

بسم الله الرحمن الرحيم

مروری بر یادگیری ماشین در

اکتشاف نفت و گاز

Machine Learning in
Oil and Gas Exploration:
A Review

Translation and Prepration by

Digi Hoosh HDM

SABA

Ali Morshedsolouk

March ۲۰۲۵ | ۱۴۰۳ Esdand



فهرست

خلاصه مقاله:..... ۴

۱. مقدمه:..... ۴

۲. روش شناسی:..... ۴

۳. اکتشاف زمین شناسی و ژئوفیزیکی:..... ۵

۴. طبقه بندی فاسیس و لیتوفاسیس:..... ۵

۵. پیش بینی خواص پتروفیزیکی:..... ۵

۶. چالش ها و راه حل ها:..... ۵

۷. نتیجه گیری خلاصه:..... ۶

یادگیری ماشین در اکتشاف نفت و گاز: یک مرور ۷

چکیده:..... ۷

۱. مقدمه ۸

۱۰. اکتشاف زمین شناسی و ژئوفیزیکی..... ۱۰

A. پردازش داده های لرزه ای..... ۱۱

B. طبقه بندی فاسیس و لیتوفاسیس ۱۳

IV. توصیف مخزن ۱۵

V. پیش بینی خواص پتروفیزیکی..... ۱۵

A. نفوذپذیری ۱۶

B. تخلخل ۲۱

C. پیش بینی اشباع آب ۲۵

VI. بحث ۲۹

VII. چالش ها ۳۲

A. مسائل مربوط به داده ها ۳۲

۳۶.....	B. شفافیت و تفسیرپذیری مدل ها
۳۸.....	C. تخصص حوزه‌ای
۳۹.....	VIII. نتیجه‌گیری
۴۵.....	معرفی تیم دیبی هوش هما (صبا)



به نام خالق هوش

مقاله "Machine Learning in Oil and Gas Exploration: A Review" به بررسی کاربردهای یادگیری ماشین در صنعت نفت و گاز، به ویژه در مراحل اکتشاف و ارزیابی مخازن می‌پردازد. این مقاله با تمرکز بر پردازش داده‌های لرزه‌ای، طبقه‌بندی فاسیس‌ها و پیش‌بینی خواص پتروفیزیکی مانند تخلخل، نفوذپذیری و اشباع آب، چالش‌ها و راه‌حل‌های بالقوه را شناسایی می‌کند.

ترجمه این مقاله را با کمک هوش مصنوعی و سپس ویراستاری آن انجام داده ام. در این ترجمه، ابتدا نگاهی به خلاصه‌ای از مقاله خواهیم داشت و سپس بخشهای اصلی مقاله را به طور کامل مرور خواهیم کرد.

خلاصه مقاله:

۱. مقدمه:

صنعت نفت و گاز یکی از پیچیده‌ترین صنایع است که شامل فعالیتهای گسترده‌ای در زنجیره ارزش خود می‌شود. بخش اکتشاف و تولید (Upstream) داده‌های زیادی مانند بررسی‌های زمین‌شناسی، لاگ‌های چاه و اطلاعات حفاری تولید می‌کند. با پیشرفت فناوری‌های جمع‌آوری داده‌های لرزه‌ای، حجم داده‌ها به طور قابل توجهی افزایش یافته است. یادگیری ماشین به عنوان یک ابزار قدرتمند برای پردازش این داده‌ها و بهبود دقت پیش‌بینی‌ها در اکتشاف نفت و گاز مطرح شده است.

۲. روش‌شناسی:

مقاله از یک روش‌شناسی سیستماتیک برای بررسی ادبیات موجود استفاده کرده است. داده‌ها از پایگاه‌های اطلاعاتی معتبر مانند ScienceDirect، IEEE Xplore و Google Scholar جمع‌آوری شده‌اند. کلیدواژه‌هایی مانند "طبقه‌بندی لیتوفاسیس"، "یادگیری ماشین"، "پیش‌بینی نفوذپذیری"، و "پردازش داده‌های لرزه‌ای" مورد استفاده قرار گرفته‌اند. در نهایت، ۱۲۸ مقاله مرتبط انتخاب و تحلیل شده‌اند.

۳. اکتشاف زمین‌شناسی و ژئوفیزیکی:

این بخش به روش‌های اکتشاف زمین‌شناسی و ژئوفیزیکی مانند بررسی‌های لرزه‌ای، الکتریکی و لاگ‌های چاه می‌پردازد. پردازش داده‌های لرزه‌ای یکی از چالش‌های اصلی است که یادگیری ماشین می‌تواند با استفاده از الگوریتم‌هایی مانند شبکه‌های عصبی کانولوشنال (CNN) و ماشین بردار پشتیبان (SVM) به بهبود آن کمک کند. این الگوریتم‌ها برای شناسایی گسل‌ها، افق‌های لرزه‌ای و سایر ویژگی‌های زمین‌شناسی استفاده شده‌اند.

۴. طبقه‌بندی فاسیس و لیتوفاسیس:

فاسیس‌ها و لیتوفاسیس‌ها نقش مهمی در تولید هیدروکربن‌ها دارند. یادگیری ماشین با استفاده از داده‌های لاگ‌های چاه و نمونه‌های هسته‌ای، می‌تواند به طبقه‌بندی لیتوفاسیس‌ها کمک کند. روش‌هایی مانند شبکه‌های عصبی (NN)، SVM و جنگل تصادفی (RF) برای این منظور استفاده شده‌اند. روش‌های ترکیبی (Ensemble Methods) نیز به دلیل دقت بالاتر مورد توجه قرار گرفته‌اند.

۵. پیش‌بینی خواص پتروفیزیکی:

پیش‌بینی خواص پتروفیزیکی مانند تخلخل، نفوذپذیری و اشباع آب برای ارزیابی مخازن ضروری است. یادگیری ماشین با استفاده از داده‌های لاگ‌های چاه و نمونه‌های هسته‌ای، می‌تواند این خواص را با دقت بالا پیش‌بینی کند. روش‌هایی مانند شبکه‌های عصبی مصنوعی (ANN)، سیستم استنتاج فازی-عصبی تطبیقی (ANFIS) و ماشین بردار پشتیبان (SVM) برای این منظور استفاده شده‌اند. همچنین، روش‌های ترکیبی و بهینه‌سازی با الگوریتم‌هایی مانند الگوریتم ژنتیک (GA) و بهینه‌سازی ازدحام ذرات (PSO) بهبود یافته‌اند.

۶. چالش‌ها و راه‌حل‌ها:

- **چالش‌های داده:** داده‌های ناقص، پراکنده و با کیفیت پایین یکی از چالش‌های اصلی است. راه‌حل‌هایی مانند ارتقاء داده‌ها (Data Augmentation)، پیش‌پردازش داده‌ها و استفاده از فناوری‌های سنجش از دور پیشنهاد شده‌اند.
- **شفافیت و تفسیرپذیری مدل‌ها:** بسیاری از مدل‌های یادگیری ماشین به عنوان “جعبه سیاه” عمل می‌کنند که تفسیر آنها دشوار است. استفاده از روش‌های Explainable AI مانند SHAP و LIME می‌تواند به بهبود تفسیرپذیری مدل‌ها کمک کند.
- **تخصص حوزه‌ای:** همکاری بین متخصصان یادگیری ماشین و متخصصان حوزه نفت و گاز برای بهبود دقت و قابلیت اطمینان مدل‌ها ضروری است.



۷. نتیجه گیری خلاصه:

یادگیری ماشین پتانسیل بالایی برای بهبود فرآیندهای اکتشاف نفت و گاز دارد. با این حال، چالش‌هایی مانند کیفیت داده‌ها، تفسیرپذیری مدل‌ها و نیاز به تخصص حوزه‌ای باید مورد توجه قرار گیرند. توسعه چارچوب‌های یادگیری ماشین مقاوم و تطبیقی، بهبود روش‌های پیش‌پردازش داده‌ها و استفاده از روش‌های ترکیبی و بهینه‌سازی می‌تواند به حل این چالش‌ها کمک کند.

این مقاله به عنوان یک مرجع ارزشمند برای محققانی که قصد ورود به این حوزه را دارند، شناخته می‌شود و راهنمایی‌های مفیدی برای تحقیقات آینده ارائه می‌دهد.

یادگیری ماشین در اکتشاف نفت و گاز: یک مرور

احمد لوال، بینگجی یانگ، هونگمی هو و ناتانائل ال. بایسا

۱. دانشکده علوم کامپیوتر و انفورماتیک، دانشگاه د مونتفورت، لستر، بریتانیا

۲. دانشکده علوم، مهندسی و محیط زیست، دانشگاه سالفورد، منچستر، بریتانیا

چکیده:

یک ارزیابی جامع از کاربردهای یادگیری ماشین انجام شده است تا روندهای در حال توسعه برای کاربردهای هوش مصنوعی (AI) در بخش نفت و گاز، به ویژه با تمرکز بر اکتشاف زمین‌شناسی و ژئوفیزیکی و توصیف مخزن، شناسایی شود. حوزه‌های حیاتی مانند پردازش داده‌های لرزه‌ای، طبقه‌بندی فاسیس و لیتوفاسیس، و پیش‌بینی خواص پتروفیزیکی ضروری (مانند تخلخل، نفوذپذیری و اشباع آب) بررسی شده‌اند. با وجود نقش حیاتی این خواص در ارزیابی منابع، پیش‌بینی دقیق آنها همچنان چالش‌برانگیز است. این مقاله مروری دقیق بر مشارکت یادگیری ماشین در پردازش داده‌های لرزه‌ای، طبقه‌بندی فاسیس و پیش‌بینی خواص مخزن ارائه می‌دهد. این مقاله پتانسیل یادگیری ماشین را برای حل چالش‌های مختلف در اکتشاف نفت و گاز، از جمله مدل‌سازی پیش‌بینانه، طبقه‌بندی و خوشه‌بندی، برجسته می‌کند. علاوه بر این، این مرور موانع منحصر به فردی را که مانع از کاربرد گسترده یادگیری ماشین در اکتشاف می‌شوند، شناسایی می‌کند، از جمله عدم قطعیت در پارامترهای زیرسطحی، تفاوت‌های مقیاس و پیچیدگی‌های مربوط به مدیریت داده‌های زمانی و مکانی. این مقاله راه‌حل‌های بالقوه را پیشنهاد می‌کند، روش‌هایی را که به دستیابی به دقت بهینه کمک می‌کنند شناسایی می‌کند و جهت‌های تحقیقاتی آینده را ترسیم می‌کند، که درک دقیقی از پویایی‌های این حوزه ارائه می‌دهد. اتخاذ یادگیری ماشین و روش‌های مدیریت داده‌های قوی برای بهبود کارایی عملیاتی در دوره‌ای که با تولید گسترده داده‌ها مشخص می‌شود، ضروری است. در حالی که محدودیت‌های ذاتی این روش‌ها تصدیق می‌شود، این روش‌ها از محدودیت‌های روش‌های تجربی و تحلیلی سنتی فراتر رفته و خود را به عنوان ابزارهای چندمنظوره برای حل چالش‌های صنعتی تثبیت کرده‌اند. این مرور جامع به عنوان یک منبع ارزشمند برای محققانی که قصد ورود به قلمروهای کمتر شناخته‌شده در این حوزه در حال تکامل را دارند، خدمت می‌کند و بینش‌ها و راهنمایی‌های ارزشمندی برای تحقیقات آینده ارائه می‌دهد.



واژه‌های کلیدی: اکتشاف نفت و گاز، یادگیری ماشین، پیش‌بینی خواص پتروفیزیکی، طبقه‌بندی فاسیس و لیتوفاسیس، پردازش داده‌های لرزه‌ای.

۱. مقدمه

صنعت نفت و گاز یک بخش پیچیده است که بسیاری از فعالیت‌های پیچیده را در زنجیره ارزش خود ترکیب می‌کند که به طور کلی به سه بخش بالادستی (Upstream)، میانی (Midstream) و پایین‌دستی (Downstream) تقسیم می‌شود، همانطور که در شکل ۱ نشان داده شده است. در هر عملیات صنعتی، حجم بی‌سابقه‌ای از داده‌ها می‌تواند از تجهیزات و گزارش‌های انسانی تولید شود. بخش بالادستی، که مربوط به اکتشاف و تولید نفت و گاز طبیعی است، داده‌هایی مانند بررسی‌های زمین‌شناسی، لاگ‌های چاه و قرائت‌های تجهیزات حفاری تولید می‌کند. این بخش همچنین انتظار می‌رود که با بهبود دستگاه‌های جمع‌آوری داده‌های لرزه‌ای، شمارش کانال‌ها و مانیتورینگ جبهه سیال، حجم داده‌های بسیار بیشتری تولید کند [۱]. بخش میانی شامل حمل و نقل و ذخیره نفت خام و گاز طبیعی با استفاده از خطوط لوله و زیرساخت‌های مرتبط مانند ایستگاه‌های پمپاژ و تانکرها است. همه این‌ها امکان تولید حجم زیادی از داده‌ها را فراهم می‌کنند. بخش پایین‌دستی شامل تبدیل نفت خام و گاز طبیعی به محصولات نهایی و بازاریابی آنها است. این بخش شامل تولید و تحلیل حجم زیادی از داده‌ها برای کسب مزیت رقابتی و کاهش هزینه‌ها است. حجم داده‌های تولید شده در صنعت نفت و گاز آنقدر زیاد است که حتی ثبت و ذخیره‌سازی آن نیز نیاز به تکنیک‌های پیچیده‌ای دارد.

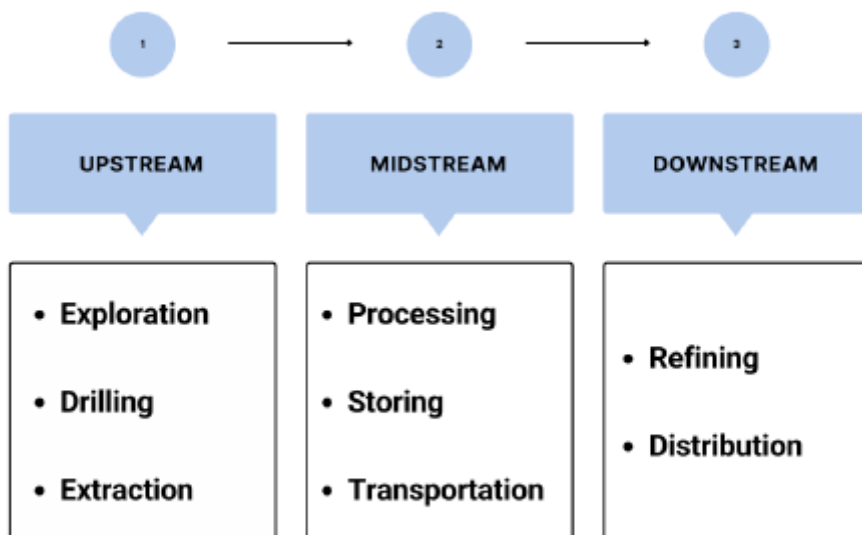


FIGURE 1. Oil and gas production process.

اکتشاف نفت و گاز عملی است که تلاش می‌کند تا تجمعات نفت و گاز طبیعی را که در زیر سطح جو زمین به دام افتاده‌اند، با استفاده از زمین‌شناسی نفت شناسایی کند. اکتشاف جهت ارائه دانش لازم برای استفاده از بهترین فرصت‌های ارائه شده توسط مناطق انتخاب شده برای اکتشاف و نظارت بر عملیات تحقیقاتی در بلوک‌های به دست آمده انجام می‌شود. اکتشاف ریسک‌های ذاتی در این فرآیند را کنترل می‌کند و معمولاً با انتخاب گزینه‌های احتمالی و اقتصادی‌تر مدیریت می‌شود.

روش‌های معمول در اکتشاف نفت و گاز برای شناسایی، ارزیابی و بهره‌برداری از منابع هیدروکربنی استفاده می‌شود. شناسایی و به دست آوردن مکان‌های امیدوارکننده اولین مرحله است، که ممکن است شامل مطالعه اطلاعات زمین‌شناسی و انجام بررسی‌های هوایی برای شناسایی مناطق با احتمال بالای وجود منابع هیدروکربنی باشد. هنگامی که یک مکان امیدوارکننده شناسایی شد، باید یک بررسی زمین‌شناسی برای درک بهتر زمین‌شناسی و پتانسیل هیدروکربنی منطقه انجام شود. این کار معمولاً با استفاده از روش‌های مختلفی مانند بررسی‌های الکترومغناطیسی، مغناطیسی، لرزه‌ای و گرانشی انجام می‌شود. بررسی‌های لرزه‌ای یکی از مهم‌ترین روش‌های مورد استفاده برای جستجوی نفت و گاز است. این روش‌ها با ارسال امواج صوتی به زمین و ثبت و تحلیل بازتاب‌های آنها کار می‌کنند. داده‌های جمع‌آوری شده می‌توانند اطلاعات گسترده‌ای در مورد زمین‌شناسی زیرین ارائه دهند و به شناسایی منابع احتمالی هیدروکربن کمک کنند.

پس از شناسایی یک مخزن احتمالی، حفاری اکتشافی ممکن است برای ارزیابی وجود، کیفیت و مقدار هیدروکربن‌های موجود در مخزن انجام شود. حفاری یک یا چند چاه اکتشافی برای جمع‌آوری نمونه‌های هسته‌ای، نمونه‌های سیال و سایر داده‌هایی که ممکن است برای شناسایی خواص مخزن مورد مطالعه قرار گیرند، معمول است. اگر هیدروکربن‌ها شناسایی شوند، مرحله بعدی ارزیابی امکان‌سنجی اقتصادی پروژه است. این شامل تعیین اندازه و بهره‌وری مخزن و هزینه‌های حفاری، تولید و حمل و نقل است. اگر این ایده از نظر تجاری امکان‌پذیر باشد، میدان توسعه می‌یابد و تولید آغاز می‌شود. ساخت تأسیسات تولیدی، حفاری چاه‌های تولیدی و استفاده از فناوری‌ها و روش‌های مختلف برای بهبود خروجی و بازیابی حداکثری انتظار می‌رود.

به طور کلی، مراحل درگیر در اکتشاف نفت و گاز پیچیده است و به مهارت‌ها و منابع فنی مختلفی نیاز دارد. از طرف دیگر، یک عملیات اکتشاف موفق ممکن است منجر به کشف ذخایر قابل توجه هیدروکربنی شود که می‌تواند منابع انرژی قابل توجهی برای بشریت فراهم کند. این مراحل در شکل ۲ نشان داده شده‌اند.



FIGURE 2. Stages in oil and gas exploration.

این ترجمه شامل دو صفحه اول مقاله است که به معرفی کلی موضوع، چالش‌ها و اهداف مقاله می‌پردازد.

