

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی معدن، نفت و ژئوفیزیک

گروه مهندسی نفت، حفاری و بهره برداری

پایان نامه کارشناسی ارشد

پیش بینی و جلوگیری از مشکل گیر رشته حفاری با استفاده از شبکه عصبی و

سری های زمانی

سجاد اسکندریان

اساتید راهنما:

دکتر بهزاد تخمچی

دکتر وامق رسولی

خرداد ۱۳۹۵



مدیریت تحصیلات تکمیلی

بسمه تعالی

شماره: ۲۰۶۹۵/۱۳۳
تاریخ: ۹۵/۳/۲۲
ویرایش:

فرم شماره ۷: صورتجلسه دفاع از پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد آقای سجاد اسکندریان به شماره دانشجویی ۹۲۳۳۷۴۴ رشته مهندسی نفت گرایش حفاری و بهره برداری تحت عنوان پیش بینی و جلوگیری از مشکل گیر رشته حفاری با استفاده از شبکه عصبی و سری های زمانی که در تاریخ ۱۳۹۵/۳/۱۷ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

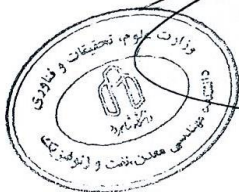
قبول (با درجه : عالی - امتیاز ۱۹.۵)	☐ دفاع مجدد	☐ مردود
-------------------------------------	------------------------------------	--------------------------------

۲- بسیار خوب (۱۸ - ۱۸/۹۹)
۴- قابل قبول (۱۵ - ۱۵/۹۹)

۱- عالی (۲۰ - ۱۹)
۳- خوب (۱۶ - ۱۷/۹۹)
۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	دکتر بهزاد تخم چی	دانشیار	
۲- استاد راهنمای دوم	دکتر وامق رسولی	استاد	
۳- استاد مشاور			
۴- نماینده شورای تحصیلات تکمیلی	دکتر سوسن ابراهیمی	استادیار	
۵- استاد ممتحن اول	دکتر مهرداد سلیمانی منفرد	استادیار	
۶- استاد ممتحن دوم	دکتر امین روشندل کاهو	استادیار	

رئیس دانشکده: دکتر رضا عرب امیری



تقدیم به جلوه‌ی مهر پروردگار،

پدر و مادری که وجودم بر ایشان همه رنج بود و وجودشان برایم همه مهر؛

دو فرشته محافظی که از نگاهشان صلابت، از رفتارشان محبت و از صبرشان
ایستادگی را آموختم.

تشکر و قدردانی

با حمد و سپاس از پروردگار مَنّان، از تمامی اساتید، کارشناسان و دوستانی که در تمامی مراحل این پژوهش مرا یاری نمودند، تشکر و قدردانی می‌نمایم.

بر خود لازم می‌دانم تا به‌طور ویژه از استاد گرانقدرم جناب آقای دکتر تخم‌چی و جناب آقای دکتر رسولی به‌عنوان اساتید راهنمای این اثر تشکر نمایم.

همچنین از مهندس جهانگیر حمیدی به‌عنوان مشاور صنعتی این پایان‌نامه و پرسنل محترم مدیریت اکتشاف ملی نفت ایران، خصوصاً ادارات پژوهش و فناوری و مهندسی حفاری، کمال تشکر را دارم.

با آرزوی سلامتی و موفقیت این بزرگواران

سجاد اسکندریان، خرداد ۱۳۹۵

تعهد نامه

اینجانب سجاد اسکندریان دانشجوی دوره کارشناسی ارشد رشته مهندسی نفت، حفاری و بهره برداری دانشکده مهندسی معدن، نفت و ژئوفیزیک دانشگاه صنعتی شاهرود نویسنده پایان نامه پیش بینی و جلوگیری از مشکل گیر رشته حفاری با استفاده از شبکه عصبی و سری های زمانی تحت راهنمایی دکتر بهزاد تخمچی و دکتر وامق رسولی متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تا کنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافت های آن ها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

