godwin ON. A new analytical method for predicting dew point pressures for gas condensate reservoirs. *SPE Niger. Annu. Int. Conf. Exhib.*, pp. SPE-162985, 2012.
- [8] Alzahabi A, El-Banbi A, Alexandre Trindade A, Soliman M. A regression model for estimation of dew point pressure from down-hole fluid analyzer data. *J Pet Explor Prod Technol*; 7:1173–83, 2017.
- [9] Hassan R, Baghban A. Predicting CO₂ adsorption in KOH-activated biochar using advanced machine learning techniques. *Sci Rep*; 15:24410, 2025.

- [10] Hassan R, Baghban A. Pioneering machine learning techniques to estimate thermal conductivity of carbon-based phase change materials: A comprehensive modeling framework. *Case Stud Therm Eng*; 73:106648, 2025.
- [11] Hassan R, Baghban A. Beyond the surface: Quasi-SMILES machine learning approaches for precise estimation of organic sorption. *Mater Today Commun*: 114126, 2025.
- [12] Soleimani R, Kouchi MM, Behtouei Z, Ghasemi Z, Baghban A. Exploring the solubility potential of anti-cancer and supportive agents in supercritical CO₂ through advanced computational intelligence techniques. *J CO₂ Util*; 102:103227, 2025.
- [13] Mirzaie M, Esfandyari H, Tatar A. Dew point pressure of gas condensates, modeling and a comprehensive review on literature data. *J Pet Sci Eng*; 211:110072, 2022.
- [14] Fang Y, Li B, Hu Y, Sun Z, Zhu Y. Condensate gas phase behavior and development. *SPE Int. Oil Gas Conf. Exhib. China*, pp. SPE-50925, 1998.
- [15] Shokir EME-M. Dewpoint pressure model for gas condensate reservoirs based on genetic programming. *SPE Unconv. Resour. Conf. Technol. Symp.*, , pp. SPE-114454, 2008.
- [16] Nowroozi S, Ranjbar M, Hashemipour H, Schaffie M. Development of a neural fuzzy system for advanced prediction of dew point pressure in gas condensate reservoirs. *Fuel Process Technol*;90:452–7, 2009.
- [17] Arabloo M, Shokrollahi A, Gharagheizi F, Mohammadi AH. Toward a predictive model for estimating dew point pressure in gas condensate systems. *Fuel Process Technol*;116:317–24, 2013.