اکتشاف زمین‌شناسی و ژئوفیزیکی

اکتشاف زمین‌شناسی و ژئوفیزیکی با استفاده از تکنیک‌های سطحی برای ارزیابی ویژگی‌های فیزیکی زمین زیرین انجام می‌شود. این روش‌ها با بررسی تغییرات در این ویژگی‌ها، وجود و محل هیدروکربن‌ها (نفت و گاز) را در مقادیر اقتصادی شناسایی یا استنباط می‌کنند. این کار با استفاده از روش‌های فیزیکی مانند لرزه‌ای، الکتریکی، نمونه‌برداری از هسته و لاگ‌های چاه انجام می‌شود تا ویژگی‌های فیزیکی سنگ‌ها ارزیابی شوند و به طور خاص، تفاوت‌های فیزیکی قابل اندازه‌گیری بین سنگ‌های حاوی هیدروکربن و سنگ‌های فاقد هیدروکربن شناسایی شوند. این اطلاعات در قرارگیری سازه‌های دریایی و تصمیم‌گیری‌های استراتژیک و اقتصادی در عملیات نفت و گاز مفید هستند.

A. پردازش داده‌های لرزه‌ای

روش اصلی ژئوفیزیکی مورد استفاده برای نقشه‌برداری از ویژگی‌های زمین‌شناسی زیر سطح زمین، چه در خشکی و چه در محیط‌های دریایی، داده‌های لرزه‌ای است. به طور ذاتی، فرآیندهای تفسیر انسانی کند، پرهزینه و غیرقابل تکرار هستند. یکی از زمان‌برترین فعالیت‌ها، تفسیر حجم زیادی از داده‌های لرزه‌ای است. به دلیل حجم زیاد، مجموعه‌های داده‌های لرزه‌ای برای الگوریتم‌های پیشرفته یادگیری ماشین مانند شبکه‌های عصبی کانولوشنال (CNN) مناسب هستند، که باید با داده‌های کافی آموزش ببینند تا به طور کارآمد و دقیق عمل کنند. چندین مشکل زمین‌شناسی پیچیده، از جمله شناسایی گسل‌ها، شناسایی بدنه‌های نمکی، نقاط شیرین (Sweet Spots) و افق‌های لرزه‌ای، با استفاده از یادگیری ماشین و داده‌های لرزه‌ای حل شده‌اند. علاوه بر این، اگرچه انسان‌ها در کشف ویژگی‌ها به طور انحصاری در دو بعد مهارت دارند، الگوریتم‌های به‌خوبی نوشته شده می‌توانند در همه ابعاد عمل کنند. کاربرد تکنیک شبکه عصبی مصنوعی (ANN) در حوزه اکتشاف نتایج مثبتی در کاهش ریسک‌های اکتشاف و افزایش کارایی چاه‌های اکتشافی داشته است [۶].

شکست‌های ساختمانی ممکن است به دلیل انواع حرکات زیرسطحی ایجاد شوند که می‌توانند منجر به تشکیل گسل‌ها شوند. پس از در نظر گرفتن وجود گسل‌ها در منطقه مورد علاقه، باید تصمیم‌گیری‌های خاصی در مورد عملیات انجام شود. در فرآیندهای سنتی، تفسیر گسل‌ها فرآیندی زمان‌بر است. از این رو، گیتون و همکاران [۷] از تکنیک ماشین بردار پشتیبان (SVM) برای تشخیص گسل‌ها در بخش‌های لرزه‌ای استفاده کردند. از بخش‌های لرزه‌ای برچسب‌گذاری شده، نویسندگان از تبدیل ویژگی‌های ناوردای مقیاس (SIFT) و هیستوگرام گرادیان‌های جهت‌دار (HOG) برای استخراج مجموعه‌ای از ویژگی‌ها استفاده کردند که برای آموزش SVM به منظور شناسایی گسل‌ها استفاده شد. مزیت ترکیب ویژگی‌های HOG و SIFT این است که از استفاده جداگانه آن‌ها بهتر عمل می‌کند. با این حال، شیونگ و همکاران [۸] ضعف SVM را در نیاز به پیش‌محاسبه ویژگی‌ها برای نقشه‌برداری از گسل‌ها نشان دادند. یک فرآیند پرزحمت نقشه‌برداری دستی از گسل‌ها باید برای هر مجموعه داده در مجموعه داده‌های آموزشی انجام شود. علاوه بر این، این تکنیک در مناطق با بازتاب‌های ضعیف عملکرد ضعیفی دارد. از این رو، برتری CNN در [۹]، [۱۰] و [۸] نشان داده شد. شیونگ و همکاران [۸] با استفاده از داده‌های لرزه‌ای سه‌بعدی، روش CNN را برای شناسایی و نقشه‌برداری خودکار از مناطق گسلی ارائه دادند که نیاز به پیش‌محاسبه انسانی را از بین می‌برد.

یانگ و سان [۱۱] یک تکنیک برای ردیابی افق‌ها با ویژگی‌های بازتاب لرزه‌ای پیچیده با استفاده از شبکه‌های عصبی کانولوشنال عمیق (Deep CNN) پیشنهاد کردند. این روش پیشنهادی می‌تواند محل گسل‌ها را تعیین کند و افق‌هایی را که از گسل‌ها عبور می‌کنند، به دقت استخراج کند. این مدل پیشنهادی نسبت به تکنیک سنتی ردیابی افق‌های سه‌بعدی سازگارتر و سریع‌تر بود. روش مبتنی بر CNN پتانسیل قابل توجهی برای بهبود کارایی و دقت ردیابی افق‌ها نشان داد.

برای یک مطالعه توموگرافی موج کامل لرزه‌ای، دی‌رسن و همکاران [۱۲] یک شبکه عصبی مصنوعی (ANN) و شبکه عصبی با اهمیت کمکی (IANN) پیشنهاد کردند. مدل‌های پیشنهادی یادگیری ماشین و تبدیل موجک پیچیده (CWT) را ترکیب می‌کنند که برای بهبود دقت طبقه‌بندی و سرعت بخشیدن به محاسبه تطابق بخش‌های موج داده‌های مشاهده‌شده و داده‌های سنتز شده امیدوارکننده است. هر دو مدل ANN و IANN نتایج مثبتی نشان دادند، با این حال IANN عملکرد کمی بهتر داشت.

با استفاده از چندین مدل شبکه عصبی، رستگاریا و همکاران [۱۳] حجم زیادی از داده‌های لرزه‌ای سه‌بعدی را پردازش کردند تا حجم الکتروفاسیس و شاخص منطقه جریان سه‌بعدی (FZI) را به دست آورند. نویسندگان یک شبکه عصبی احتمالی (PNN) برای مدل الکتروفاسیس پیشنهاد کردند که از خوشه‌بندی مبتنی بر گراف چندمقیاسی (MRGC) به عنوان بهینه‌ساز استفاده می‌کند. از طرف دیگر، مدل سه‌بعدی FZI از یک روش چندویژگی با استفاده از شبکه تابع پایه شعاعی (RBF)، شبکه پیش‌خور چندلایه (MLFFN) و PNN برای بهبود مدل استفاده کرد. بر اساس نتایج، این دو مدل با یکدیگر توافق خوبی دارند و مدل‌های مبتنی بر PNN می‌توانند برای تخمین FZI و حجم الکتروفاسیس به طور موثر استفاده شوند.

در مطالعه خود، لا مارکا و همکاران [۱۴] یک روش ارزیابی کمی جدید برای الگوریتم‌های یادگیری ماشین بدون نظارت معرفی کردند که از تکنیک‌هایی مانند K-means، نقشه‌های توپوگرافی تولیدی (GTM) و تحلیل مؤلفه‌های اصلی (PCA) در تفسیر لرزه‌ای استفاده می‌کند. روش‌شناسی آن‌ها با استفاده از داده‌های لرزه‌ای مصنوعی چندبعدی نشان داده شد که داده‌ها را به گروه‌های زمین‌شناسی معنادار خوشه‌بندی می‌کند. یادگیری ماشین دامنه ویژگی‌های تحلیل شده را گسترش می‌دهد و جزئیات پیچیده‌ای را آشکار می‌کند که اغلب توسط مفسران انسانی نادیده گرفته می‌شوند. قابل توجه است که الگوریتم‌های یادگیری ماشین معمولاً با استفاده از لاگ‌های چاه کالیبره می‌شوند؛ با این حال، تخصص انسانی نیز نقش محوری در فرآیند تفسیر دارد.

مطالعه دیگری توسط کیان و همکاران [۱۵] یک روش ماشین بردار پشتیبان را با ترکیب داده‌های زمین‌شناسی، حفاری چاه، لاگ‌های چاه و بررسی‌های لرزه‌ای توسعه دادند تا یک تخمین چندویژگی از نقاط شیرین مخزن و یک توصیف کمی جامع از مخازن شیل ارائه دهند. این تکنیک با ارائه یک روش ارزیابی کمی کارآمد و دقیق برای ارزیابی مخازن شیل، از روش‌های سنتی برتر بود.

این ترجمه ادامه متن مقاله را تا بخش پردازش داده‌های لرزه‌ای و طبقه‌بندی فاسیس‌ها پوشش می‌دهد.

B. طبقه‌بندی فاسیس و لیتوفاسیس

فاسیس‌ها یک ویژگی زمین‌شناسی پایه‌ای هستند که بر تولید هیدروکربن‌ها تأثیر می‌گذارند، بنابراین درک سنگ‌های فاسیس در اکتشاف نفت و گاز حیاتی است [۱۶]. داده‌های هسته‌ای و لاگ‌های پیشرفته چاه می‌توانند این اطلاعات را ارائه دهند، اما دسترسی به این نوع داده‌های غنی به دلیل هزینه و زمان مورد نیاز برای جمع‌آوری آن‌ها محدود است. بسیاری از تکنیک‌های یادگیری ماشین کم‌هزینه که از داده‌های لاگ‌های چاه ارزان‌قیمت استفاده می‌کنند، برای تحقیقات زیرسطحی پیشنهاد شده‌اند. مزیت این داده‌های لاگ‌های چاه شامل دسترسی پیوسته در طول عمق و جمع‌آوری آسان داده‌ها است. در نتیجه، آن‌ها منبع ارزشمندی از اطلاعات درباره سنگ‌های زیرسطحی هستند. لیتوفاسیس‌ها معمولاً با ترکیب ویژگی‌های پتروفیزیکی و زمین‌شناسی تعیین می‌شوند و می‌توانند ابزاری ضروری برای توصیف مخزن باشند [۱۷]. از زمان ظهور لاگ‌های چاه، چندین روش ریاضی برای پیش‌بینی لیتولوژی با استفاده از لاگ‌های چاه توسعه یافته‌اند [۱۸]. طبقه‌بندی لیتوفاسیس‌ها معمولاً با استفاده از نمونه‌های هسته‌ای و داده‌های لاگ‌های چاه و با کمک یادگیری ماشین انجام می‌شود.

اخيراً از یادگیری ماشین برای کمک به ارزیابی زمان‌بر لاگ‌های چاه برای طبقه‌بندی لیتوفاسیس‌ها استفاده شده است. این روش می‌تواند پس از آموزش با استفاده از چاه‌های هسته‌ای در منطقه، لیتوفاسیس‌ها را در چاه‌های بدون هسته طبقه‌بندی کند. برای آموزش مدل، طبقه‌بندی لیتوفاسیس‌ها بر اساس اندازه‌گیری عمق با استفاده از ترکیب داده‌های لاگ‌های چاه و داده‌های هسته‌ای انجام می‌شود [۱۹]. لاگ‌های اشعه گاما (GR)، مقاومت (Rt)، نوترون (NPHI)، چگالی (RHOB) و لیتولوژی از رایج‌ترین لاگ‌ها برای شناسایی فاسیس‌ها هستند. این لاگ‌ها امکان تولید ویژگی‌های پیچیده‌ای را فراهم می‌کنند که می‌توانند پیش‌بینی‌ها را بهبود بخشند، از جمله کل مواد آلی (TOC) و چگالی دانه‌های ماتریس [۲۰] (RHOMAA).

محققان از تکنیک‌های مختلف یادگیری ماشین مانند شبکه‌های عصبی (NN)، ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) استفاده کرده‌اند [۱۷]، [۱۹]، [۲۱]، [۲۲]، [۲۳]، [۲۴]، [۲۵]، [۲۶]، [۲۷]. آن‌ها لیتوفاسیس‌ها را از لاگ‌های چاه در انواع مختلف مخازن شناسایی کردند و بر اساس یافته‌های خود نتیجه گرفتند که این استراتژی‌ها مؤثر و خودکار هستند و به زمان و منابع کمتری نسبت به روش‌های سنتی نیاز دارند.

چندین محقق [۱۷]، [۱۹]، [۲۱]، [۲۲] تحقیقات خود را با استفاده از NN برای طبقه‌بندی لیتوفاسیس‌ها انجام دادند. آن‌ها نتیجه گرفتند که این روش از روش‌های سنتی بهتر عمل می‌کند. اما هنگامی که با حجم کمی از داده‌ها سروکار داریم، سبتوشیخ و صالحی [۲۳] اشاره کردند که SVM عملکرد بهتری دارد. با این حال، شی و همکاران [۱۶] تحقیقات خود را با استفاده از هر دو روش NN و SVM انجام دادند و نتیجه گرفتند که هر دو روش تحت تأثیر تعداد ویژگی‌های موجود قرار می‌گیرند و هنگامی که ویژگی‌های محدودی استفاده می‌شوند، با مشکل مواجه می‌شوند. این کار نشان داد که روش‌های ترکیبی (Ensemble Methods) برتر هستند. روش‌های ترکیبی چندین مدل پایه را برای ایجاد



یک مدل پیش‌بینی واحد ترکیب می‌کنند [۲۵]. دل‌آورسانا و همچنین تیواری و دوییدی [۲۴]، [۲۶] نیز از روش‌های ترکیبی به عنوان روش‌های قوی‌تر، قابل اعتمادتر و دقیق‌تر حمایت کردند. در تحقیقی دیگر، هو و همکاران [۲۸] مدل‌های چندلایه پرسپترون (MLP)، SVM و روش‌های ترکیبی مانند eXtreme Gradient Boosting (XGBoost) و RF را برای طبقه‌بندی لیتوفاسیس‌ها در شیل گولونگ مقایسه کردند. بر اساس عملکرد مدل‌ها، می‌توان نتیجه گرفت که روش‌های ترکیبی دقت بیشتری دارند.

با این حال، تحقیقات جدید نشان داده‌اند که روش Gradient Boosting (GB) به دلیل مقاومت بالاتر، از سایر الگوریتم‌های یادگیری ماشین بهتر عمل می‌کند [۱۹]. اما هنگامی که با حجم زیادی از داده‌ها سروکار داریم، RF عملکرد بهتری دارد. مطالعات باتاچاریا و میشر [۲۷] نیز برتری RF نسبت به GB را نشان دادند، زیرا زمان محاسباتی را در مرحله آموزش کاهش می‌دهد.

در یک مطالعه تطبیقی که توسط المضحفر و همکاران [۲۹] انجام شد، چندین الگوریتم boosting برای طبقه‌بندی لیتوفاسیس‌ها در یک مخزن کربناتی عراق ارزیابی شدند. این مطالعه عملکرد چندین الگوریتم boosting از جمله Generalized Boosting Modelling، Logistic Boosting Regression (LogitBoost)، XGBoost، GBM، Adaptive Boosting Model (AdaBoost) و K-nearest neighbour (KNN) را با استفاده از داده‌های لاگ‌های چاه و داده‌های هسته‌ای بررسی کرد. در بین این الگوریتم‌ها، XGBoost بالاترین سطح دقت را در طبقه‌بندی لیتوفاسیس‌ها نشان داد.

در مطالعه دیگری توسط کیم [۳۰]، یک روش پیشگامانه برای طبقه‌بندی لیتوفاسیس‌ها در سازندهای چالش‌برانگیز آستین چاک و ایگل فورد، که به دلیل کیفیت پایین مخزن شناخته شده‌اند، پیشنهاد شد. محققان از یک شبکه عصبی کانولوشنال (CNN) برای انجام این کار طبقه‌بندی با استفاده از لاگ‌های چاه معمولی استفاده کردند و جالب اینجاست که مدل CNN از مدل سنتی ANN بهتر عمل کرد. این تحقیق اهمیت استفاده از روش‌های پیشرفته مانند CNN را برای افزایش دقت طبقه‌بندی لیتوفاسیس‌ها برجسته می‌کند. مزیت اضافی مدل CNN این است که وابستگی کمتری به لاگ‌های تفسیرشده مانند تخلخل، اشباع و شکنندگی دارد، که عدم قطعیت‌های ناشی از تفسیرهای ذهنی دستی را کاهش می‌دهد.

برای مقابله با پیچیدگی‌های طبقه‌بندی لیتوفاسیس‌ها در یک محیط زیرسطحی پویا، داتا و همکاران [۳۱] یک روش جدید معرفی کردند. این روش یک فرآیند چندمرحله‌ای تشخیص تغییرات را دنبال می‌کند. این فرآیند با تشخیص تغییرات قابل توجه در لاگ‌های چاه آغاز می‌شود، سپس این تغییرات با دسته‌های لیتوفاسیس‌ها هماهنگ می‌شوند، پس از آن مجموعه داده با مدیریت کلاس‌های بیش‌نمایش بهینه می‌شود و در نهایت از SVM برای طبقه‌بندی استفاده می‌شود. این روش به طور چشمگیری از الگوریتم سنتی SVM بهتر عمل کرد.

۱۷. توصیف مخزن

فرآیند تعیین عینی ویژگی‌های مخزن بر اساس اطلاعات زمین‌شناسی و شناسایی عدم قطعیت‌ها در تغییرات جغرافیایی، به عنوان توصیف مخزن شناخته می‌شود [۳۲].

به طور کلی، توصیف مخزن در مرحله اکتشاف انجام می‌شود تا محل و میزان احتمالی ذخایر نفتی ارزیابی شود. پس از تعیین محل و مقدار هیدروکربن‌های موجود در مخزن، می‌توان میدان نفتی را برای استخراج این ذخایر توسعه داد. حفاری اکتشافی معمولاً در چندین چاه جداگانه در مرحله اول این فرآیند انجام می‌شود. هدف هر چاه ارائه اطلاعاتی درباره ویژگی‌های سازند سنگی اطراف چاه و انواع ذخایر هیدروکربنی است که ممکن است در آنجا وجود داشته باشد.

هدف از توصیف مخزن، کسب دانش عمیق‌تر درباره ویژگی‌های فیزیکی و شیمیایی مخازن است تا تصمیم‌گیری‌های آگاهانه‌تری درباره توسعه و بهره‌برداری از آن‌ها انجام شود، که بر سودآوری عملیات نفتی و تأثیرات محیطی آن‌ها تأثیر می‌گذارد. این کار به تعیین بهترین روش‌های تولید برای حداکثر کردن خروجی کمک می‌کند و نشان می‌دهد که رفتار سیالات مخزن تحت شرایط مختلف چگونه تغییر می‌کند. هدف توصیف مخزن، ایجاد یک مدل زمین‌شناسی است که از داده‌های موجود برای پیش‌بینی خواص پتروفیزیکی در سرتاسر میدان نفتی استفاده می‌کند [۳۳]. توسعه یک تصویر دقیق از ویژگی‌های یک مخزن می‌تواند چالش‌برانگیز و زمان‌بر باشد. در نتیجه، نیاز مداوم به بهبود روش‌های خودکار توصیف مخزن وجود دارد.