گیر رشته‌ی حفاری، یکی از مشکلات رایج در صنعت حفر چاه‌های نفت و گاز می‌باشد. وقوع این پدیده می‌تواند منجر به بروز مشکلات عدیده‌ای، از اتلاف وقت و هزینه تا از دست رفتن چاه، شود. در این پژوهش سعی گردید تا با ساخت مدلی، وقوع گیر رشته‌ی حفاری پیش‌بینی شود. برای ساخت این مدل از پارامترهای عملیات و سیال حفاری، مربوط به حفر چاه‌های یکی از میادین عظیم جنوب غربی ایران، استفاده شد. از ۹۷۳ مجموعه داده، که هر یک شامل پارامترهای مختلف عملیات و سیال حفاری است، ۱۲۶ مورد مربوط به داده‌های وقوع گیر رشته‌ی حفاری و مابقی مربوط به داده‌های عدم وقوع گیر می‌باشند. در این پایان‌نامه ابتدا داده‌ها توسط الگوریتم‌های دسته‌بندی مانند الگوریتم K نزدیک‌ترین همسایه و ماشین بردار پشتیبان، بر اساس الگوهای گیر و عدم گیر دسته‌بندی شدند. الگوریتم K نزدیک‌ترین همسایه ۹۳/۶ درصد و ماشین بردار پشتیبان ۹۷ درصد در شناسایی دسته‌های گیر و عدم گیر موفق بودند. سپس توسط شبکه عصبی پرسپترون چند لایه و سری‌های زمانی تحت شبکه عصبی، وقوع گیر رشته‌ی حفاری شناسایی و پیش‌بینی گردید. از سری زمانی به‌عنوان یکی از روش‌های دینامیکی جهت پیش‌بینی گیر رشته‌ی حفاری بر اساس داده‌های روزهای پیشین استفاده شد. شبکه عصبی پرسپترون چند لایه ۹۴/۹ درصد و سری زمانی ۹۴/۲ درصد، گیر و عدم گیر را برای داده‌های آزمون به‌درستی پیش‌بینی نمودند. با اطلاع از احتمال وقوع پدیده‌ی گیر رشته‌ی حفاری، امکان جلوگیری از آن به‌وسیله‌ی تغییر پارامترهای مؤثر فراهم خواهد شد.

کلمات کلیدی: گیر رشته‌ی حفاری، شبکه عصبی پرسپترون چند لایه، سری‌های زمانی، الگوریتم K نزدیک‌ترین همسایه، ماشین بردار پشتیبان.

لیست مقالات مستخرج از پایان نامه

- ❖ اسکندریان س.، تخمچی ب.، رسولی و.، حمیدی ج.، (۱۳۹۵) "به کارگیری ماشین بردار پشتیبان (SVM) در بررسی و تشخیص پدیده گیر رشته حفاری، به همراه تحلیل مؤلفه اصلی جهت بهبود عملکرد"، سومین کنفرانس بین المللی نفت، گاز، پالایش و پتروشیمی با رویکرد توسعه ارتباط بین دولت، دانشگاه و صنعت، تهران، ایران.
- ❖ اسکندریان س.، تخمچی ب.، رسولی و.، حمیدی ج.، (۱۳۹۵) "تخمین نرخ نفوذ مته حفاری (Bit ROP) با در نظر گرفتن پارامترهای حفاری، سیال حفاری و نگرش بر لیتولوژی منطقه با به کارگیری شبکه عصبی مصنوعی (ANN)"، سومین کنفرانس بین المللی نفت، گاز، پالایش و پتروشیمی با رویکرد توسعه ارتباط بین دولت، دانشگاه و صنعت، تهران، ایران.