۱۷. پیش‌بینی خواص پتروفیزیکی

جمع‌آوری داده‌های دقیق درباره ویژگی‌های مخزن برای توصیف مخزن ضروری است. هدف اصلی توصیف مخزن، توسعه نمایش‌های سه‌بعدی از خواص پتروفیزیکی است. این کار شامل جمع‌آوری داده‌ها درباره ویژگی‌های پتروفیزیکی است که بیش از بقیه درباره تجمع سیالات درون سازند سنگی ارائه می‌دهد. دقیق‌ترین روش برای تخمین خواص پتروفیزیکی، روش آزمایشگاهی است؛ اما این روش پرهزینه و زمان‌بر است. به دلیل این محدودیت، تنها تعداد محدودی نمونه برای برخی چاه‌ها در دسترس است و این نمونه‌ها تنها بخش‌هایی از بازه‌های عمق را پوشش می‌دهند [۳۴]. تعداد قابل توجهی نمونه برای تعریف دقیق یک سازند زیرسطحی به دلیل رفتارهای زمین‌شناسی پیچیده و ناهمگنی فضایی مخازن مورد نیاز است. روش‌های مبتنی بر لاگ‌های چاه به طور گسترده‌ای برای حل این مشکل استفاده شده‌اند.

نمونه‌های واقعی سنگ‌ها در آزمایشگاه بررسی شدند و از روش‌های ابزاری که ویژگی‌های فیزیکی را اندازه‌گیری می‌کنند، به عنوان منابع داده برای پارامترهای پتروفیزیکی استفاده شد [۳۵]. این داده‌ها شامل داده‌های هسته‌ای، لرزه‌ای



و لاگ‌های چاه بودند. بر اساس گفته‌های شو و همکاران [۳۶]، داده‌های پتروفیزیکی را می‌توان به عنوان داده‌های بزرگ در نظر گرفت، زیرا ویژگی‌های آن‌ها را دارا هستند. جدول ۱ داده‌های پتروفیزیکی را نشان می‌دهد. یادگیری ماشین به طور گسترده‌ای برای پیش‌بینی خواص پتروفیزیکی مانند تخلخل، نفوذپذیری، فشار موینگی و اشباع آب استفاده شده است. یادگیری ماشین نیاز به پردازش انسانی و پیچیدگی‌های زمین‌شناسی که روش‌های سنتی با آن‌ها مواجه هستند را از بین می‌برد و در عین حال زمان پردازش را به طور قابل توجهی کاهش می‌دهد، در حالی که کیفیت و ثبات خروجی را حفظ می‌کند [۳۷].

پارامترهای زیرسطحی مهمی باید شناسایی یا ارزیابی شوند؛ اما مهم‌ترین عوامل نفوذپذیری و تخلخل هستند. این پارامترها شاخص‌های کلیدی کیفیت و امکان‌پذیری اقتصادی مخازن نفتی هستند. تخلخل، که اندازه‌گیری نسبت فضاهای خالی یا منافذ در یک سنگ است، یک عامل حیاتی برای تخمین مقدار احتمالی هیدروکربن‌های موجود در یک مخزن است. این فضاهای خالی می‌توانند به عنوان مناطق ذخیره‌سازی هیدروکربن‌ها عمل کنند. از طرف دیگر، نفوذپذیری یک عامل مهم در توصیف چگونگی اتصال فضاهای خالی مجزا در یک سنگ است. نفوذپذیری توانایی هیدروکربن‌ها برای جریان یافتن از طریق منافذ به سمت سطح، جایی که می‌توانند استخراج شوند، را اندازه‌گیری می‌کند. بدون یک تخمین دقیق از نفوذپذیری، حل بسیاری از مسائل مهندسی نفت غیرممکن است.

A. نفوذپذیری

مطالعه‌ای توسط هوانگ و همکاران [۳۸] کاربرد یک شبکه عصبی مصنوعی (ANN) را برای پیش‌بینی نفوذپذیری در یک میدان گازی دریایی در شرق کانادا بررسی کرد. نویسندگان یک ANN با انتشار برگشتی (Back Propagation) را با استفاده از داده‌های لاگ‌های چاه از شش چاه پیشنهاد کردند. مدل پیشنهادی از روش‌های سنتی مانند رگرسیون خطی چندگانه (MLR) و رگرسیون غیرخطی چندگانه (MNL) بهتر عمل کرد.

TABLE 1. Common Petrophysical data and their attributes.

Data	Source	Type	Depth-index	Dimension
Routine Core Analysis	Core	Numerical	Discrete	Low
X-Ray Diffraction (XRD) Mineralogy	Core	Numerical	Discrete	High
SEM	Core	Image	Discrete	High
Thin Sections	Core	Image	Discrete	High
Core Photo	Core	Image	Discrete	High
Capillary Pressure	Core	Array	Discrete	High
Electrical properties	Core	Numerical	Discrete	High
Relative Permeability	Core	Array	Discrete	High
Facies Description	Core	Text	Discrete	High
Mud logging	Mud Log	Numerical and Text	Continuous	Low
Conventional Logs	Log	Numerical	Continuous	Low
Dielectric Log	Log	Numerical	Continuous	Low
Nuclear Spectroscopy	Log	Array	Continuous	High
NMR Log	Log	Waveforms/maps	Continuous	High
Image Log	Log	Image	Continuous	High
Sonic Log	Log	Waveforms/maps	Continuous	High
Formation Testing (pressure build-up/drawdown)	Log	Array (pressure vs. time)	Discrete	High
Pressure Transient	Production	3-D	Discrete/Time-index	High
Seismic Attributes	Seismic	3-D	Continuous	High

به طور مشابه، هله و همکاران [۳۹] این یافته را با پیش‌بینی تخلخل و نفوذپذیری مخزن دریای شمال تأیید کردند. این مدل نیز از روش‌های سنتی بهتر عمل کرد. به همین ترتیب، سینگ [۴۰] از ANN برای تخمین نفوذپذیری از لاگ‌های چاه معمولی استفاده کرد. نویسندگان توانایی این تکنیک را در ایجاد تطابق ثابت بین خروجی‌های پیش‌بینی شده و واقعی برجسته کردند. مطالعه‌ای توسط عبدیده [۴۱] نفوذپذیری را در یک میدان نفتی در ایران با استفاده از تکنیک ANN انتشار برگشتی پیش‌خور پیش‌بینی کرد. با استفاده از لاگ‌های چاه برای پیش‌بینی، این تکنیک از نظر دقت پیش‌بینی نسبت به MLR برتری داشت. مدل ANN در بن‌عونه و پادمانابان [۴۲] برای پیش‌بینی نفوذپذیری یک مخزن ماسه‌سنگی توسعه داده شد. با این حال، تنها تخلخل به عنوان ورودی مدل استفاده شد. با استفاده از تنها سه ویژگی لاگ‌های چاه: شاخص تحرک، تخلخل نوترون و چگالی توده، الکتاری و همکاران [۴۳] یک فرمول تجربی از ANN برای تخمین نفوذپذیری در یک مخزن کربناتی ناهمگن ساختند. مدل ANN پیشنهادی دقت کمی کمتر از سیستم استنتاج فازی-عصبی تطبیقی (ANFIS) داشت، اما از SVM بهتر بود. با این حال، مدل یک فرمول تجربی ارائه داد. با این حال، باشوگ و کارپین [۴۴] رابطه بین نفوذپذیری و تخلخل، سطح ویژه و اشباع آب کاهش‌ناپذیر را بررسی کردند. نویسندگان استفاده از مدل ANN را برای پیش‌بینی نفوذپذیری پیشنهاد کردند. این روش سطح قابل قبولی از دقت را نشان داد. مطالعه‌ای توسط ایرانی و نصیمی ANN [۴۵] تکاملی را برای پیش‌بینی نفوذپذیری معرفی کرد. این مدل از یک الگوریتم ژنتیک (GA) به عنوان بهینه‌ساز در ANN برای جستجوی پارامترهای بهینه شبکه استفاده کرد. نویسندگان خاطرنشان کردند که مدل پیشنهادی دقت بالاتری نسبت به ANN سنتی ارائه داد. پس از اعمال تحلیل مؤلفه‌های اصلی (PCA) برای استخراج ویژگی‌های مرتبط از لاگ‌های چاه، باقری‌پور [۴۶] یک مدل ترکیبی (CM) متشکل از MLP، RBF و شبکه عصبی رگرسیون عمومی (GRNN) را با استفاده از GA برای



پیش‌بینی نفوذپذیری ساخت. مدل CM پیشنهادی دقت بهتری نسبت به روش‌های فردی ارائه داد. علاوه بر GA، ماتینکیا و همکاران [۴۷] الگوریتم بهینه‌سازی ازدحام ذرات (PSO) و الگوریتم Social Ski-Driver (SSD) را برای پیش‌بینی نفوذپذیری با استفاده از MLP در سازند فهلپان چاهی بررسی کردند. هیبرید MLP-SSD پس از حذف داده‌های پرت و انتخاب ویژگی با استفاده از توضیحات افزودنی شیلی (SHAP) بهترین دقت را ارائه داد. همچنین، ژائو و همکاران [۴۸] از SHAP برای تجسم و تفسیر پیش‌بینی‌های خود با استفاده از الگوریتم‌های LR، SVM، RF، BPNN و XGBoost استفاده کردند. با این حال، XGBoost دقیق‌ترین نتایج را در پیش‌بینی‌های آن‌ها ارائه داد. به طور مشابه، لیو و لیو [۴۹] نفوذپذیری را در حوضه اردوس با استفاده از هیبرید PSO و XGBoost پیش‌بینی کردند. نویسندگان همچنین از SHAP برای انتخاب ویژگی و تفسیر استفاده کردند تا مدل را قابل تفسیرتر کنند. مدل پیشنهادی از CNN، حافظه کوتاه‌مدت بلندمدت (LSTM) و واحد بازگشتی دروازه‌دار (GRU) بهتر عمل کرد.

اگرچه ANN برای پیش‌بینی نفوذپذیری مؤثر نشان داده شده است، اما معایبی مانند همگرایی کند و گیر کردن در مینیمم‌های محلی دارد. مطالعه‌ای توسط طهماسی و هزارخانی [۵۰] یک شبکه عصبی مدولار (MNN) را برای پیش‌بینی نفوذپذیری بررسی کرد. مدل MNN شامل چندین شبکه عصبی به هم پیوسته است که به طور مؤثر یک مسئله بزرگ را به اجزای کوچکتر تقسیم می‌کند. این امر پیش‌بینی‌های سریع‌تر، ساده‌تر و دقیق‌تر را امکان‌پذیر می‌کند. مدل پیشنهادی از نظر دقت پیش‌بینی و عملکرد از شبکه عصبی سنتی بهتر عمل کرد. در تحقیقی که توسط جیمی‌الله‌مدی و جوادپور [۵۱] انجام شد، یک شبکه عصبی تابع پایه شعاعی (RBF) برای پیش‌بینی نفوذپذیری از تخلخل پیشنهاد شد. این مدل به دلیل تقریب جهانی و سرعت یادگیری بالاتر از شبکه‌های عصبی سنتی متمایز است. به طور مشابه، [۵۲] استفاده از GA را به عنوان بهینه‌ساز درون یک ANN برای تعیین بهترین پارامترها به منظور کاهش زمان و دستیابی به بالاترین عملکرد ممکن پیشنهاد کرد. این تکنیک برای پیش‌بینی نفوذپذیری به طور جداگانه در یک مخزن ایرانی بر اساس تقسیم‌بندی زمین‌شناسی استفاده شد. با این حال، آیف و همکاران [۵۳] کارایی مدل‌های هیبریدی را برای پیش‌بینی نفوذپذیری و تخلخل با استفاده از لاگ‌های چاه بررسی کردند. نویسندگان یک سیستم عصبی-فازی را پیشنهاد کردند که ANN و منطق فازی (FL) را ترکیب می‌کند تا مزایای هر دو روش را به دست آورد و از روش‌های فردی بهتر عمل کند. برای غلبه بر برخی محدودیت‌های ANN، سج‌جوهی و هزارخانی [۵۴] نظریه موجک را معرفی کردند. تکنیک پیشنهادی از انواع مختلف موجک به عنوان توابع فعال‌سازی برای تخمین نفوذپذیری استفاده کرد. این تکنیک از لاگ‌های چاه به عنوان ورودی استفاده کرد و برتری خود را نسبت به ANN سنتی نشان داد. در همین حال، بازیر و تادیونی [۵۵] عملکرد سیستم استنتاج عصبی-فازی هم‌فعال (CANFIS)، MLP و SVM را برای تخمین نفوذپذیری در یک مخزن ماسه‌سنگی فشرده مقایسه کردند. CANFIS بهترین دقت را به دست آورد، اما با هزینه سرعت محاسباتی کندتر. با استفاده از تنها تخلخل، سطح ویژه و اشباع آب کاهش‌ناپذیر، کمالی و همکاران [۵۶] استفاده از الگوریتم گروه‌بندی داده‌ها (GMDH) را برای پیش‌بینی نفوذپذیری در مخازن کربناتی روسیه و ایران

پیشنهاد کردند. الگوریتم پیشنهادی توانست نفوذپذیری را به دقت پیش‌بینی کند و از رگرسیون چندجمله‌ای، رگرسیون بردار پشتیبان (SVR) و درخت تصمیم (DT) بهتر عمل کرد.

مطالعه‌ای توسط حماده و الشافعی [۵۷] رزونانس مغناطیسی هسته‌ای (NMR) را برای تکمیل لاگ‌های چاه معمولی به منظور مقابله با ناهمگنی مخازن ماسه‌سنگی گازی معرفی کرد. NMR به دلیل ارائه تخلخل کمی مستقل از لیتولوژی و تخمین قابل اعتماد از پتانسیل هیدروکربنی مورد توجه قرار گرفته است. نویسندگان از ANN پیش‌خور برای پیش‌بینی تخلخل و نفوذپذیری یک مخزن ماسه‌سنگی گازی ناهمگن استفاده کردند. بر اساس یافته‌ها، پیش‌بینی‌ها با استفاده از NMR همراه با لاگ‌های چاه معمولی دقت بیشتری نسبت به پیش‌بینی‌هایی که تنها از لاگ‌های چاه معمولی استفاده می‌کردند، ارائه دادند.

برخی از نویسندگان تحقیقاتی را با استفاده از تکنیک‌های مختلف یادگیری ماشین انجام داده‌اند. مطالعه‌ای توسط السباعی و همکاران [۵۸] از تکنیک شبکه عملکردی (FN) برای پیش‌بینی نفوذپذیری در یک مخزن کربناتی استفاده کرد. با استفاده از یک پایه چندجمله‌ای، عملکرد پیش‌بینی‌کننده مدل FN همبستگی بهتری نسبت به ANN، ANFIS و رگرسیون آماری نشان داد و از معماری ساده مدل بهره برد. برعکس، اولاتونجی و همکاران [۵۹] ماشین‌های یادگیری سریع (Extreme Learning Machines) را برای پیش‌بینی نفوذپذیری در مخازن کربناتی خاورمیانه بررسی کردند. روش پیشنهادی از نظر عملکرد، دقت و سرعت یادگیری سریع‌تر از ANN و SVM برتر بود. از طرف دیگر، غلامی و همکاران [۶۰] رگرسیون بردار مرتبط (RVR) را برای پیش‌بینی نفوذپذیری در یک مخزن کربناتی با استفاده از GA به عنوان بهینه‌ساز بررسی کردند. هنگامی که دقت روش پیشنهادی با SVM مقایسه شد، مزیت کمی نشان داد. در مطالعه‌ای دیگر، عبدالرحیم و همکاران [۶۱] تکنیک FL را برای پیش‌بینی نفوذپذیری در یک مخزن کربناتی خاورمیانه بررسی کردند. نویسندگان کارایی خوشه‌بندی تفریقی را نسبت به تکنیک تقسیم‌بندی شبکه‌ای تأیید کردند. روش پیشنهادی تطابق عالی‌ای نشان داد و برای پیش‌بینی نفوذپذیری مؤثر بود. علاوه بر این، وانگ و همکاران [۶۲] FL را با استفاده از تکنیک مهندسی ویژگی Student-Newman-Keuls بهینه کردند. مدل پیشنهادی از تکنیک سستی بدون بهینه‌ساز بهتر عمل کرد.

در مطالعه‌ای توسط ژانگ و همکاران [۶۳] عملکرد MLP، SVR و MLR را در پیش‌بینی نفوذپذیری در یک مخزن ماسه‌سنگی گازی فشرده ناهمگن مقایسه کردند. تخلخل و لاگ‌های چاه به عنوان ورودی استفاده شدند. MLP و SVR دقت پیش‌بینی بالایی نشان دادند، با این حال SVR همبستگی کمی بالاتر و MLP خطای کمی کمتر داشت. در مطالعه‌ای دیگر، شیخ‌نسب و همکاران [۶۴] پیش‌بینی نفوذپذیری مخزن کربناتی را با استفاده از الگوریتم‌های ماشین

بردار پشتیبان حداقل مربعات (LSSVM) و ماشین یادگیری سریع چندلایه (MELM) پیشنهاد کردند. نویسندگان از الگوریتم بهینه‌سازی کوکوی (COA)، PSO و GA برای بهینه‌سازی مدل‌ها استفاده کردند. پس از استفاده از روش توکی برای حذف داده‌های پرت، هیبرید MELM و COA دقیق‌ترین نتایج را ارائه داد.

از طرف دیگر، آنیفووس و همکاران [۶۵] از یک پارادایم یادگیری ماشین ترکیبی برای غلبه بر فرضیه‌های تک‌گانه تکنیک‌های هوش محاسباتی (CI) سنتی و سیستم‌های هوشمند ترکیبی (HIS) و انتخاب پارامترهای مدل CI/HIS استفاده کردند. مطالعه‌ای توسط بات [۶۶] تلاش کرد تا تخلخل، نفوذپذیری، اشباع سیال و لیتوفاسیس‌ها را در میدان اوسبرگ با استفاده از تکنیک بگینگ (Bagging) ماشین‌های کمیته پیش‌بینی کند. نویسنده از لاگ‌های سیم‌لین و اندازه‌گیری در حین حفاری (MWD) برای پیش‌بینی بلادرنگ استفاده کرد. ماشین‌های کمیته عملکرد برتر نسبت به یک شبکه عصبی تک‌گانه نشان دادند.