فهرست مطالب

فصل اول: کلیات	۱
۱-۱ تعریف مسأله	۲
۱-۱-۱ طبقه‌بندی انواع مکانیسم‌های گیر رشته‌ی حفاری	۲
۱-۱-۲ طبقه‌بندی گیر رشته حفاری به روشی دیگر	۴
۱-۲-۱-۱ گیر مکانیکی	۵
۱-۲-۲-۱ گیر اختلاف فشاری	۵
۱-۳-۱ مهم‌ترین وظایف سیالات حفاری	۸
۱-۴-۱ پارامترهای مؤثر گل حفاری بر گیر اختلاف فشاری	۸
۱-۵-۱ فاکتورهای افزایش دهنده‌ی نیروی گیر اختلاف فشاری	۹
۱-۶-۱ روش‌های تعیین نقطه آزادسازی	۹
۱-۶-۱-۱ روش کشش	۱۰
۱-۶-۱-۲ ابزار تعیین نقطه آزاد (FPIT)	۱۰
۲-۱ تاریخچه تحقیق	۱۱
۳-۱ رویکرد این پایان‌نامه در حل مسأله	۱۵
۴-۱ ساختار پایان‌نامه	۱۶

فصل دوم: داده‌های تحقیق ۱۷

۱-۲ داده‌ها و پارامترهای ورودی ۱۸

۲-۲ معرفی پارامترهای عملیات و سیال حفاری ۱۹

۳-۲ سازندها و لیتولوژی ۲۲

۱-۳-۲ ناحیه‌ی زاگرس ۲۲

۱-۱-۳-۲ دشت خوزستان ۲۳

۲-۱-۳-۲ زاگرس چین‌خورده ۲۳

۳-۱-۳-۲ زاگرس رورانده ۲۳

۲-۳-۲ سازندهای مهم ناحیه‌ی زاگرس ۲۴

فصل سوم: روش‌شناسی تحقیق ۲۷

۱-۳ مقدمه ۲۸

۲-۳ الگوریتم K نزدیک‌ترین همسایه ۲۸

۱-۲-۳ قانون نزدیک‌ترین همسایه ۳۰

۳-۳ ماشین بردار پشتیبان ۳۱

۱-۳-۳ روش عملکرد ماشین بردار پشتیبان در دسته‌بندی ۳۲

- ۳-۴ معیار ارزیابی روش‌های دسته‌بندی..... ۳۵
- ۳-۵ تحلیل مؤلفه‌های اصلی (PCA) ۳۶
- ۳-۶ شبکه عصبی مصنوعی ۳۷
- ۳-۶-۱ ساختار شبکه‌های عصبی ۳۸
- ۳-۶-۲ مدل ریاضی یک نرون در شبکه عصبی مصنوعی ۴۰
- ۳-۶-۲-۱ تابع ترکیب ۴۰
- ۳-۶-۲-۲ تابع تحریک ۴۲
- ۳-۶-۳ الگوریتم یادگیری ۴۴
- ۳-۶-۳-۱ الگوریتم یادگیری نظارتی (با سرپرست) ۴۴
- ۳-۶-۳-۲ الگوریتم یادگیری تقویت‌شده یا تشدید ۴۶
- ۳-۶-۳-۳ الگوریتم یادگیری غیرنظارتی (بدون سرپرست) ۴۶
- ۳-۶-۴ انتخاب داده‌های آموزشی ۴۷
- ۳-۶-۵ اهمیت نسبی ورودی‌های شبکه ۴۷
- ۳-۶-۶ معیارهای پایان آموزش شبکه ۴۷
- ۳-۷ سری‌های زمانی ۴۸
- ۳-۷-۱ انواع سری زمانی ۴۹

۴۹..... ۲-۷-۳ اجزای تشکیل دهنده‌ی سری زمانی

۴۹..... ۱-۲-۷-۳ روند

۵۰..... ۲-۲-۷-۳ تغییرات فصلی

۵۰..... ۳-۲-۷-۳ تغییرات دوره‌ای

۵۰..... ۴-۲-۷-۳ تغییرات نامنظم

۵۱..... ۳-۷-۳ الگوسازی برای یک سری زمانی

۵۱..... ۸-۳ پیاده‌سازی مدل‌های شبکه عصبی و سری زمانی تحت شبکه عصبی، در نرم افزار.....