با این حال، چن و لین [۶۷] از فرمول‌های تجربی معمولاً استفاده‌شده در توصیف مخزن برای ساخت یک مدل ترکیبی جدید برای محاسبه نفوذپذیری استفاده کردند. مدل ترکیبی از فرمول‌های تجربی ویلی و رز [۶۸]، کوتس و دومانویر [۶۹] و شلومبرگر [۷۰] برای تشکیل یک ماشین کمیته استفاده کرد. روش پیشنهادی پیش‌بینی‌های بسیار قابل‌اطمینان‌تری نسبت به روش‌های فردی ارائه داد و تعمیم‌پذیری بسیار بهتری داشت. با این حال، [۷۱] از فرمول‌های تجربی و رگرسیون چندگانه در ماشین کمیته خود استفاده کردند. ترکیب بهینه وزن‌ها با استفاده از GA تعیین شد. نویسندگان نفوذپذیری یک مخزن کربناتی در میدان نفتی بالال را با استفاده از داده‌های لاگ‌های چاه معمولی پیش‌بینی کردند. به طور مشابه، ماشین کمیته پیش‌بینی‌های دقیق‌تری نسبت به روش‌های فردی ارائه داد. در مدل ترکیبی هلمی [۷۲]، این مدل شامل SVM، ANN و ANFIS بود. نفوذپذیری در یک میدان نفتی در خاورمیانه با استفاده از لاگ‌های چاه پیش‌بینی شد. این نشان می‌دهد که مدل‌های ترکیبی ناهمگن ممکن است عملکرد بهتری نسبت به مدل‌های فردی داشته باشند، همانطور که در دقت و تعمیم‌پذیری مشاهده شد. از طرف دیگر، آنیفووس و همکاران [۷۳] از لاگ‌های چاه یک مخزن کربناتی خاورمیانه استفاده کردند و از سه الگوریتم انتخاب ویژگی برای پیش‌بینی نفوذپذیری استفاده کردند. SVM و منطق فازی نوع-۲ (T₂FL) با استفاده از استراتژی‌های انتخاب ویژگی FN، DT و آنترپی اطلاعات فازی (FIE) آموزش داده شدند. روش ترکیبی FN-SVM عملکرد بسیار خوبی نسبت به سایر مدل‌های ترکیبی و مستقل نشان داد. در مقابل، یک رویکرد نوآورانه ارائه‌شده توسط ماسرور و همکاران [۷۴] شبکه عصبی کانولوشنال عمیق باقیمانده چندورودی (MIRes CNN) را برای پیش‌بینی نفوذپذیری در میدان نفتی آزادگان ایران معرفی می‌کند. این تکنیک منحصر به فرد به طور همزمان از دو مجموعه داده متمایز استفاده می‌کند: لاگ‌های عددی چاه (NWLS) و تصاویر ویژگی‌های گرافیکی (GFIs). GFIs ها با تبدیل بردار یک‌بعدی NWLS به ماتریس‌های دوبعدی ایجاد شدند. در حالی که مجموعه داده‌های NWL توسط یک شبکه عصبی کانولوشنال عمیق باقیمانده یک‌بعدی تک‌ورودی (SIRes 1D-CNN) پردازش می‌شوند، GFIs ها توسط یک شبکه عصبی کانولوشنال عمیق

باقیمانده دوبعدی تک‌ورودی (SIRes ۲D-CNN) پردازش می‌شوند. تحلیل تطبیقی نشان داد که این رویکرد پیشنهادی از SIRes ۱D-CNN، SIRes ۲D-CNN، GMDH و RF بهتر عمل کرد.

با استفاده از شبکه عصبی پیش‌خور انتشار برگشتی، آنیفوووس و همکاران [۷۵، ۷۶] یک مدل ترکیبی برای پیش‌بینی نفوذپذیری و تخلخل تشکیل دادند. اساس این مدل ترکیبی، شبکه‌های عصبی با تعداد مختلف نورون‌های مخفی بهینه، با تعداد تصادفی نورون‌های مخفی و با الگوریتم‌های یادگیری مختلف است. مطالعه‌ای توسط آنیفوووس و همکاران [۷۷] یک مدل ترکیبی SVM را برای پیش‌بینی تخلخل و نفوذپذیری پیشنهاد کرد. مدل پیشنهادی پیش‌بینی‌هایی را بر اساس مقادیر مختلف پارامترهای تنظیم‌شده بهینه ارائه می‌دهد. مقایسه عملکرد مدل با SVM پیاده‌شده با استفاده از تکنیک بگینگ، یک SVM استاندارد و یک ترکیب از درخت‌های تصمیم، برتری مدل پیشنهادی را نشان داد. مطالعه‌ای توسط آنیفوووس و همکاران [۷۸] یک ترکیب از ماشین‌های یادگیری سریع (ELM) را برای پیش‌بینی تخلخل و نفوذپذیری پیشنهاد کرد. مدل پیشنهادی از تکنیک FN برای انتخاب ویژگی پیشرفته استفاده می‌کند، که آن را به یک مدل ترکیبی تبدیل می‌کند. عملکرد مدل از ELM سنتی و جنگل تصادفی بهتر بود. اوچره و همکاران [۷۹] یک مدل ترکیبی ساختند که از جنگل تصادفی و تکنیک انتخاب ویژگی تنظیم Lasso همراه با XGBoost برای پیش‌بینی دقیق اشباع آب و نفوذپذیری استفاده کرد. بر اساس نتایج، مدل ترکیبی پیشنهادی از مدل XGBoost سنتی و مدل ترکیبی که PCA و XGBoost را ادغام می‌کرد، بهتر عمل کرد.

اگرچه روش‌های جدید لاگ‌گیری چاه دقیق‌تر از روش‌های قدیمی هستند، محققان علاقه کمی به بهبود الگوریتم‌های خود نشان داده‌اند. اگرچه محققان علاقه کمی به توسعه الگوریتم‌های خود نشان داده‌اند، روش‌های مدرن لاگ‌گیری چاه دقت بیشتری نسبت به روش‌های سنتی نشان داده‌اند. بر اساس مرور ادبیات برای پیش‌بینی نفوذپذیری، جدول ۲ خلاصه‌ای جامع از رویکردهای مختلف یادگیری ماشین استفاده‌شده، پارامترهای ورودی و محل مخزن بررسی‌شده را ارائه می‌دهد.

B. تخلخل

محققان معمولاً از ANN برای پیش‌بینی تخلخل در سازندهای مختلف استفاده کرده‌اند [۳۹، ۵۷، ۶۶، ۸۰، ۸۱]، [۸۲، ۸۳، ۸۴]. هله و همکاران [۳۹] با استفاده از یک ANN انتشار برگشتی، تخلخل و نفوذپذیری را در دریای شمال پیش‌بینی کردند. نویسندگان با استفاده از چگالی، تخلخل نوترون، سرعت صوت و اشعه گاما توانستند تخلخل و نفوذپذیری را در مخازن ژوراسیک با دقت قابل قبولی پیش‌بینی کنند. یک مطالعه تطبیقی توسط کوناته و همکاران [۸۲] دو مدل ANN را برای پیش‌بینی نفوذپذیری در میدان نفتی ژنزیگ بررسی کردند. مدل‌های مورد بررسی شامل

شبکه عصبی رگرسیون عمومی (GRNN) و شبکه عصبی پیش‌خور انتشار برگشتی (FFBP) بودند. GRNN از نظر دقت پیش‌بینی برتری نشان داد. به طور مشابه، ژانگ و همکاران [۸۵] شبکه عصبی GRU را برای پیش‌بینی تخلخل بررسی کردند. GRU سریع‌تر است و به منابع محاسباتی کمتری برای پیش‌بینی نیاز دارد. مدل پیشنهادی شامل یک تابع کوپولا به عنوان تحلیل همبستگی (CA) برای انتخاب ویژگی بود. برتری این مدل نسبت به مدل‌های GRU مستقل، شبکه عصبی بازگشتی (RNN) و MLP نشان داده شد. در مطالعه دیگری، حماده و الشافعی [۵۷] مدلی را توسعه دادند که از لاگ‌های NMR برای تکمیل لاگ‌های چاه معمولی در ذخایر ماسه‌سنگی گازی استفاده می‌کند. این مطالعه نشان داد که پیش‌بینی‌های با استفاده از NMR همراه با لاگ‌های چاه معمولی دقت بیشتری نسبت به پیش‌بینی‌هایی که تنها از لاگ‌های چاه معمولی استفاده می‌کنند، دارند.

محققان با موفقیت ANN را با سایر روش‌ها ترکیب کرده‌اند تا محدودیت‌های ذاتی ANN را برطرف کنند. در یک مطالعه تطبیقی، زرگری و همکاران [۸۱] ANN و ANFIS را برای پیش‌بینی تخلخل و نفوذپذیری در یک مخزن کربناتی ایران مقایسه کردند. ANFIS دقت بهتری نسبت به مدل ANN ارائه داد. نویسندگان همچنین پتانسیل الگوریتم‌های ژنتیک را برای بهبود پیش‌بینی ANN تأیید کردند. همچنین، الکتاری و همکاران [۸۳] ANN، ANFIS و SVM را مقایسه کردند. با این حال، نویسندگان خاطرنشان کردند که ANN دقت بهتری ارائه می‌دهد. برعکس، نوربانی و همکاران [۸۴] از فلورسانس پرتو ایکس دستی (HH-XRF) به عنوان ورودی برای پیش‌بینی تخلخل در یک مخزن گچی استفاده کردند. نویسندگان به سرعت و دقت ارائه‌شده توسط روش HH-XRF برای توصیف ژئوشیمیایی تکیه کردند. تکنیک‌های RF، ANN، GA-ANN و GA-RF برای تعیین دقیق‌ترین روش پیش‌بینی استفاده شدند. با این حال، GA-RF بالاترین سطح دقت را ارائه داد. با این حال، لیم و کیم [۸۰] از منطق فازی برای انتخاب پارامترهای ورودی بین لاگ‌های چاه قبل از اعمال ANN برای پیش‌بینی استفاده کردند. در مطالعه دیگری، احمدی و چن [۸۶] یک الگوریتم رقابتی امپریالیستی (ICA) و یک ترکیب از GA و PSO (HGAPSO) را برای پیش‌بینی تخلخل با استفاده از ANN اعمال کردند. نویسنده همچنین بهینه‌سازی HGAPSO را بر روی LSSVM برای پیش‌بینی تخلخل اعمال کرد. مدل‌ها با ANN مستقل و درخت‌های تصمیم فازی (FDT) مقایسه شدند. با این حال، مدل HGAPSO-LSSVM دقت بالاتری ارائه داد. علاوه بر این، سان و همکاران [۸۷] بهینه‌سازی شبکه عصبی المان با استفاده از الگوریتم بهینه‌سازی نهنگ (WOA) را برای پیش‌بینی تخلخل در چاه‌های نفتی غرب چین پیشنهاد کردند. الگوریتم WOA-Elman دقت بهتری نسبت به المان مستقل و الگوریتم BP ارائه داد. مطالعه‌ای توسط وانگ و کائو [۸۸] پیش‌بینی تخلخل با استفاده از یک روش یادگیری عمیق به نام شبکه عصبی یکپارچه را پیشنهاد کرد. روش پیشنهادی، که یک CNN یک‌بعدی را با GRU دوطرفه ترکیب می‌کند، دقت بالاتری نسبت به biGRU، GRU، LSTM، RNN و MLR نشان داد.

TABLE 2. Summary of literature on the prediction of permeability using machine learning.

References	Technique	Inputs Parameters	Field/Reservoir Source
[38]	ANN	SP, GR, RHOB, DT, NPHI, DRHO, latitude, longitude	Offshore gas field, Eastern Canada
[51]	RBF	core porosity	Asmari oil reservoir, south of Iran
[39]	ANN	RHOB, GR, NPHI and DT	North Sea case
[66]	CM	wireline and MWD logs	Oseberg field
[40]	ANN	GR, NPHI logs and RHOB	Uinta Basin, Southwest Utah field
[67]	CMEF	GR, LLD, LLS, MSFL, CAL, NPHI, RHOB and DT	Southern Taiwan
[61]	FL	CT, DRHO, DT, DR, MSFL, NPHI, PHIT, RHOB, RT, SWT	Middle Eastern carbonate reservoir
[44]	ANN	PHIT, SSA, and Swirr from core analysis	Carbonate reservoirs, USSR
[57]	ANN	GR, RHOB, NPHI, RT, NMR log data	Lower-Mesozoic reservoir
[52]	GA-ANN	SGR, RHOB, NP, PHIT, CT, SWT	Mansuri oil field, Iran
[71]	GA-Committee Machine	conventional well log data	Babal oil field, Persian Gulf
[45]	GA-ANN	GR, RHOB, NPHI, DT transit time, PHIT	Mansuri field
[50]	MNN, ANN	SGR, RT, SWT, PHIT and Secondary Porosity	Persian Gulf, Iranian off-shore
[58]	FN, ANFIS, ANN, NR	CT, DRHO, DT, MSFL, NPHI, PHIT, RHOB, RT, SWT	Middle Eastern Carbonate reservoir
[41]	ANN, MLR	GR, DT, NPHI, RHOB, RT, SWT	Oil field, southwest of Iran
[59]	ANN, SVM and ELM	CT, DRHO, DT, MSFL, NPHI, PHIT, RHOB, RT and SWT	Middle Eastern reservoir
[65]	T2FL, SVM and ANN	NPHI, RHOB, CALI, GR, SWT, and seismic parameters	Middle Eastern carbonate reservoir
[76]	Ensemble ANN	GR, PHIT, RHOB, SWT, RT, MSFL, NPHI and CALI	Middle Eastern carbonate reservoir
[72]	Ensemble of SVM, ANN, ANFIS	MSFL, NPHI, PHIT, RHOB, SWT, CALI, CT, DRHO, GR, RT	Middle Eastern reservoir
[73]	SVM, FN, T2FLS, FN-T2FLS, FIE-SVM, FN-SVM, DT-SVM	GR, PHIT, RHOB, RT, MSFL, CALI, CT, NPHI, ILD, EL, Magnetic resonance	Middle Eastern carbonate reservoir
[53]	ANN, FL, NF	GR, RHOB, RT, NPHI, SWT	Hassi R'Mel field, Algeria
[60]	GA-RVR, GA-SVR	GR, DT, RHOB, NPHI, LLD, PEF and MSFL	Carbonate oil field, Kuzestan, Iranian
[46]	MLP, RBF, GRNN, CM-GA(MLP-RBF-GRNN)	DT, RHOB, NPHI, RHOB	-
[54]	ANN, WNN	RT, SWT, PHIT, SGR, Sand, Dol, shale and secondary porosity	Persian Gulf, Iranian off-shore
[75], [77]	SVM, RF, SVM bagging, Stacked ensemble SVM	GR, PHIE, RHOB, SWT, RT, SFL, NPHI and CALI	Middle Eastern reservoir
[63]	SLR, MLR, MLP, and SVR	GR, AC, RT, RHOB, CNL, RT, RS and Coordinates(X, Y)	Chuanxi Depression, Sichuan Basin, China
[43]	ANN, SVM, and ANFIS	RHOB, NPHI, LLS, and LLD	Middle Eastern carbonate reservoir
[55]	CANFIS, MLP and SVM	GR, NPHI, RT, RHOB and DT	Measverde tight gas sandstones
[62]	FL	GR, SP, LLD, LLS, RHOB, AC and Elec	Mesozoic buried hill
[79]	RF-LR-XGBoost, PCA-XGBoost	CALI, DRHO, DT, GR, NPHI, PEF, RACEHM, RACELM, RD, RT, RHOB and ROP	Volve Oil Field, North Sea



TABLE 2. (Continued.) Summary of literature on the prediction of permeability using machine learning.

[48]	SHAP with LR, BPNN, SVM, RF, KNN, GBDT and XGBoost	GR, CAL, AC, DEN, CNC, RFOC, RILD, RILM, and SP	Wenchang A Sag
[49]	SHAP-PSO-XGBoost, CNN, LSTM, GRU	Por, CAL, GR, AC, RILD, SP, VSH, RFOC	Ordos Basin
[56]	GMDH, PR, SVR, DT	PHIT, SWT, SP	Carbonate Reservoir in Russia and Iran
[47]	MLP-SSD, MLP-PSO, MLP-GA	HCAL, CGR, PEF, NPHI, RHOB, RT, DTCO	Fahlian Chahi Formation
[64]	MELM-COA, MELM-PSO, MELM-GA, LSSVM-COA, LSSVM-PSO, LSSVM-GA, LSSVM, CNN	HCAL, SGR, PEF, NPHI, RHOB, RT, DTCO	Fahlian Formation, Iran
[74]	SIRes 1D-CNN, SIRes 2D-CNN, MIRes CNN, GMDH, RF	GFI, RHOB, NPHI, PEF, LLD, LLS, SGR, CGR, CAL, DCALI, and DRHO	Sarvak formation, Iran

Abbreviations: AC= Acoustic Log, ANN= Artificial Neural Network, BPNN= Backpropagation Neural Network, CAL= Caliper Log, CANFIS= Complex-Valued Adaptive Neuro-Fuzzy Inference System, CM= Committee Machine, CMEF= Committee Machine with Empirical Formulae, CNN= Convolutional Neural Network, CNC= Neutron Porosity Log, CNL= Compensated Neutron Log, COA= Cuckoo Optimization Algorithm, DCALI= Density Caliper Log, DEN= Density Log, DR= Deep Resistivity Log, DRHO= Bulk Density Log, DT= Sonic Log, ELM= Extreme Learning Machines, FL= Fuzzy Logic, GA= Genetic Algorithm, GBDT= Gradient Boosting Decision Trees, GMDH= Group Method of Data Handling, GR= Gamma Ray Log, GRNN= General Regression Neural Network, GRU= Gated Recurrent Unit, HCAL= High-Resolution Caliper Log, ILD= Deep Induction Log, KNN= k-Nearest Neighbors, LLD= Deep Laterolog Log, LLS= Shallow Laterolog Log, LSTM= Long Short-Term Memory, MWD= Measurement While Drilling, MSFL= Micro Spherically Focused Log, NPHI= Neutron Porosity Log, PEF= Photoelectric Effect Log, PCA= Principal Component Analysis, PHIE= Effective Porosity, PHIT= Porosity, PR= Polynomial Regression, RACEHM= Resistivity Log, RACELM= Resistivity Log, RBF= Radial Basis Function Network, RHOB= Bulk Density Log, RF= Random Forest, RILD= Deep Laterolog Resistivity Log, RILM= Shallow Laterolog Resistivity Log, RT= Resistivity Log, RVR= Relevance Vector Regression, RT= Resistivity Log, SFL= Shallow Micro Spherically Focused Log, SGR= Shale-Gas Ratio Log, SHAP= Shapley Additive Explanations, SIRes 1D-CNN= Single-Input deep Residual 1D-CNN, SIRes 2D-CNN= Single-Input deep Residual 2D-CNN, SLR= Simple Linear Regression, SSA= Specific Surface Area, SVM= Support Vector Machine, SVM= Support Vector Regression, Swirr= Irreducible Water Saturation, SWT= Water Saturation, T2FL= Two-Step Fuzzy Logic, T2FLS= Type-2 Fuzzy Logic System, VSH= Volume of Shale Log, WNN= Wavelet Neural Network, XGBoost= Extreme Gradient Boosting