۵۲..... ۱-۸-۳ بارگذاری داده‌ها

۵۲..... ۲-۸-۳ تقسیم داده‌ها

۵۳..... ۳-۸-۳ نرمال‌سازی داده‌ها

۵۳..... ۴-۸-۳ ساختار شبکه

۵۴..... ۵-۸-۳ آموزش شبکه

۵۴..... ۶-۸-۳ ارزیابی شبکه

۵۵..... فصل چهارم: نتایج شناسایی و پیش‌بینی گیر رشته‌ی حفاری

۵۶..... ۱-۴ مقدمه

۲-۴ پردازش بر روی داده‌ها	۵۶
۳-۴ نتایج الگوریتم K نزدیک‌ترین همسایه	۶۶
۴-۴ نتایج الگوریتم ماشین بردار پشتیبان	۶۷
۵-۴ نتایج الگوریتم‌های دسته‌بندی به همراه تحلیل مؤلفه‌های اصلی	۶۹
۶-۴ تشخیص الگو و پیش‌بینی به‌وسیله‌ی شبکه عصبی	۷۲
۶-۴-۱ عملکرد شبکه عصبی با افزودن پارامتر لیتولوژی	۷۴
۷-۴ پیش‌بینی به‌وسیله‌ی سری زمانی	۸۰
۸-۴ بررسی پارامترهای ورودی به مدل	۸۵
فصل پنجم: نتیجه‌گیری و پیشنهادات	۸۷
۵-۱ نتیجه‌گیری	۸۸
۵-۲ پیشنهادات	۸۹
منابع	۹۱

فهرست اشکال

- شکل ۱-۱ گیر رشته‌ی حفاری در اثر تجمع خرده‌های حفاری ۳
- شکل ۱-۲ شرایط رشته‌ی حفاری در حالت نرمال (عدم وقوع گیر) و وقوع گیر ۶
- شکل ۱-۳ ابرصفحه با حداکثر مرز جداکننده (خطوط) به همراه مرزهای جداکننده (خط‌چین‌ها) ۳۲
- شکل ۲-۳ نحوه ساخت ابرسطح جداکننده بین دو طبقه داده در فضای دو بعدی ۳۳
- شکل ۳-۳ نمایش طرح کلی یک نرون ۴۱
- شکل ۴-۳ نمایش تابع تحریک پله‌ای ۴۲
- شکل ۵-۳ نمایش تابع تحریک خطی ۴۳
- شکل ۶-۳ نمایش تابع تحریک سیگموئیدی ۴۳
- شکل ۷-۳ نمایش تابع تحریک تانژانت هایپربولیک ۴۳
- شکل ۸-۳ نمایش مدل سری زمانی تحت شبکه عصبی، در نرم افزار متلب ۵۴
- شکل ۱-۴ نمودار هیستوگرام پارامترهای عملیات حفاری ۵۸
- شکل ۲-۴ نمودار هیستوگرام پارامترهای حفاری مرتبط با مسیر چاه ۶۱
- شکل ۳-۴ نمودار هیستوگرام پارامترهای سیال حفاری ۶۲

شکل ۴-۴ ماتریس طبقه‌بندی نتایج KNN ۶۷

شکل ۴-۵ ماتریس طبقه‌بندی نتایج SVM ۶۸

شکل ۴-۶ ماتریس طبقه‌بندی نتایج KNN ارتقاء یافته با PCA ۶۹

شکل ۴-۷ ماتریس طبقه‌بندی نتایج SVM ارتقاء یافته با PCA ۶۹

شکل ۴-۸ نمایش مواردی که الگوریتم SVM در دسته بندی آن‌ها ناموفق بود، در برابر افزایش عمق ۷۱

شکل ۴-۹ مدل شبکه عصبی پرسپترون چند لایه ۷۳

شکل ۴-۱۰ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP برای داده‌های آموزش ۷۳