سایر تکنیک‌های یادگیری ماشین نیز برای پیش‌بینی تخلخل استفاده شده‌اند. مطالعه‌ای توسط الانازی و گیتس [۸۹] تکنیک SVR را برای تخمین تخلخل بررسی کرد. مدل پیشنهادی از نظر دقت و مقاومت از MLP، GRNN و شبکه عصبی تابع پایه شعاعی (RBFNN) بهتر عمل کرد. با این حال، مقاومت SVR به انتخاب تابع هسته بستگی دارد. این مزیت با بار استفاده از منابع محاسباتی بسیار بیشتر از روش‌های جایگزین همراه است. با این حال، آنیفوووس و همکاران [۷۳] سه تکنیک انتخاب ویژگی را برای پیش‌بینی تخلخل با استفاده از اندازه‌گیری‌های آزمایشگاهی از میدان نفتی شمالی ماریون اعمال کردند. تکنیک‌های انتخاب ویژگی FN، DT و آنتروپی اطلاعات فازی (FIE) بر روی SVM و T2FL اعمال شدند. تکنیک ترکیبی FN-SVM در بین مدل‌های ترکیبی و مستقل جایگزین برجسته بود. علاوه بر این، احمدی و همکاران [۹۰] از توانایی بهینه‌سازی GA برای انجام پیش‌بینی‌ها با استفاده از FL و LSSVM استفاده کردند. مدل‌های پیشنهادی تخلخل و نفوذپذیری چاه‌ها از میدان‌های نفتی خلیج فارس شمالی را پیش‌بینی کردند. GA-LSSVM دقت کمی بهتر نسبت به روش جایگزین ارائه داد. همچنین، ژونگ و کار [۹۱] یک مدل ترکیبی SVM با

تابع هسته ترکیبی (MKF) را بررسی کردند. این مدل با استفاده از بهینه‌سازی ازدحام ذرات (PSO) برای بهبود قابلیت‌های پیش‌بینیکننده آن بهینه شد. از نظر دقت، روش پیشنهادی از SVM سنتی، LSSVM، ANN و RBF بهتر عمل کرد. در یک مطالعه جداگانه، اندرسن و همکاران [۹۲] بهینه‌سازی LSSVM را برای پیش‌بینی تخلخل و اشباع آب در میدان وارگ در نروژ انجام دادند. نویسندگان مدل‌های پیش‌بینیکننده را با استفاده از ترکیبات مختلف لاگ‌های چاه بررسی کردند. جالب توجه است که یافته‌های آن‌ها نشان داد که دقیق‌ترین پیش‌بینی‌ها زمانی حاصل می‌شود که تنها بر روی تخلخل تمرکز شود و تنها از سه لاگ چاه خاص استفاده شود: چگالی، مقاومت عمیق و اشعه گاما. علاوه بر این، تحقیقات آن‌ها نشان داد که گنجاندن لاگ‌های چاه اضافی بهبود قابل توجهی در عملکرد پیش‌بینیکننده مدل ایجاد نکرد. علاوه بر این، آنیفووس و همکاران [۷۷] یک مدل ترکیبی با استفاده از SVM ارائه دادند. مدل پیشنهادی پیش‌بینی‌هایی را بر اساس چندین مقدار پارامتر تنظیم‌شده بهینه ارائه می‌دهد. مقایسه عملکرد مدل پیشنهادی با SVM پیاده‌شده با استفاده از تکنیک بگینگ، یک SVM استاندارد و یک ترکیب از DT برتری آن را نشان داد. در مطالعه دیگری، طارق و همکاران [۹۳] شبکه عصبی عمیق (DNN)، RF، DT، KNN، XGBoost و AdaBoost را برای پیش‌بینی تخلخل NMR با استفاده از لاگ‌های چاه معمولی مقایسه کردند. بر اساس نتایج، DNN، RF و XGBoost بالاترین سطح دقت را نشان دادند. نتایج آزمایشی به شدت نشان می‌دهد که استفاده از DNN، RF یا XGBoost می‌تواند دقت پیش‌بینی‌ها را به طور قابل توجهی افزایش دهد.

از طرف دیگر، حقیقی و همکاران [۹۴] پیش‌بینی تخلخل در میدان دامار، اندونزی را با استفاده از الگوریتم XGBoost بهینه‌شده با تکنیک GridSearchCV (GS) پیشنهاد کردند. با این حال، پان و همکاران [۹۵] پیش‌بینی تخلخل را با استفاده از یک مدل XGBoost بهینه‌شده با GS و GA پیشنهاد کردند. وجود دو استراتژی بهینه‌سازی مختلف به مدل سود می‌رساند، همانطور که دقت آن نشان داده شد. مدل پیشنهادی زمانی که با یک مدل بهینه‌سازی GS به تنهایی بررسی شد، از LR، SVR، RF و XGBoost بهتر عمل کرد. بر اساس مرور ادبیات برای پیش‌بینی تخلخل، جدول ۳ خلاصه‌ای جامع از رویکردهای مختلف یادگیری ماشین استفاده‌شده، پارامترهای ورودی و محل مخزن بررسی‌شده را ارائه می‌دهد.

C. پیش‌بینی اشباع آب

اشباع آب یکی دیگر از ویژگی‌های مهم مخزن است که نشان‌دهنده بخش آب موجود در فضاها منفذی خاص است. این پارامتر در محاسبات عمق سوراخ کاری برای سایت‌های تولید هیدروکربن در خشکی و دریا کمک می‌کند [۹۶]. این پارامتر برای محاسبه صحیح حجم هیدروکربن ضروری است. در طول چند دهه گذشته، روش‌های تجربی مختلفی برای پیش‌بینی اشباع آب با استفاده از داده‌های پتروفیزیکی از لاگ‌های چاه، از جمله مقاومت، سرعت صوت، چگالی و

تخلخل نوترون معرفی شده‌اند. مدل تجربی پیشگام برای پیش‌بینی اشباع، مدل آرچی [۹۷] برای مخازن ماسه‌سنگی تمیز بود. چندین محقق تلاش کرده‌اند تا رابطه بین اشباع آب و داده‌های لاگ‌های چاه را برای پیش‌بینی اشباع آب در سازندهای مختلف استخراج کنند [۹۸]، [۹۹]، [۱۰۰]، [۱۰۱]، [۱۰۲]. با این حال، این روش‌ها به دلیل محدودیت‌های سازندگی خود، تنها در لیتولوژی‌های محدود قابل استفاده هستند. این مدل‌ها فاقد تعمیم‌پذیری هستند و نمی‌توانند به طور جهانی اعمال شوند. علاوه بر این، پارامترهای مرتبط با هر مدل دارای عدم قطعیت‌های ذاتی هستند که ممکن است منجر به نتایج اشتباه شود. بنابراین، تکنیک‌های یادگیری ماشین به طور گسترده‌ای برای پیش‌بینی اشباع آب استفاده شده‌اند.

ANN و FL نمونه‌هایی از تکنیک‌های محبوب هوش مصنوعی (AI) هستند که برای پیش‌بینی اشباع آب استفاده می‌شوند. در میان رویکردهای مختلف یادگیری ماشین، ANN بیشترین دامنه کاربردهای بالقوه را دارد و در زمینه‌های مختلف موفق بوده است. چندین مدل ANN با موفقیت بر روی داده‌های هسته‌ای و لاگ‌های چاه اعمال شده‌اند. اولین مورد هله و بات [۱۰۳] بود که یک شبکه عصبی کمیته را پیشنهاد کرد که از لاگ‌های سرعت صوت، چگالی، تخلخل نوترون و مقاومت به عنوان ورودی استفاده می‌کرد. سپس، شوکر [۱۰۴] یک مدل ANN را پیاده‌سازی کرد که لاگ پتانسیل خودی (SP log) را به ویژگی‌های ورودی اضافه کرد. برتری این مدل با مقایسه پیش‌بینی‌های اشباع آب تولید شده توسط ANN با پیش‌بینی‌های حاصل از تحلیل پتروفیزیکی سنتی اثبات شد. از طرف دیگر، مدل کاملیار [۱۰۵] تنها تخلخل و نفوذپذیری از هسته و همچنین ارتفاع بالاتر از سطح آب آزاد را در نظر گرفت. به طور مشابه، البلوشی و همکاران [۱۰۶] یک ANN را با استفاده از یک الگوریتم یادگیری انتشار برگشتی مقاوم آموزش دادند. نویسندگان همچنین تأثیر چندین پارامتر مختلف لاگ چاه را که ورودی‌های مدل بودند، با استفاده از یک روش رتبه‌بندی ویژگی‌ها بر روی لاگ‌های چاه بررسی کردند. مدل پیشنهادی برای پیش‌بینی اشباع آب استفاده شد و دقت بهتری نسبت به روش آماری ارائه داد. علاوه بر این، مردی و همکاران [۱۰۷] نیز ایده استفاده از یک مدل ANN را برای پیش‌بینی نه تنها اشباع آب، بلکه ضریب سیمان شدگی و ضریب اشباع در دو مخزن کربناتی در ایران مطرح کردند. مدل استفاده شده توسط نویسندگان نه تنها شامل اندازه‌گیری‌های لاگ چاه، بلکه تخلخل هسته نیز بود. اشباع آب، تخلخل و نفوذپذیری در منطقه دلتای نیجر توسط اوکون و همکاران [۱۰۸] با استفاده از یک ANN پیش‌خور انتشار برگشتی پیش‌بینی شد. مدل پیشنهادی شامل رتبه‌بندی ویژگی‌ها بود و دقت بالایی به دست آورد. در مطالعه البلوشی و همکاران [۱۰۹]، چگالی، نوترون، مقاومت و لاگ‌های سیم‌لاین فوتوالکتریک به عنوان ویژگی‌های ورودی انتخاب شدند تا یک مدل با استفاده از تکنیک ANN برای پیش‌بینی اشباع آب ساخته شود. بر اساس گفته‌های ناین و همکاران [۱۱۰]، برتری مدل ANN نسبت به مدل‌های سنتی در پیش‌بینی اشباع آب و تخلخل در یک مخزن ماسه‌سنگی شیل‌دار گزارش شد. داده‌های هسته‌ای از دو چاه تطابق عالی با مدل پیشنهادی نشان دادند، که تطابق خوبی را نشان داد.

TABLE 3. Summary of literature on the prediction of porosity using machine learning.

References	Technique	Inputs parameter	Field/reservoir source
[39]	ANN	DEN, DT, RT	-
[80]	FL-ANN	NPHI, CAL, LLD, LLS, RHOB, and SP	-
[89]	SVM	GR, CNL, DT, DEN	-
[39]	ANN	NMR logs and GR, RHOB, CNL, Deep and Shallow RL	Gas sand reservoir
[57]	ANN	GR, RHOB, CNL, Deep and Shallow RL	Gas sand reservoir
[81]	ANN	DT, NPHI, RHOB, CT, and SGR	Carbonate reservoir, Southern Iran
[73]	SVM, FN, T2FLS, FN-T2FLS, FIE-SVM, FN-SVM, DT-SVM	petrographic measurements - top interval grain density, grain volume, length and diameter	Northern Marion Platform of North America
[90]	GA-FL, GA-LSSVM	DT, DEN, CNL, PHIT	Northern Persian Gulf oil-fields
[78]	SVM, RF, SVM bagging, Stacked Ensembles	petrographic measurements - top interval, grain density, grain volume, length and diameter	Northern Marion Platform of North America
[82]	ANN	RHOB, CNL, AC, and NPHI	-
[43]	ANN, ANFIS, SVM	DT, RHOB, and NPHI	Carbonate reservoir
[86]	FDT, ANN, ICA-ANN, HGAPSO-LSSVM	DT, RHOB, PHIT, NPHI	Iranian oil fields
[91]	PSO-MKF-SVM	GR, DEN, Shale content, Slope of GR, Slope of density	Jacksonburg-Stringtown oil field
[85]	GRU NN, CA-GGRU, MLP, RNN		Ordos basin
[84]	RF, ANN, GA-RF, GA-ANN	Hand-held X-ray fluorescence (HH-XRF)	-
[88]	INN, BiGRU, GRU, LSTM, RNN, MLR	GR, DEN, CNL, DT, NPHI,	Western Sichuan Basin
[87]	Elman, WOA-Elman, BP	PE, DEN, M2R1, AC, GR, R25, R4 and CNL	Oilfield in Western China
[95]	LR, SVR, RF, XGBoost, GS-XGBoost, GS-GA-XGBoost	AC, CAL, CNL, DEN, GR and array induction resistivity(AT90)	Oilfield in Northern Shaanxi, China
[92]	LSSVM-PSO	RHO, RD, GR	Varg field, Norway
[94]	GS-XGBoost	ILD, NPHI, RHOB, GR, rock formation	Damar field, Indonesia
[93]	DNN, DT, RF, KNN, AdaBoost, XGBoost	GR, CALI, NPHI, PE, RHOB	-

Abbreviations: AC= Caliper Log, AdaBoost= Adaptive Boosting, ANN= Artificial Neural Network, ANFIS= Adaptive Neuro-Fuzzy Inference System, BP= Backpropagation, CA= Correlation Analysis, CAL= Caliper Log, CNL= Neutron Density Log, DNN= Deep Neural Network, DT= Sonic Log, DEN= Density Log, FL= Fuzzy Logic, FN= Functional Network, GA= Genetic Algorithm, GB= Gradient Boosting, GR= Gamma Ray Log, GRU= Gated Recurrent Unit, GS= Genetic Search, HGAPSO= Hybrid Particle Swarm Optimization and Genetic Algorithm, ICA= Imperialist Competitive Algorithm, ILD= Deep Induction Log, LLD= Deep Laterolog Log, LLS= Shallow Laterolog Log, LR= Lasso Regularization, LSSVM= Least Square Support Vector Machine, MKF= Mixed Kernels Function, MLP= Multilayer Perceptrons, NPHI= Neutron Porosity Log, PCA= Principal Component Analysis, PE= Photoelectric Effect, PSO= Particle Swarm Optimization, RHO= Resistivity Log, RD= Deep Resistivity Log, RF= Random Forest, RNN= Recurrent Neural Network, SVM= Support Vector Machine, SVR= Support Vector Regression, T2FLS= Type-2 Fuzzy Logic System, WOA= Whale Optimization Algorithm, XGBoost= eXtreme Gradient Boosting.

مطالعه دیگری توسط کناری و مشهور [۱۱۱] یک ANFIS را با ترکیب ANN و منطق فازی برای تخمین اشباع آب در یک میدان کربناتی ایران فرموله کردند. این مدل از ANN سنتی بهتر است، زیرا می‌تواند دقت، مقاومت و تعمیم‌پذیری بیشتری نسبت به هر یک از اجزای جداگانه ارائه دهد. علاوه بر این، ابراهیم و همکاران [۱۱۲] معادلات تجربی را با ANN و ANFIS برای پیش‌بینی اشباع آب مقایسه کردند. ANFIS کمی بهتر از ANN در نتیجه پیش‌بینی عمل کرد، اما به طور قابل توجهی بهتر از فرمول‌های تجربی بود. علاوه بر این، خان و همکاران [۱۱۳]

ANN و ANFIS را در یک مخزن کربناتی در خاورمیانه مقایسه کردند. نتایج نشان داد که ANFIS دقت خروجی کمی بهتر از ANN ارائه داد. در همین حال، ANN و FL توسط باقری و همکاران [۱۱۴] در یک مخزن کربناتی در خاورمیانه مقایسه شدند. خروجی نشان داد که مدل FL دقت بالاتری نسبت به مدل ANN ارائه می‌دهد.

برعکس، امیری و همکاران [۱۱۵] مدل ANN خود را با استفاده از یک الگوریتم رقابتی امپریالیستی (ICA) در یک مخزن غیرمتعارف بهینه کردند. علاوه بر این، نویسندگان تأثیر داده‌های پرت را مشاهده کردند، که هنگامی که به درستی شناسایی و حذف شدند، نتیجه پیش‌بینی را به طور قابل توجهی بهبود بخشیدند. در مطالعه دیگری، غلانو و همکاران [۱۱۶] مفهوم استفاده از یک شبکه عصبی پایه شعاعی را برای پیش‌بینی اشباع آب در سازند کربناتی سروک در ایران پیشنهاد کردند. مزایای مدل RBF نسبت به سایر مدل‌های شبکه عصبی شامل ساختار ساده و توانایی کسب دانش سریع است.

با این حال، آدینران و همکاران [۳۴] کارایی FN را در پیش‌بینی اشباع آب و تخلخل مخزن با استفاده از لاگ‌های چاه گزارش کردند. این مدل به دلیل ارائه راه‌حل‌های سریع و منحصر به فرد که از شبکه‌های عصبی بهتر عمل می‌کند، مورد توجه قرار گرفته است. همچنین، طارق و همکاران [۱۱۷] یک مدل FN را برای پیش‌بینی اشباع آب پیشنهاد کردند. دقت مدل با آزمایش چندین الگوریتم بهینه‌سازی، مانند تکامل تفاضلی (DE)، PSO و استراتژی تطبیق ماتریس کوواریانس (CMAES) برای توسعه دقیق‌ترین نسخه بهبود یافت. PSO بهترین انتخاب در بین آن‌ها بود. علاوه بر این، اندرسن و همکاران [۹۲] بهینه‌سازی ماشین بردار پشتیبان حداقل مربعات (LSSVM) را برای پیش‌بینی تخلخل و اشباع آب در میدان وارگ، نروژ انجام دادند. تحقیق آن‌ها شامل پیش‌بینی با استفاده از مجموعه‌های مختلف لاگ‌های چاه بود. جالب توجه است که بهترین نتایج زمانی حاصل شد که پیش‌بینی اشباع آب تنها با استفاده از چهار لاگ انجام شد: مقاومت متوسط، اشعه گاما، کالیبر تنظیم‌شده و لاگ پتانسیل خودی. جالب‌تر این که، مطالعه آن‌ها نشان داد که گنجاندن لاگ‌های اضافی منجر به بهبود عملکرد پیش‌بینی‌کننده نشد.

SVM یک تکنیک جایگزین دیگر برای یادگیری ماشین است که توسط [۹۶]، [۱۱۸] و [۱۱۹] برای پیش‌بینی اشباع آب ارائه شده است. بر اساس گفته‌های ملاجان و همکاران [۱۱۸]، این مدل از ANN از نظر دقت پیش‌بینی‌ها بهتر عمل کرد. علاوه بر این، میاه و همکاران [۹۶] نسخه دیگری از SVM را بررسی کردند که از حداقل مربعات به عنوان تابع هسته خود استفاده می‌کرد و به آن ماشین بردار پشتیبان حداقل مربعات (LS-SVM) می‌گفتند. نویسندگان همچنین اهمیت رتبه‌بندی ویژگی‌ها را در نظر گرفتند، زیرا با در نظر گرفتن تنها مهم‌ترین ویژگی‌های ورودی، زمان و پیچیدگی مدل را کاهش می‌دهد. LS-SVM پیشنهادی آن‌ها از ANN از نظر دقت پیش‌بینی‌کننده بهتر عمل کرد.