شکل ۴-۱۱ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP برای داده‌های آزمون ۷۴

شکل ۴-۱۲ نمودار هیستوگرام معادل عددی لیتولوژی سازندهای میدان مورد مطالعه ۷۷

شکل ۴-۱۳ نمایش مدل شبکه عصبی MLP با افزودن پارامتر معادل عددی لیتولوژی ۷۸

شکل ۴-۱۴ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP با ۲۱ پارامتر ورودی برای داده‌های

آموزش ۷۸

شکل ۴-۱۵ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP با ۲۱ پارامتر ورودی برای داده‌های

آزمون ۷۹

شکل ۴-۱۶ نمایش مدل سری زمانی ساخته شده ۸۱

شکل ۴-۱۷ ماتریس طبقه‌بندی نتایج مدل سری زمانی برای داده‌های آموزش ۸۱

شکل ۴-۱۸ ماتریس طبقه‌بندی نتایج مدل سری زمانی برای داده‌های آزمون ۸۲

شکل ۴-۱۹ نمایش مدل سری زمانی با تأخیر زمانی ۲..... ۸۳

شکل ۴-۲۰ ماتریس طبقه‌بندی نتایج مدل سری زمانی با تأخیر زمانی ۲، برای داده‌های آموزش..... ۸۳

شکل ۴-۲۱ ماتریس طبقه‌بندی نتایج مدل سری زمانی با تأخیر زمانی ۲، برای داده‌های آزمون..... ۸۴

شکل ۴-۲۲ نمایش میزان همبستگی پارامترهای ورودی بر خروجی..... ۸۶

فهرست جداول

- جدول ۱-۲ طبقه‌بندی چاه‌ها بر اساس عمق حفاری شده نهایی..... ۱۸
- جدول ۲-۲ نمایش لیتولوژی هر سازند و عمق شروع سازند در میدان..... ۲۴
- جدول ۱-۴ مقادیر حداقل، حداکثر و میانگین پارامترهای ورودی..... ۵۶
- جدول ۲-۴ مشخصات و ساختار الگوریتم K نزدیک‌ترین همسایه ۶۶
- جدول ۳-۴ اطلاعات ساختار الگوریتم ماشین بردار پشتیبان..... ۶۸
- جدول ۴-۴ مقایسه نتایج الگوریتم‌های KNN و SVM..... ۷۰
- جدول ۵-۴ ساختار و ویژگی‌های مدل شبکه‌ی عصبی MLP..... ۷۲
- جدول ۶-۴ معادل عددی تعیین شده برای هر سنگ..... ۷۵
- جدول ۷-۴ معادل عددی برای ترکیبات سنگی مختلف..... ۷۶
- جدول ۸-۴ مشخصات مدل سری زمانی..... ۸۰
- جدول ۹-۴ مقایسه نتایج مدل‌های شبکه عصبی MLP و سری زمانی..... ۸۵

فصل اول: کلیات

۱-۱ تعریف مسأله

۲-۱ تاریخچه تحقیق

۳-۱ رویکرد این پایان نامه در حل مسأله

۴-۱ ساختار پایان نامه

۱-۱ تعریف مسأله

گیر رشته‌ی حفاری یکی از رایج‌ترین و مهم‌ترین مشکلات حفاری است. این پدیده از مشکلات کوچکی که موجب افزایش نسبی هزینه‌ها می‌شوند تا موارد پیچیده‌ای که می‌توانند منجر به نتایج نامطلوبی مانند از دست دادن قسمتی از رشته‌ی حفاری یا چاه به‌صورت کامل گردد، رده‌بندی می‌شوند [۱].

گیر رشته‌ی حفاری را می‌توان به صورت نیروی درون چاهی که مانع چرخش و خروج لوله‌ها از چاه می‌شود، تعریف نمود. همچنین محدودیت حرکت رشته‌ی حفاری به سبب نیروهای درون‌چاهی، تنگی چاه نامیده می‌شود. لذا می‌توان گیر رشته حفاری را یکی از عمده‌ترین مشکلات در حفر چاه‌های نفتی به حساب آورد که نه تنها باعث اتلاف وقت جهت عملیات آزادسازی می‌شود، بلکه موجب ایجاد هزینه‌های اضافی یا در صورت عدم موفقیت در آزادسازی، سبب از دست دادن چاه می‌شود. این مشکل همواره یک خطر اصلی برای عملیات حفاری، مخصوصاً در حفر چاه‌های جهت‌دار که دارای هزینه‌های بالایی هستند، به حساب می‌آید. بنابراین برای جلوگیری از این مشکل و حذف هزینه‌های ناشی از آن، باید به دنبال راهی جهت ممانعت از وقوع این پدیده بود [۲].

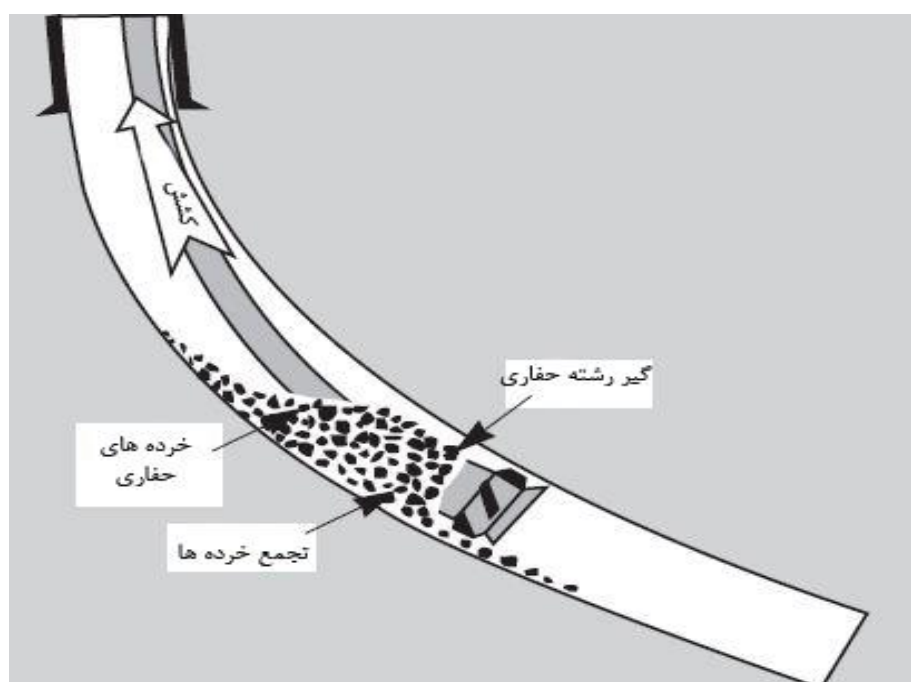
۱-۱-۱ طبقه‌بندی انواع مکانیسم‌های گیر رشته‌ی حفاری

گیر رشته‌ی حفاری را بر اساس مکانیسم وقوع آن، می‌توان به سه دسته تقسیم نمود. دسته اول گیرهای ناشی از ته‌نشین شدن و ریزش مواد جامد موجود در گل می‌باشند که عبارتند از:

- ته‌نشین شدن کنده‌های حفاری و مواد جامد موجود در گل
- ریزش شیل‌های دیواره چاه و آماس‌پذیری آن

- ریزش سازندهای سست
- ریزش سازندهای دارای شکستگی
- گیرهای مرتبط با سیمان کاری
- گیرهای مرتبط با مانده^۱ درون چاه

در شکل ۱-۱ طرح کلی گیر رشته‌ی حفاری ناشی از تجمع خرده‌های حفاری نمایش داده شده است.



شکل ۱-۱ گیر رشته‌ی حفاری در اثر تجمع خرده‌های حفاری [۱].