بازیر و همکاران [۱۲۰] عملکرد SVM، ANN، جنگل تصادفی و گرادیان بوستینگ را برای پیش‌بینی اشباع آب با استفاده از یک مجموعه داده کوچک در یک مخزن ماسه‌سنگی مقایسه کردند. اگرچه نویسندگان قابلیت اطمینان تمام تکنیک‌های مختلف استفاده‌شده را گزارش کردند، SVM بهترین عملکرد را ارائه داد. در مطالعه دیگری، هاداویموقدام

و همکاران [۱۲۱] دقت الگوریتم‌های مختلف بوستینگ، یعنی XGBoost، LightGBM، AdaBoost، CatBoost و Super Learner را برای پیش‌بینی اشباع آب در یک مخزن ماسه‌سنگی در فدراسیون روسیه مقایسه کردند. از بین همه گزینه‌ها، XGBoost دقیق‌ترین بود. با این حال، دقت تنها کمی بهتر از Super Learner بود.

اوپر و همکاران [۷۹] یک مدل ترکیبی متشکل از یک مدل ترکیبی از جنگل تصادفی و تنظیم Lasso به عنوان تکنیک انتخاب ویژگی و XGBoost به عنوان پیش‌بینی‌کننده برای پیش‌بینی اشباع آب و نفوذپذیری ساختند. مدل ترکیبی پیشنهادی بهتر از مدل XGBoost سنتی و یک مدل ترکیبی که PCA و XGBoost را ادغام می‌کرد، عمل کرد. بر اساس مرور ادبیات برای پیش‌بینی اشباع آب، جدول ۴ خلاصه‌ای جامع از رویکردهای مختلف یادگیری ماشین استفاده‌شده، پارامترهای ورودی و محل مخزن بررسی‌شده را ارائه می‌دهد.

۷.۱. بحث

یادگیری ماشین رشد قابل توجهی در اکتشاف نفت و گاز داشته است. این رشد را می‌توان به توانایی آن در حل چالش‌های مختلف در صنعت، مانند پردازش داده‌های لرزه‌ای، طبقه‌بندی لیتوفاسیس‌ها و توصیف مخزن نسبت داد.

مدل‌های یادگیری ماشین مزایای متعددی نسبت به روش‌های سنتی اکتشاف نفت و گاز دارند، که از مدل‌های تجربی و نیمه‌تجربی برای تخمین پارامترهای مخزن مشتق شده‌اند. مدل‌های یادگیری ماشین می‌توانند با تعاملات پیچیده و غیرخطی بین پارامترهای لاگ چاه، بینش‌هایی را کشف کنند که مدل‌های سنتی از آن‌ها غافل بوده‌اند. علاوه بر این، آن‌ها نتایج بسیار دقیقی را با استفاده از زمان و منابع بسیار کمتری نسبت به روش‌های سنتی ارائه می‌دهند [۱۱۲]. مزایای یادگیری ماشین را نمی‌توان نادیده گرفت، زیرا واضح است که آن‌ها می‌توانند زمان مورد نیاز برای پردازش داده‌های لرزه‌ای، طبقه‌بندی لیتوفاسیس‌ها و توصیف مخزن را به طور چشمگیری کاهش دهند. به طور مشابه، در نتیجه، میزان نیروی کار و منابع مورد نیاز برای حل مشکلات در صنعت کاهش می‌یابد [۱۳]، [۲۵]، [۶۳].

با این حال، محدودیت‌هایی برای هر روش وجود دارد و یادگیری ماشین نیز از این قاعده مستثنی نیست. با وجود پیشرفت‌های قابل توجه در حل مسائل خطی، غیرخطی و پیچیده، از جمله طبقه‌بندی، رگرسیون و پیش‌بینی، چندین نقطه ضعف وجود دارد. MLP که معمولاً استفاده می‌شود، کند آموزش می‌بیند، مستعد گیر کردن در مینیمم‌های محلی است و به آزمایش و خطای زیادی برای تعیین توپولوژی بهینه نیاز دارد. علاوه بر این، به مقدار داده بیشتری نسبت به مدل‌های مشابه نیاز دارد.



علاوه بر این، یک همبستگی مستقیم بین کیفیت داده‌های مورد استفاده برای آموزش یک الگوریتم یادگیری ماشین و عملکرد الگوریتم وجود دارد. کیفیت داده‌های مورد استفاده برای آموزش مدل به طور مستقیم بر دقت آن تأثیر می‌گذارد [۱۲۲]. این موضوع اغلب به عنوان اصل GIGO (زباله داخل، زباله خارج) شناخته می‌شود. این بدان معناست که اگر داده‌های ورودی ضعیف باشند، خروجی نیز ضعیف خواهد بود. داده‌های اصلی ممکن است به همبستگی‌های غیرخطی فشرده شده باشند که تنها پس از پیش‌پردازش گسترده داده‌ها آشکار می‌شوند. داده‌ها همچنین ممکن است به دلایل مختلفی ناقص باشند، مانند مقادیر خارج از محدوده، اطلاعات متناقض یا تغییرات تصادفی جزئی در مشاهدات. در نتیجه، پیش‌پردازش قابل توجهی باید انجام شود تا تعامل پیچیده متغیرهایی که ممکن است در منابع داده‌های مختلف در صنعت بالادستی نفت و گاز کشف شوند، نمایان شود.

علیرغم اثربخشی آن‌ها، مدل‌های یادگیری ماشین فردی به اندازه کافی مقاوم نیستند تا مسائل پیچیده و عدم قطعیت‌های موجود در بخش نفت و گاز را حل کنند. محققان اخیراً بر استفاده از روش‌های ترکیبی و هیبرید یادگیری ماشین برای غلبه بر این مشکل تمرکز کرده‌اند. این موضوع با افزایش تعداد مقالات اخیر در مورد استفاده از روش‌های ترکیبی و هیبرید یادگیری ماشین برای پردازش داده‌های لرزه‌ای، طبقه‌بندی فاسیس و لیتوفاسیس‌ها و توصیف مخزن نشان داده شده است. با این حال، هنوز کار زیادی باید انجام شود تا روش‌های ادغام ترکیبی استاندارد شوند.

یادگیری ماشین هیبرید برای تکمیل مدل‌های فردی با نقاط قوت دیگران استفاده شد. یادگیری ماشین هیبرید محاسبات یا فرآیندهای مختلفی از چندین مدل را ترکیب می‌کند، که همه آن‌ها با هدف بهبود یکدیگر هستند. چندین مدل پایه با هم همکاری می‌کنند تا یکدیگر را تکمیل و تقویت کنند و نتایج بهبود یافته‌ای نسبت به معادل‌های تک مدلی خود ارائه دهند. الگوریتم‌های بهینه‌سازی مانند GA ممکن است با انتخاب بهترین ابرپارامترها، مدل‌ها را بهبود بخشند. الگوریتم‌های کاهش ابعاد، مانند PCA، ممکن است پیچیدگی مدل را کاهش دهند و در عین حال نویز را از داده‌ها حذف کنند. گزینه‌های هیبریدسازی مختلفی برای بهبود مدل‌های یادگیری ماشین فردی که قادر به حل چالش‌های پیچیده بخش نفت و گاز هستند، وجود دارد.

یک ماشین کمیت برای توسعه بیشتر شبکه عصبی استفاده شد. ترکیب تخصص چندین متخصص به جای تمرکز صرف بر یک متخصص برتر، دقت، مقاومت و توانایی تعمیم‌پذیری بهتری را به دست می‌آورد. این به این دلیل است که تعمیم‌پذیری اعضای فردی منحصر به فرد نیست. هرس متخصصان ممکن است محدودیت‌های منابع اضافی تحمیل شده توسط ماشین کمیت را دور بزنند.

علاوه بر این، یادگیری ترکیبی بیشتر برای بهبود عملکرد مدل‌های یادگیری ماشین فردی برای حل مسائل پیچیده در اکتشاف نفت و گاز مورد تحقیق قرار گرفته است. ترکیب‌ها می‌توانند نتایج مختلفی را ادغام کنند، از جمله چندین تکنیک یادگیری، تفسیرهای متضاد داده‌ها، ملاحظات داده‌های نمونه‌گیری تصادفی، ساختارهای چند مدلی و سایر ویژگی‌های مورد علاقه تعریف شده. از آنجایی که مدل‌های ترکیبی می‌توانند چندین فرضیه را به طور همزمان مدیریت کنند،

می‌توانند به غلبه بر درجه بالای عدم قطعیت موجود در ویژگی‌های مخزن و متغیرهای تنظیم مدل کمک کنند. این امر نتایج قابل اطمینان‌تر و دقیق‌تری را فراهم می‌کند و نتیجه‌گیری کلی را با کمترین احتمال خطا و ابهام ارائه می‌دهد. یادگیری ترکیبی می‌تواند سنتز داده‌های بسیار بعدی و چندوجهی، مانند داده‌های موجود در بخش نفت و گاز را مدیریت کند. فرصت‌های بی‌پایانی برای بررسی و تجزیه و تحلیل یادگیری ترکیبی برای دستیابی به پتانسیل و عملکرد بهبود یافته وجود دارد.

یادگیری ماشین می‌تواند به طور قابل توجهی بر تصمیم‌گیری‌های متخصصان صنعت نفت و گاز تأثیر بگذارد. انتظار می‌رود که یادگیری ماشین در سال‌های آینده اهمیت بیشتری در بخش‌های نفت و گاز پیدا کند [۲۰]. با این حال، محققان هنوز با مشکل دسترسی به داده‌های آزمایشگاهی و میدانی مواجه هستند، که مانعی برای بهبود ادبیات است.

با ادامه تولید حجم عظیمی از داده‌ها در اکتشاف نفت و گاز، ایجاد، بهبود و ادغام روش‌های مدیریت داده‌های بزرگ در حوزه هوش مصنوعی اهمیت فزاینده‌ای پیدا می‌کند. استفاده حداکثری از داده‌های موجود یک تمرکز فعلی است و احتمالاً در آینده نیز ادامه خواهد داشت. برای دستیابی به بهینه‌سازی، باید از منابع قدرتمند هوش مصنوعی استفاده کرد.

نقشه راه نشان داده شده در شکل ۳ فرآیندهای حیاتی برای دستیابی به دقت بهینه در کاربرد یادگیری ماشین در اکتشاف نفت و گاز را نشان می‌دهد. اولین مرحله شامل جمع‌آوری داده‌های با کیفیت بالا درباره مخزن با استفاده از جدیدترین ابزارها و روش‌های لاگ‌گیری چاه است. این فناوری‌ها ممکن است طیف گسترده‌ای از اندازه‌گیری‌ها را ارائه دهند که می‌توانند برای تعیین جنبه‌های زمین‌شناسی، ژئوفیزیکی و توصیف مخزن استفاده شوند. مرحله دوم اطمینان از رعایت روش‌های با بالاترین کیفیت داده است. اعتبارسنجی جامع داده‌ها، پاک‌سازی و نرمال‌سازی برای اطمینان از صحت و قابلیت اطمینان داده‌ها مورد نیاز است. کیفیت داده‌های مورد استفاده برای مدل‌سازی بر دقت و اثربخشی الگوریتم یادگیری ماشین تأثیر می‌گذارد. آماده‌سازی داده‌ها مرحله سوم است. این فرآیند شامل انتخاب ویژگی‌های مهم داده‌ها، مقیاس‌بندی و تبدیل داده‌ها به قالبی است که برای الگوریتم‌های یادگیری ماشین مناسب باشد. شناسایی و حذف ویژگی‌های نامربوط به مسئله مورد نظر بسیار مهم است. مرحله چهارم انتخاب الگوریتم یادگیری ماشین مناسب برای داده‌ها و کار است. روش‌های یادگیری ماشین مانند طبقه‌بندی، رگرسیون، هیبرید و ترکیب‌ها ممکن است استفاده شوند. الگوریتم مورد استفاده به مسئله خاص و مجموعه داده بستگی دارد. مرحله پنجم ارزیابی الگوریتم یادگیری ماشین با استفاده از معیارهای مختلف مانند دقت، precision، میانگین خطای مطلق، میانگین مربعات خطا و R-squared است. این مرحله به تعیین صحت مدل و شناسایی زمینه‌های بهبود کمک می‌کند. مرحله نهایی افزایش دقت و عملکرد الگوریتم یادگیری ماشین است. این ممکن است با تنظیم ابرپارامترها یا استفاده از یک روش کاملاً جدید انجام شود. هدف دستیابی به بالاترین دقت پیش‌بینی ممکن با استفاده از داده‌های ارائه شده بود. به طور کلی، مراحل نشان داده شده در شکل نقطه شروع خوبی برای محققان و متخصصانی است که علاقه‌مند به استفاده از یادگیری ماشین برای پیش‌بینی ویژگی‌های مخزن در مرحله اکتشاف نفت و گاز هستند.

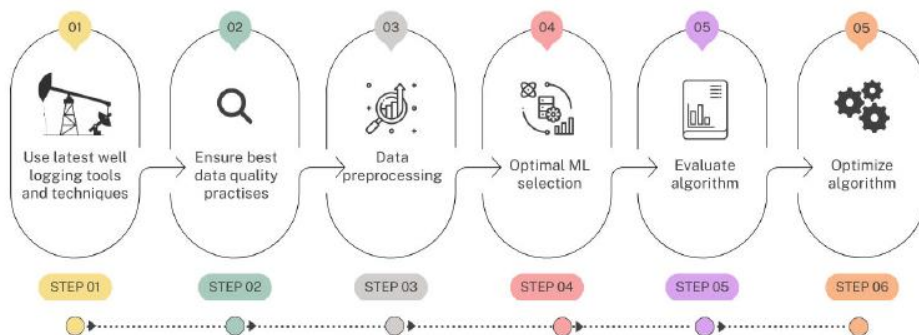


FIGURE 3. Road-map for optimal accuracy of machine learning techniques in oil and gas exploration.

۷.۱. چالش‌ها

صنعت نفت و گاز از طریق فعالیت‌های اکتشاف، حفاری، تولید و پالایش، حجم عظیمی از داده‌ها را تولید می‌کند و آن را به یکی از صنایع داده‌محور در جهان تبدیل می‌کند. مزایای یادگیری ماشین در این صنعت را نمی‌توان نادیده گرفت، زیرا می‌تواند کارایی را افزایش دهد، هزینه‌ها را کاهش دهد و ایمنی را بهبود بخشد. با این حال، برخی چالش‌های فنی مهم باید برطرف شوند تا پتانسیل یادگیری ماشین در مرحله اکتشاف این صنعت به طور کامل مورد استفاده قرار گیرد. این چالش‌ها در ادامه توضیح داده شده‌اند.

A. مسائل مربوط به داده‌ها

۱) دسترسی به داده‌ها

یکی از موانع اصلی در استفاده گسترده از یادگیری ماشین در صنعت نفت و گاز، کمبود داده‌های با کیفیت و در دسترس است. صنعت نفت و گاز حجم عظیمی از داده‌ها تولید می‌کند، اما بسیاری از این داده‌ها ساختار نیافته، پراکنده و دسترسی به آن‌ها دشوار است [۱۲۳]، [۱۲۴]. این موضوع یک مشکل جدی برای الگوریتم‌های یادگیری ماشین است، زیرا آن‌ها بهترین عملکرد را زمانی دارند که با حجم زیادی از داده‌های با کیفیت تغذیه شوند [۱۲۵].

مرحله اکتشاف نفت و گاز نیز به کمبود داده در صنعت کمک می‌کند. در مراحل اولیه اکتشاف، داده‌های محدود رایج است، زیرا حفاری چاه‌ها در شرایط سخت مانند اعماق دریا یا مناطق قطبی دشوار است. این موضوع باعث می‌شود

جمع‌آوری و انتقال داده‌ها از چنین مناطقی پرزحمت و پرهزینه باشد. استفاده از داده‌های به دست آمده در کاربردهای یادگیری ماشین ممکن است دشوار باشد، به ویژه اگر داده‌ها محدود، ناقص، ناسازگار یا با کیفیت پایین باشند [۱].

۲) پیش‌پردازش داده‌ها

کیفیت پایین داده‌ها منجر به چالش پیش‌پردازش داده‌ها می‌شود، زیرا داده‌های مورد نیاز برای آموزش مدل‌های یادگیری ماشین بسیار پیچیده هستند [۱۲۶]. این داده‌ها ممکن است شامل بررسی‌های لرزه‌ای، داده‌های حفاری، لاگ‌های چاه، داده‌های تولید و سایر داده‌های ژئوفیزیکی با کیفیت و فرمت‌های مختلف باشند. ممکن است حجم قابل توجهی از داده‌های تکراری، ناسازگار و مقادیر گم‌شده در داده‌ها وجود داشته باشد که نیاز به پاک‌سازی و استانداردسازی گسترده قبل از استفاده مفید از داده‌ها دارد.

علاوه بر این، ترکیب این منابع داده‌های مختلف منجر به حجم زیادی از داده‌ها می‌شود که چالش‌های مربوط به ابعاد بالا را به دلیل تعداد زیادی از ویژگی‌های اندازه‌گیری شده در اعماق و مکان‌های مختلف ایجاد می‌کند. همچنین، شرایط زمین‌شناسی ذاتاً نامشخص، عدم قطعیت‌های بیشتری را در داده‌های زیرسطحی ناشی از خطاهای اندازه‌گیری، کالیبراسیون، پردازش، درون‌یابی و برون‌یابی ایجاد می‌کند.

مخازن ویژگی‌های زمین‌شناسی را در مقیاس‌های مختلف، از ساختارهای میکروسکوپی در سطح منافذ تا ساختارهای ماکروسکوپی در سطح میدان، نشان می‌دهند. ادغام داده‌های جمع‌آوری شده در مقیاس‌های مختلف برای توسعه مدل‌های مخزن دقیق و جامع بسیار مهم است. فعالیت‌های اکتشاف اغلب شامل داده‌های مکانی، مانند نقشه‌های زمین‌شناسی، بررسی‌های لرزه‌ای و تصاویر ماهواره‌ای است که چالش‌های منحصر به فردی در ادغام داده‌ها، مهندسی ویژگی‌ها و تقاضای محاسباتی ایجاد می‌کند. اطلاعات زمانی موجود در برخی از مجموعه داده‌های اکتشافی، که تغییرات تاریخی در زمین‌شناسی یا عوامل محیطی را ثبت می‌کنند، لایه دیگری از پیچیدگی و عدم قطعیت را اضافه می‌کنند و نیاز به تکنیک‌های تخصصی مانند تحلیل سری‌های زمانی و ادغام داده‌ها برای استخراج بینش‌های معنادار دارند.