دسته دوم گیرهای ناشی از نیروهای تفاضل فشاری می‌باشند و دسته سوم گیرهای مرتبط با هندسه چاه می‌باشند که عبارتند از:

- مجموعه‌ی ته‌چاهی^۱ خشک و غیرقابل انعطاف
- جاکلیدی^۲
- سگ‌دست^۳
- سازندهای متحرک
- حفره کم سایز
- مچاله شدن لوله‌های جداری^۴ [۳].

۱-۱-۲ طبقه‌بندی گیر رشته حفاری به روشی دیگر

به صورت کلی‌تر، گیر رشته‌ی حفاری را به دو دسته‌ی کلی تقسیم می‌کنند؛ یکی گیرهای مکانیکی و دیگری گیر اختلاف فشاری است. تشخیص گیر لوله‌ها به صورت اختلاف فشاری به این صورت است که چرخش لوله و حرکت آن به سمت بالا و پایین امکان‌پذیر نیست، اما هنوز گردش گل حفاری امکان‌پذیر است و این در تناقض با گیر لوله‌ها به صورت مکانیکی است [۴].

1 Bottom Hole Assembly (BHA)
2 key seat
3 dog leg
4 collapsed casing

۱-۲-۱-۱ گیر مکانیکی

عمدتاً به دلیل هم‌آمدن و سد شدن مکانیکی یا اختلالات هندسی چاه مانند حفره جاکلیدی و سگ‌دست رخ می‌دهد. اما گیر اختلاف فشاری به دلیل نیروی ناشی از اختلاف فشار هیدرواستاتیک ستون گل و فشار سازند به وجود می‌آید. این در شرایطی است که رشته‌ی حفاری در مقابل سازند تراوا که اندود گل بر روی آن تشکیل شده، قرار می‌گیرد [۱، ۵ و ۶].

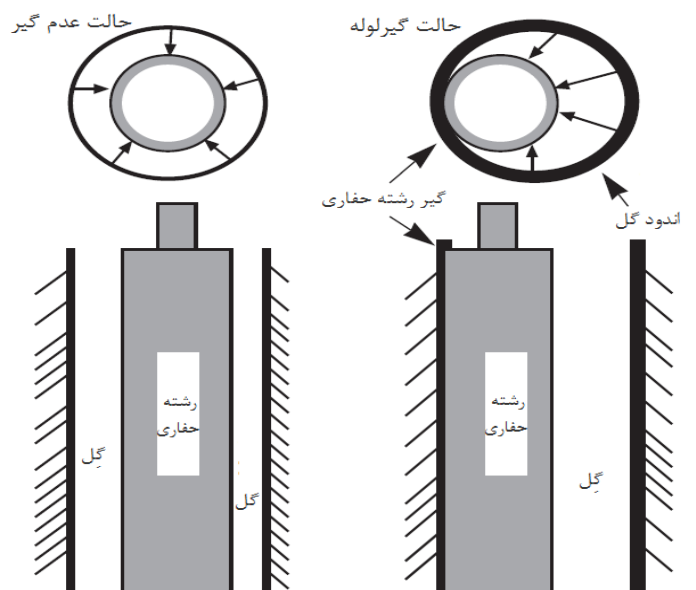
گیر مکانیکی معمولاً هنگامی اتفاق می‌افتد که رشته‌ی حفاری در حال حرکت است و به‌طور معمول با عدم گردش گل مشخص می‌شود. بعضی اوقات، به هر حال یک مقدار محدود حرکت بالا و پایین با کمی دوران آزاد، هنگام گیر مکانیکی لوله ملاحظه می‌گردد. گیر اختلاف فشاری زمانی که لوله ساکن است اتفاق می‌افتد. مانند هنگامی که اتصالات برقرار می‌گردد یا وقتی که مسیریابی صورت می‌گیرد. گیر اختلاف فشاری زمانی مشخص می‌گردد که هیچ حرکت بالا یا پایین یا حرکت دورانی آزاد وجود ندارد. اما گردش گل هنوز امکان‌پذیر است و این در شرایطی است که یک طرف لوله گیر افتاده باشد. اما در شرایط گیر کامل نه تنها امکان دوران و جابه‌جایی لوله حفاری وجود ندارد، بلکه گردش گل هم ممکن نیست [۴، ۷ و ۸]. به دلیل اهمیت گیر اختلاف فشاری، در ادامه مکانیسم، شرایط و عوامل مؤثر بر آن مورد بررسی قرار می‌گیرد.