پیش‌پردازش در مخازن کربناتی به دلیل ناهمگنی شدید و ساختار منافذ پیچیده‌ای که از تخلخل ماتریس، حفره‌ها، شکاف‌ها و سایر ویژگی‌های زمین‌شناسی تشکیل شده‌اند، به طور قابل توجهی چالش‌برانگیزتر است [۴۳]. این موضوع منجر به رابطه ضعیف بین تخلخل و نفوذپذیری می‌شود. به دلیل این موضوع، پیش‌بینی نفوذپذیری با استفاده از لاگ NMR دشوارتر می‌شود، زیرا به شدت به همبستگی بین تخلخل و نفوذپذیری وابسته است. علاوه بر این، ممکن است داده‌های پرت و ناهنجاری در داده‌ها وجود داشته باشد که دقت مدل‌های یادگیری ماشین را کاهش می‌دهد.

۳) پراکندگی داده‌ها و محدودیت‌های دسترسی

ساختار پراکنده صنعت نیز یکی از دلایل کمبود داده است [۱۲۴]. بسیاری از شرکت‌ها، پیمانکاران و ارائه‌دهندگان خدمات در عملیات اکتشاف و تولید در صنعت نفت و گاز مشارکت دارند، که آن را به یک صنعت بسیار غیرمتمرکز تبدیل



می‌کند. این پراکندگی منجر به ایجاد سیلوهای داده و دسترسی محدود می‌شود، که انتقال داده‌ها بین نهادهای مختلف را دشوار می‌کند.

در نهایت، نگرانی‌های قانونی و حریم خصوصی دسترسی به داده‌ها را در صنعت نفت و گاز محدود می‌کند. به دلیل ماهیت بالقوه حساس داده‌های جمع‌آوری شده در طول اکتشاف و تولید، مقررات سخت‌گیرانه‌ای بر جمع‌آوری، استفاده و انتشار آن‌ها اعمال می‌شود.

۴) راه‌حل‌های مقابله با چالش‌ها

راه‌حل‌های بالقوه‌ای برای غلبه بر این چالش‌ها وجود دارد. یکی از راه‌حل‌ها این است که ذینفعان صنعت همکاری کنند و اطلاعات را به اشتراک بگذارند. توسعه استانداردهای داده مشترک، انتشار داده‌ها در مخازن عمومی و همکاری با مؤسسات دانشگاهی برای ایجاد زیرساخت‌های اشتراک‌گذاری داده‌ها گزینه‌های عملی هستند.

افزایش داده‌ها (Data Augmentation) یک رویکرد جایگزین است که در آن داده‌های جدید به داده‌های موجود اضافه می‌شوند. این ممکن است شامل تولید داده‌های مصنوعی با استفاده از ابزارهای شبیه‌سازی، افزایش تصاویر لرزه‌ای با روش‌های بینایی کامپیوتری یا استفاده مجدد از داده‌های سایر منابع از طریق یادگیری انتقالی باشد. روش‌های مختلفی مانند چرخش، برش، تغییر اندازه و افزودن نویز به داده‌های اصلی برای ایجاد داده‌های آموزشی بیشتر از مجموعه‌های داده موجود استفاده می‌شوند. این روش‌ها می‌توانند برای انواع داده‌های تصویری مانند SEM، مقاطع نازک، هسته‌ها و تصاویر لرزه‌ای اعمال شوند تا کیفیت و کمیت داده‌ها افزایش یابد. روش‌هایی مانند کاهش نمونه‌برداری، افزایش نمونه‌برداری، درون‌یابی، برون‌یابی و هموارسازی می‌توانند بر روی لاگ‌های چاه اعمال شوند.

تلاش‌هایی باید برای جمع‌آوری داده‌های بیشتر با استفاده از تکنیک‌های نوآورانه انجام شود تا دسترسی به داده‌ها افزایش یابد. داده‌ها از مناطق غیرقابل دسترس را می‌توان با استفاده از فناوری‌های سنجش از دور مانند پهپادها و تصاویر ماهواره‌ای جمع‌آوری کرد. داده‌های حفاری و تولید را می‌توان با استفاده از فناوری‌های سنسور مدرن به صورت بلادرنگ جمع‌آوری کرد.

بهبود کیفیت داده‌های موجود راه‌حل دیگری برای مقابله با مشکل کمبود داده است. سرمایه‌گذاری در روش‌های پیش‌پردازش داده‌ها می‌تواند به اطمینان از کیفیت کافی داده‌ها کمک کند. این ممکن است شامل پاک‌سازی داده‌ها، نرمال‌سازی و تبدیل باشد. یک رویکرد کنترل کیفیت نیز می‌تواند برای اطمینان از کیفیت کافی داده‌ها برای کاربردهای یادگیری ماشین استفاده شود. این ممکن است شامل ایجاد دستورالعمل‌های استاندارد برای جمع‌آوری داده‌ها و انجام بررسی‌های مختلف داده‌ها برای سازگاری و اعتبارسنجی باشد. تکنیک‌های کاهش ابعاد می‌توانند برای حفظ اطلاعات ضروری در حالی که تعداد ویژگی‌ها کاهش می‌یابد، استفاده شوند. روش‌های انتخاب ویژگی نیز می‌توانند برای شناسایی و حفظ مرتبط‌ترین ویژگی‌ها اعمال شوند.

علاوه بر این، تکنیک‌های مدل‌سازی چندمقیاسی که هم ویژگی‌های میکروسکوپی و هم ماکروسکوپی را در نظر می‌گیرند، می‌توانند استفاده شوند. این شامل تطبیق الگوریتم‌ها برای مدیریت داده‌ها در مقیاس‌های مختلف و ادغام مجموعه‌های داده متنوع برای ایجاد یک مدل مخزن جامع است. همچنین، تکنیک‌های کمی‌سازی عدم قطعیت می‌توانند در پیش‌پردازش داده‌ها ادغام شوند تا به مدل‌سازی و مدیریت عدم قطعیت‌ها کمک کنند و نمایش قوی‌تری از شرایط زمین‌شناسی ارائه دهند. تکنیک‌های تخصصی مانند ادغام داده‌های مکانی، مهندسی ویژگی‌ها و روش‌های محاسباتی برای مقابله با چالش‌های منحصر به فرد داده‌های مکانی می‌توانند اجرا شوند. برای داده‌های زمانی، از تحلیل سری‌های زمانی و تکنیک‌های ادغام داده‌ها برای استخراج بینش‌های معنادار از تغییرات تاریخی استفاده می‌شود.

مدل‌هایی که به طور خاص برای مخازن کربناتی طراحی شده‌اند، باید ویژگی‌های منحصر به فرد تخلخل ماتریس، حفره‌ها، شکاف‌ها و سایر ویژگی‌ها را در نظر بگیرند. بررسی تکنیک‌های پیشرفته یادگیری ماشین که می‌توانند با رابطه ضعیف بین تخلخل و نفوذپذیری مقابله کنند، مفید خواهد بود. علاوه بر این، بررسی دقیق اجرای روش‌های تشخیص داده‌های پرت مهم است، زیرا برخی از ویژگی‌های منحصر به فرد زیرسطحی ممکن است به عنوان داده‌های پرت در نظر گرفته شوند، که می‌تواند دقت مدل را بهبود بخشد.

روش‌های جامعی که بتوانند عدم قطعیت‌ها، ناهمگنی و ساختارهای پیچیده را مدیریت کنند، باید توسعه یابند. آینده نیازمند چارچوب‌های یادگیری ماشین قوی و انعطاف‌پذیر است که بتوانند با چالش‌های کیفیت داده‌ها، عدم قطعیت‌ها و محدودیت‌ها مقابله کنند. این امر امکان ایجاد رابطه قوی بین ورودی و خروجی را فراهم می‌کند. رویکردهای پیشرفته یادگیری ماشین، مانند انتخاب ویژگی، کاهش ابعاد و تنظیم مناسب، می‌توانند برای کشف روابط پیچیده داده‌ها و بهبود دقت پیش‌بینی استفاده شوند.

تحقیقات می‌توانند بر بهبود تکنیک‌های درون‌یابی برای افزایش دقت و قابلیت اطمینان پیش‌بینی‌ها، به ویژه در مناطق با داده‌های پراکنده، تمرکز کنند. روش‌های پیشرفته آمار فضایی و یادگیری ماشین، مانند فرآیندهای گاوسی، می‌توانند بررسی شوند. توسعه فرمت‌های استاندارد داده‌ها، هستی‌شناسی‌ها و استانداردهای فراداده برای داده‌های مکانی می‌تواند به ادغام داده‌ها کمک کند. ابزارهای خودکار برای هماهنگ‌سازی داده‌ها می‌توانند ایجاد شوند. تکنیک‌هایی برای ادغام مؤثر داده‌های زمانی و مکانی می‌توانند توسعه یابند. این می‌تواند شامل تحقیقات در پایگاه‌های داده مکانی-زمانی و کاربردهای GIS (سیستم اطلاعات جغرافیایی) باشد.

علاوه بر این، تعامل با سازمان‌های نظارتی برای تعیین قوانین تبادل و استفاده از داده‌ها به شدت انطباق نظارتی را افزایش می‌دهد. روش‌های انجام این کار شامل توافق‌نامه‌های اشتراک‌گذاری داده‌ها و ناشناس‌سازی داده‌ها است.

تحقیقات آینده می‌توانند بر رویکردهای نوآورانه برای مقابله مؤثر با چالش کمبود داده تمرکز کنند. استخراج اطلاعات بهبود یافته ممکن است با توسعه الگوریتم‌های جدید داده‌کاوی که می‌توانند با مجموعه‌های داده بزرگ و پیچیده مقابله



کنند، امکان‌پذیر باشد. به طور مشابه، توانایی پیش‌بینی‌کننده می‌تواند با توسعه الگوریتم‌های جدید یادگیری ماشین که برای یادگیری از مجموعه‌های داده کوچک و پرنویز بهینه شده‌اند، بهبود یابد. علاوه بر این، پیشرفت‌ها در تکنیک‌های ادغام داده‌ها امکان ادغام کارآمدتر داده‌ها از منابع مختلف را فراهم می‌کند. در نهایت، تکنیک‌های بهبود یافته یادگیری ماشین برای کشف داده‌ها در داده‌های محدود می‌توانند برای شناسایی الگوها و بینش‌های جدید در داده‌های محدود بررسی شوند.

هنگام مقابله با مشکل مسائل داده‌ها در صنعت نفت و گاز برای اهداف یادگیری ماشین، یک رویکرد ترکیبی به احتمال زیاد بهترین نتایج را ارائه می‌دهد. یک رویکرد جامع شامل پاک‌سازی داده‌ها، کاهش ابعاد، مدیریت عدم قطعیت، مدل‌سازی چندمقیاسی و تکنیک‌های تخصصی برای داده‌های مکانی و زمانی برای مقابله با چالش‌های پیش‌پردازش داده‌ها در بخش اکتشاف نفت و گاز ضروری است. تطبیق راه‌حل‌ها با شرایط زمین‌شناسی خاص، مانند آن‌هایی که در مخازن کربناتی وجود دارند، اثربخشی مدل‌های یادگیری ماشین را بیشتر افزایش می‌دهد. علاوه بر این، صنعت نفت و گاز می‌تواند در صورت همکاری اعضا برای بهبود اشتراک‌گذاری، جمع‌آوری و کیفیت داده‌ها، از مزایای کامل یادگیری ماشین در اکتشاف و تولید بهره‌مند شود. جهت‌های تحقیقاتی آینده نیز باید در توسعه رویکردهای نوآورانه گنجانده شوند.

B. شفافیت و تفسیرپذیری مدل‌ها

(۱) شفافیت مدل‌ها

شناسایی نحوه تولید پیش‌بینی‌ها توسط مدل‌های یادگیری ماشین و عوامل مؤثر بر این پیش‌بینی‌ها، یکی از چالش‌های مهم در اکتشاف نفت و گاز است. این مشکل به این دلیل ایجاد می‌شود که بسیاری از الگوریتم‌های محبوب یادگیری ماشین، مانند شبکه‌های عصبی، به عنوان “جعبه سیاه” در نظر گرفته می‌شوند، به این معنی که فرآیند تصمیم‌گیری آن‌ها اساساً غیرشفاف است و توضیح آن‌ها دشوار است [۲۰].

(۲) تفسیرپذیری مدل‌ها

به دلیل ماهیت پیچیده و چندبُعدی داده‌های مورد استفاده در اکتشاف نفت و گاز، تفسیر این مدل‌ها دشوار است. الگوریتم‌های یادگیری ماشین ممکن است بر روی طیف گسترده‌ای از داده‌های ژئوفیزیکی، از جمله بررسی‌های لرزه‌ای، لاگ‌های چاه و داده‌های تولیدی آموزش ببینند که همه این‌ها می‌توانند دارای ویژگی‌های متعدد و روابط پیچیده باشند. به دلیل پیچیدگی این داده‌ها، تفسیر پیش‌بینی‌های مدل‌های یادگیری ماشین ممکن است چالش‌برانگیز باشد [۴۹].

هنگامی که خروجی مدل‌ها به راحتی قابل تفسیر نباشند، ارزیابی نتایج مدل و شناسایی خطاها یا سوگیری‌ها دشوار می‌شود.

۳) تجسم داده‌های چندبعدی

تجسم داده‌های چندبعدی به دلیل پیچیدگی ذاتی و مسائلی مانند خطاها، ناسازگاری‌ها، مقادیر گم‌شده و کیفیت پایین داده‌ها، چالش‌های قابل توجهی در صنعت نفت و گاز ایجاد می‌کند. این عوامل به طور کلی منجر به تجسم‌های نادرست می‌شوند و بینش‌هایی که می‌توان از داده‌ها استخراج کرد را محدود می‌کنند. علاوه بر این، هنگامی که با مجموعه‌های داده چندبعدی بسیار بزرگ سروکار داریم، محدودیت‌های محاسباتی، مشکلات تجسم مؤثر اطلاعات را بیشتر تشدید می‌کند [۱۲۷]، [۱۲۸].

۴) راه‌حل‌های مقابله با چالش‌ها

مشکل تفسیر مدل‌های یادگیری ماشین مانع مهمی برای نفوذ این فناوری در صنعت نفت و گاز است، اما راه‌حل‌هایی برای غلبه بر آن وجود دارد. این راه‌حل‌ها شامل تحلیل اهمیت ویژگی‌ها است. این روش بررسی می‌کند که چگونه هر ویژگی در ورودی‌های مدل بر پیش‌بینی‌های آن تأثیر می‌گذارد. با تعیین اینکه کدام ویژگی‌ها بیشترین اهمیت را دارند، محققان می‌توانند درک بهتری از روابط بین داده‌ها و پیش‌بینی‌های مدل به دست آورند.

استفاده از **هوش مصنوعی تفسیرپذیر (Explainable AI)** می‌تواند تفسیرپذیری مدل را بهبود بخشد. تکنیک‌هایی مانند مقادیر (SHAP (SHapley Additive exPlanations)، LIME (Local Interpretability Model-agnostic Explanations)، اهمیت جایگشتی (Permutation Importance) و نمودار وابستگی جزئی (Partial Dependence Plot) درک عمیق‌تری از نحوه تعامل مدل با ویژگی‌های ورودی و تولید خروجی فراهم می‌کنند. با بررسی این نمودارها، محققان می‌توانند درک بهتری از تأثیر متغیرهای ورودی مختلف بر پیش‌بینی‌های مدل به دست آورند.

استفاده از **مدل‌های ترکیبی (Ensemble Models)** یک رویکرد جایگزین دیگر است. هدف از مدل ترکیبی ارائه پیش‌بینی‌های دقیق‌تر و قابل اعتمادتر با ترکیب چندین مدل یادگیری ماشین است. محققان می‌توانند با ترکیب مدل‌های مختلف که نقاط قوت و ضعف مکمل دارند، پیش‌بینی‌پذیری و شفافیت نتایج را بهبود بخشند.

رویکرد مفید دیگر، **تجسم مدل** است که شامل بررسی مکانیسم‌های داخلی یک مدل برای درک بهتر فرآیند پیش‌بینی آن است. تکنیک‌هایی مانند تجسم درخت تصمیم، بیشینه‌سازی فعال‌سازی (Activation Maximization) و نقشه‌برداری برجستگی (Saliency Mapping) بینش‌هایی در مورد ارتباطات و الگوهای پنهان در داده‌ها ارائه می‌دهند که پایه‌های دقت مدل هستند. با این حال، برای داده‌های چندبعدی، توجه ویژه‌ای لازم است تا پیچیدگی‌های



مرتبط با این تجسم‌ها به طور مؤثر مدیریت شود. اجرای فرآیندهای قوی مدیریت کیفیت داده‌ها برای اطمینان از دقت و بینش‌های معنادار ضروری است.

علاوه بر این، نیاز مبرمی به پیشرفت در قابلیت‌های محاسباتی برای تسهیل تجسم مجموعه‌های داده حتی بزرگ‌تر و پیچیده‌تر وجود دارد. با افزایش حجم و پیچیدگی داده‌ها در صنعت نفت و گاز، سرمایه‌گذاری در توسعه منابع محاسباتی که بتوانند تقاضای تجسم چنین مجموعه‌های داده عظیمی را برآورده کنند، ضروری است.

در نهایت، ترکیب این روش‌ها برای غلبه بر مشکل تفسیر مدل‌های یادگیری ماشین در اکتشاف نفت و گاز ضروری است تا درک بهتری از روابط بین داده‌ها و پیش‌بینی‌های مدل به دست آید. اگر مدل‌ها تفسیرپذیرتر باشند، محققان می‌توانند اطمینان بیشتری به پیش‌بینی‌های خود داشته باشند و از آن‌ها به طور مؤثرتری در اکتشاف نفت و گاز استفاده کنند. علاوه بر این، با بهره‌گیری از پیشرفت‌ها در تجسم، صنعت می‌تواند بینش‌های عمیق‌تری از داده‌های بزرگ و پیچیده خود به دست آورد که منجر به فرآیندهای تصمیم‌گیری آگاهانه‌تر و بهبود عملکرد کلی می‌شود.

C. تخصص حوزه‌ای

تخصص حوزه‌ای در این زمینه به معنای آشنایی با زمین‌شناسی و مهندسی پیچیده اکتشاف نفت و گاز است که از سال‌ها تجربه در این زمینه به دست می‌آید. علاوه بر این، تخصص در فناوری‌ها و فرآیندهای یادگیری ماشین برای پیاده‌سازی موثرترین تکنیک‌ها نیز ضروری است. دانش حوزه‌ای برای اطمینان از دقت و قابلیت اطمینان مدل‌های یادگیری ماشین در صنعت نفت و گاز حیاتی است.

(۱) تخصص در نفت و گاز

از آنجا که مدل‌های یادگیری ماشین به داده‌های ورودی وابسته هستند، نیاز به دانش حوزه‌ای مطرح می‌شود. داده‌هایی که در طول توسعه نفت و گاز جمع‌آوری می‌شوند، مانند بررسی‌های لرزه‌ای، لاگ‌های چاه و سوابق تولید، همگی نیاز به درک عمیق از اصول زمین‌شناسی و مهندسی دارند.

بدون دانش حوزه‌ای، تعیین اینکه کدام ویژگی‌های ورودی برای مدل‌های یادگیری ماشین مهم‌تر هستند و آیا آن‌ها به درستی فرآیندهای زمین‌شناسی یا مهندسی زیربنایی را نشان می‌دهند، دشوار خواهد بود. دانش حوزه‌ای برای کالیبره کردن و اعتبارسنجی مدل‌ها به منظور اطمینان از دقت پیش‌بینی‌های آن‌ها ضروری است.

۲) تخصص در یادگیری ماشین

تخصص در تکنیک‌های یادگیری ماشین نیز به اندازه تخصص در نفت و گاز مهم است. این تخصص اطمینان می‌دهد که تکنیک مناسب در زمان مناسب استفاده شود. رویکردهای یادگیری ماشین یکسان نیستند و هر مسئله و مجموعه داده نیاز به راه‌حل منحصر به فردی دارد. هر تصمیم، از انتخاب مدل تا روش بهینه‌سازی آن، حیاتی است. عملکرد یک مدل یادگیری ماشین ممکن است به پارامترهای ساختار مدل بستگی داشته باشد که باید با داده‌ها تطابق داشته باشد. بهینه‌سازی پارامترها در یک مدل یادگیری ماشین می‌تواند عملکرد مدل را بهبود بخشد. بنابراین، انتخاب بهترین روش‌های بهینه‌سازی، از جمله مدل‌ها و روش‌های ارزیابی، یک چالش بزرگ است که بر اثربخشی و قابلیت اطمینان مدل‌های یادگیری ماشین در صنعت تأثیر می‌گذارد.

۳) راه‌حل‌های مقابله با چالش‌ها

یکی از راه‌حل‌های غلبه بر این چالش، همکاری نزدیک با متخصصان حوزه‌ای مانند زمین‌شناسان، پتروفیزیک‌دان‌ها و مهندسان مخزن برای آموزش و ارزیابی مدل‌های یادگیری ماشین است. این همکاری می‌تواند به روش‌های مختلفی انجام شود، از جمله دریافت بازخورد از متخصصان حوزه‌ای در طول فرآیند ساخت مدل و اعتبارسنجی مدل‌ها بر اساس اصول شناخته‌شده زمین‌شناسی یا مهندسی.

استفاده از الگوریتم‌های یادگیری ماشین که به طور صریح دانش حوزه‌ای را در نظر می‌گیرند، یک گزینه عملی دیگر است. به عنوان مثال، برخی از محققان مدل‌های یادگیری ماشین مبتنی بر فیزیک را بررسی کرده‌اند که از اصول علمی ثابت‌شده برای بهبود دقت مدل و تفسیرپذیری آن توسط انسان استفاده می‌کنند.

به طور کلی، چالش دانش حوزه‌ای در استفاده از یادگیری ماشین در صنعت نفت و گاز بر نیاز به همکاری نزدیک بین مهندسان داده و متخصصان حوزه‌ای تأکید می‌کند تا اطمینان حاصل شود که مدل‌ها به درستی کالیبره، اعتبارسنجی و تفسیر شده‌اند. با همکاری، محققان می‌توانند مدل‌های یادگیری ماشین را توسعه دهند که بهتر با پیچیدگی‌ها و تخصص‌های صنعت نفت و گاز سازگار باشند.

VIII. نتیجه‌گیری

این مطالعه یک بررسی جامع و دقیق از کاربرد یادگیری ماشین در بخش بالادستی نفت و گاز، با تمرکز ویژه بر نقش آن در حوزه اکتشاف نفت و گاز ارائه می‌دهد. تلاش‌های تحقیقاتی ما شامل طیف وسیعی از منابع داده‌ای، از جمله مقالات تحقیقاتی بررسی‌شده دقیق، پایان‌نامه‌ها و بینش‌های ارائه‌شده در کنفرانس‌ها است. یک نگرانی قابل توجه که به طور مداوم با آن مواجه می‌شویم، کمبود داده‌های قابل دسترس برای مطالعه در این حوزه بسیار تخصصی است.



همان‌طور که بررسی ما نشان می‌دهد، الگوریتم‌های یادگیری ماشین توانایی قابل توجهی در پردازش داده‌های لرزه‌ای، طبقه‌بندی دقیق فاسیس و لیتوفاسیس‌ها و تخمین خواص پتروفیزیکی ضروری مانند اشباع آب، نفوذپذیری و تخلخل در طیف گسترده‌ای از سازندهای زمین‌شناسی نشان داده‌اند. طیف الگوریتم‌های مورد استفاده در این بررسی‌ها به طور چشمگیری متنوع است و شامل تکنیک‌های پایه‌ای مانند ANN، SVM، CNN، XGBoost، FL، FN و CM می‌شود. قابل توجه است که مدل‌های ترکیبی که چندین الگوریتم یا مدل یادگیری ماشین را با تکنیک‌های پیشرفته انتخاب ویژگی ترکیب می‌کنند، به طور مداوم دقت بالاتری نسبت به روش‌های مستقل ارائه می‌دهند.

با وجود این پیشرفت‌های امیدوارکننده، چندین چالش عمده باید برطرف شوند تا یادگیری ماشین بتواند به طور کامل در مرحله اکتشاف صنعت نفت و گاز به کار گرفته شود:

- **کیفیت و دسترسی به داده‌ها:** کیفیت و دسترسی به داده‌ها همچنان یک مانع بزرگ است. داده‌ها در اکتشاف اغلب محدود، ساختارنیافته، ناسازگار و ممکن است دارای عدم قطعیت‌هایی باشند. باید راه‌حل‌هایی برای بهبود کیفیت داده‌ها، افزایش اشتراک‌گذاری داده‌ها بین ذینفعان و استفاده از فناوری‌های نوظهور مانند سنسور از دور و جمع‌آوری داده‌های بلادرنگ توسعه یابد.
- **شفافیت و تفسیرپذیری:** ماهیت “جعبه سیاه” بسیاری از مدل‌های یادگیری ماشین چالش‌هایی را در درک نحوه رسیدن به پیش‌بینی‌ها ایجاد می‌کند. روش‌هایی برای افزایش شفافیت و تفسیرپذیری مدل‌ها، مانند تکنیک‌های هوش مصنوعی تفسیرپذیر (Explainable AI)، باید بیشتر مورد بررسی قرار گیرند و در شیوه‌های صنعتی ادغام شوند.
- **تخصص حوزه‌ای:** شکاف بین متخصصان یادگیری ماشین و متخصصان حوزه‌ای باید پر شود. همکاری بین دانشمندان داده، زمین‌شناسان، پتروفیزیک‌دان‌ها و مهندسان مخزن برای اطمینان از دقت مدل‌ها و همسویی آن‌ها با اصول زمین‌شناسی و مهندسی ضروری است.
- **ملاحظات اخلاقی و نظارتی:** همان‌طور که با هر فناوری دیگری، استفاده از یادگیری ماشین در صنعت نفت و گاز باید به استانداردهای اخلاقی و مقررات صنعت پایبند باشد. توجه به مسائل حریم خصوصی، امنیت و انطباق نظارتی برای کاربرد مسئولانه این ابزارهای قدرتمند ضروری است.

در نگاه به آینده، چندین جهت تحقیقاتی امیدوارکننده وجود دارد:

- **تقویت چارچوب‌های یادگیری ماشین مقاوم:** توسعه چارچوب‌های یادگیری ماشین مقاوم یک جهت تحقیقاتی مهم است. داده‌های اکتشاف نفت و گاز اغلب محدود، ضعیف و با عدم قطعیت‌های ذاتی همراه هستند. مسیر آینده یادگیری ماشین در این حوزه در ایجاد چارچوب‌های انعطاف‌پذیر و مقاوم نهفته است که

بتوانند حتی در مواجهه با چالش‌های داده‌های محدود، ضعیف و نامشخص، بینش‌های قابل اعتمادی ارائه دهند. چنین نوآوری‌هایی برای اطمینان از اثربخشی مداوم یادگیری ماشین در اکتشاف نفت و گاز ضروری است.

در نگاه به آینده، چندین جهت تحقیقاتی امیدوارکننده وجود دارد:

- **تقویت چارچوب‌های یادگیری ماشین مقاوم:** توسعه چارچوب‌های یادگیری ماشین مقاوم یک جهت تحقیقاتی مهم است. داده‌های اکتشاف نفت و گاز اغلب محدود، ضعیف و با عدم قطعیت‌های ذاتی همراه هستند. مسیر آینده یادگیری ماشین در این حوزه در ایجاد چارچوب‌های انعطاف‌پذیر و مقاوم نهفته است که بتوانند حتی در مواجهه با چالش‌های داده‌های محدود، ضعیف و نامشخص، بینش‌های قابل اعتمادی ارائه دهند. چنین نوآوری‌هایی برای اطمینان از اثربخشی مداوم یادگیری ماشین در اکتشاف نفت و گاز ضروری است.
- **ادغام یادگیری ماشین با فیزیک و مدل‌سازی عددی:** یکی دیگر از جهت‌های تحقیقاتی امیدوارکننده، ادغام یادگیری ماشین با مدل‌سازی فیزیکی و عددی است. این رویکرد ترکیبی می‌تواند از مزایای هر دو روش بهره‌مند شود و مدل‌هایی ایجاد کند که هم دقیق‌تر و هم تفسیرپذیرتر باشند. به عنوان مثال، مدل‌های فیزیکی مبتنی بر یادگیری ماشین می‌توانند از قوانین فیزیکی شناخته‌شده برای بهبود پیش‌بینی‌ها و کاهش وابستگی به داده‌های آموزشی استفاده کنند.
- **بهبود روش‌های پیش‌پردازش و مهندسی ویژگی‌ها:** با توجه به پیچیدگی و ناهمگنی داده‌های اکتشاف نفت و گاز، توسعه روش‌های پیش‌پردازش و مهندسی ویژگی‌های پیشرفته می‌تواند به طور قابل توجهی دقت مدل‌های یادگیری ماشین را بهبود بخشد. این شامل تکنیک‌هایی برای مدیریت داده‌های پرت، نویز و عدم قطعیت‌ها، و همچنین استخراج ویژگی‌های معنادار از داده‌های پیچیده است.
- **استفاده از یادگیری عمیق و شبکه‌های عصبی پیشرفته:** با پیشرفت‌های اخیر در یادگیری عمیق، استفاده از شبکه‌های عصبی پیچیده‌تر مانند شبکه‌های عصبی کانولوشنال (CNN) و شبکه‌های عصبی بازگشتی (RNN) می‌تواند به بهبود دقت مدل‌ها در کاربردهای اکتشاف نفت و گاز کمک کند. این مدل‌ها می‌توانند الگوهای پیچیده در داده‌های لرزه‌ای و لاگ‌های چاه را بهتر تشخیص دهند.



- **توسعه مدل‌های ترکیبی و هیبرید:** مدل‌های ترکیبی که از چندین الگوریتم یادگیری ماشین استفاده می‌کنند، می‌توانند به طور قابل توجهی دقت و قابلیت اطمینان پیش‌بینی‌ها را افزایش دهند. این مدل‌ها می‌توانند از نقاط قوت هر الگوریتم بهره‌مند شوند و نقاط ضعف آن‌ها را جبران کنند.
- **افزایش شفافیت و تفسیرپذیری مدل‌ها:** با توجه به ماهیت “جعبه سیاه” بسیاری از مدل‌های یادگیری ماشین، توسعه روش‌هایی برای افزایش شفافیت و تفسیرپذیری مدل‌ها یک جهت تحقیقاتی مهم است. این شامل استفاده از تکنیک‌هایی مانند SHAP، LIME و سایر روش‌های تفسیر مدل است که به متخصصان حوزه‌ای کمک می‌کند تا درک بهتری از پیش‌بینی‌های مدل داشته باشند.
- **استفاده از داده‌های بلادرنگ و اینترنت اشیا (IoT):** با پیشرفت فناوری‌های جمع‌آوری داده‌های بلادرنگ و اینترنت اشیا، استفاده از این داده‌ها در مدل‌های یادگیری ماشین می‌تواند به بهبود دقت و به‌روزرسانی مدل‌ها در زمان واقعی کمک کند. این امر به ویژه در عملیات اکتشاف و تولید نفت و گاز که نیاز به تصمیم‌گیری سریع و دقیق دارند، بسیار مفید خواهد بود.
- **توجه به مسائل اخلاقی و محیط زیستی:** با افزایش استفاده از یادگیری ماشین در صنعت نفت و گاز، توجه به مسائل اخلاقی و محیط زیستی نیز اهمیت بیشتری پیدا می‌کند. این شامل توسعه مدل‌هایی است که نه تنها دقت بالایی دارند، بلکه از نظر محیط زیستی پایدار و از نظر اخلاقی مسئولانه هستند.
- در نهایت، یادگیری ماشین پتانسیل بالایی برای تحول صنعت نفت و گاز دارد، به ویژه در مرحله اکتشاف که نیاز به دقت و سرعت بالا در تصمیم‌گیری‌ها وجود دارد. با این حال، برای دستیابی به این پتانسیل، باید چالش‌های مربوط به کیفیت داده‌ها، تفسیرپذیری مدل‌ها، تخصص حوزه‌ای و مسائل اخلاقی و نظارتی به طور جدی مورد توجه قرار گیرند. با ادامه تحقیقات و توسعه در این زمینه، یادگیری ماشین می‌تواند به یکی از ابزارهای کلیدی در بهبود کارایی و کاهش هزینه‌ها در صنعت نفت و گاز تبدیل شود.
- **تجسم پیشرفته:** نوآوری در تکنیک‌های تجسم داده‌ها از اهمیت بالایی برخوردار است، به‌ویژه برای مدیریت داده‌های پیچیده و چندبعدی در صنعت نفت و گاز. محققان باید بر روش‌های تحلیل بصری تمرکز کنند که امکان استخراج بینش‌های معنادار از مجموعه داده‌های بزرگ و پیچیده را فراهم می‌کنند، زیرا قابلیت‌های محاسباتی به‌طور مداوم در حال پیشرفت هستند.
- **همکاری میان‌رشته‌ای:** تشویق همکاری بین دانشگاه‌ها، صنعت و نهادهای نظارتی می‌تواند پیشرفت را تسریع کند. اشتراک‌گذاری داده‌ها، تأمین مالی تحقیقات و توسعه استانداردهای مشترک می‌تواند به حل مسائل مربوط به کیفیت و دسترسی به داده‌ها کمک کند.

- **ابزارهای انطباق با مقررات:** توسعه ابزارها و چارچوب‌هایی که به عبور از پیچیدگی‌های محیط نظارتی صنعت نفت و گاز کمک می‌کنند، ضروری است. این ابزارها باید ضمن تسهیل انطباق با مقررات، امنیت و حریم خصوصی داده‌ها را نیز تضمین کنند.

- **قابلیت‌های محاسباتی:** سرمایه‌گذاری مستمر در منابع محاسباتی برای مدیریت حجم فزاینده داده‌ها و پاسخگویی به نیازهای محاسباتی الگوریتم‌های یادگیری ماشین حیاتی است. این شامل بررسی راه‌حل‌هایی مانند رایانش ابری، رایانش توزیع‌شده و رایانش با عملکرد بالا می‌شود.

این بررسی سهم قابل توجهی در درک چالش‌های منحصر به فرد اعمال یادگیری ماشین در مرحله اکتشاف صنعت نفت و گاز دارد، مانند عدم قطعیت در پارامترهای اکتشاف، تفاوت‌های مقیاس و پیچیدگی‌های مدیریت داده‌های زمانی و مکانی. به‌طور قابل توجهی، این بررسی فراتر از شناسایی چالش‌ها می‌رود و راه‌حل‌های بالقوه ارائه می‌دهد، روش‌هایی که به دستیابی به دقت بهینه کمک می‌کنند را شناسایی می‌کند و جهت‌های تحقیقاتی آینده را ترسیم می‌کند. این تحلیل جامع، نقشه‌راهی برای غلبه بر چالش‌ها و غنی‌سازی پایگاه دانش برای محققان و ذینفعان صنعت ارائه می‌دهد.



شرکت صبا (دیجی هوش هما)

راه های تماس با ما

ایمیل cmo@sabaind.com

تلگرام [@DigiHooshHoma](https://www.linkedin.com/company/maaia/)

لینکداین <https://www.linkedin.com/company/maaia/>

وب سایت www.sabaind.com

موبایل +98-0905 611 0895

تلفن +98-021-22146343

آدرس تهران - سعادت آباد - میدان بهرود -
خیابان عابدی - ساختمان صبا

معرفی تیم دیجی هوش هما (صبا)

□ ماموریت تیم دیجی هوش هما، آشناسازی اعضای ممتزم هیات مدیره، مدیران عامل، مدیران ارشد، و متفحصین سازمانها برای چگونگی پیادهسازی و استفاده از فناوریهای هوشمندسازی، و به ویژه هوش مصنوعی به عنوان فناوری شالوده شکن عصر حاضر، در سازمان یا دپارتمان مدنظر آنها می باشد.

□ ما باور داریم که امروزه تحول دیجیتال و پیادهسازی هوش مصنوعی در سازمان ها نه یک انتخاب، بلکه یک ضرورت است، چرا که هزینه عدم استفاده از آن برای شرکت زیان بار خواهد بود. ما در مجموعه دیجی هوش هما (صبا) با افتخار آماده ایم تا شما را در این مسیر همراهی کنیم.

□ مجموعه دیجی هوش هما (صبا) در نقش مشاوره، آموزش، و ارائه راهمهای هوش مصنوعی به شما و سازمان شما متعهد بوده و پیادهسازی بهینه راهمهای هوش مصنوعی را با کمک برترین متفحصین و شرکتهای ارائه دهنده خدمات هوش مصنوعی برای شما رقم خواهد زد.

□ تکنولوژی های هوشمندسازی و هوش مصنوعی می توانند تغییرات اساسی در نحوه کارکرد و بهره وری سازمان ها به وجود آورند. با استفاده از راهم های ما، می توانید فرایندهای پیچیده را ساده سازی کنید، تصمیم گیری های بهتری انجام دهید و بهره وری سازمان خود را به شکل چشمگیری افزایش دهید.

□ چشم انداز ما در دیجی هوش هما (در نقش اپراتور هوش مصنوعی AI Integrator)، ایجاد تحولی واقعی و ملموس در سازمان های متقاضی است. ما با تیمی مجرب و متفحص، آماده ایم تا بهترین مشاوره ها و راهم های هوشمند را به شما ارائه دهیم. از مشاوره، آموزش های درون سازمانی تا پیاده سازی کامل راهم های هوش مصنوعی و اینترنت اشیا، ما در تمامی مراحل همراه شما هستیم.

□ مدیران تاثیرگذار با شهادت دنبال سافت آینده ای هوشمندتر و پر رونق تر برای مملکت خود هستند. برای دریافت اطلاعات بیشتر و مشاوره تخصصی، با ما تماس بگیرید.