۱-۲-۱-۲ گیر اختلاف فشاری^۱

این نوع گیر عمدتاً در سازندهایی با تراوایی بالا و دارای شکستگی مانند ماسه‌سنگ، سنگ آهک و دولومیت و در شرایطی که فشار هیدرواستاتیک گل از فشار سازند بیش‌تر باشد، رخ می‌دهد. اگر رشته‌ی حفاری

1 differential pressure sticking

(عموماً مجموعه‌ی ته چاهی) به علت ضخیم بودن اندود گل (که خود در اثر بالا بودن مواد جامد گل و هرزروی سیال حفاری بالا به وجود می‌آید)، با دیواره‌ی اندود گل تماس حاصل کند (مثلاً به علت زاویه‌ی چاه یا ناپایداری رشته‌ی حفاری) و حرکت رشته و گردش گل در مقابل سازند تراوا متوقف شود، ضخامت اندود گل افزایش یافته و دور رشته را نشت‌بندی می‌کند. همین امر باعث عدم انتقال فشار هیدرواستاتیک گل به ناحیه‌ای از رشته که در تماس با دیواره است، می‌شود. در نتیجه نیروی ناشی از اختلاف فشار بین هیدرواستاتیک گل و سازند کم فشارتر، رشته را کاملاً به دیواره چاه می‌چسباند (شکل ۱-۲). در این حالت محل نشت‌بندی شده به مرور زمان شدیداً گسترش یافته و رشته در اندود گل مدفون می‌شود.



شکل ۱-۲ شرایط رشته‌ی حفاری در حالت نرمال (عدم وقوع گیر) و وقوع گیر [۱].

این نوع گیر در سازندهای با تراوایی بالا رخ می‌دهد و در سازندهای شیلی و کم‌تراوا به وجود نمی‌آید. اگر لوله‌ها برای مدت کوتاهی حرکت نداشته باشند، در اثر اختلاف فشار ستون گل و فشار سازند، لوله‌ها، خصوصاً لوله‌های وزنه، به دیواره‌ی چاه چسبیده و در اندود گل روی دیواره فرو رفته و گیر می‌کنند.

گیر اختلاف فشاری از پارامترهایی مثل خصوصیات گل و خصوصیات سازند تأثیر می‌پذیرد. وزن گل سبب ایجاد اختلاف فشار می‌گردد. به طوری که اگر در یک طول نسبتاً زیاد از لوله امتداد یابد، سبب ایجاد چندین هزار پوند نیرو می‌شود که برای جلوگیری از حرکت لوله کافی است. به هر حال برای رخداد گیر اختلاف فشاری باید اندود گل موجود باشد. زیرا اندود گل به طور طبیعی به وجود نمی‌آید. علامت گیر اختلاف فشاری عدم قابلیت چرخش رشته‌ی حفاری یا حرکت آن‌ها به سمت بالا و پایین است. این در حالی است که گردش گل حفاری برقرار است.

مقدار فشار فراتعادلی در ناحیه‌های پرخطر بستگی به خصوصیات سازند، زاویه‌ی چاه، اندازه‌ی چاه، تجهیزات درون‌چاهی و سطح تماس بین لوله و سازند تراوا دارد. خصوصیات سازند را نمی‌توان تغییر داد. همچنین در مواردی که بایستی چاه کنترل شود یا پایدار بماند استفاده از گل با وزن بالا غیر قابل اجتناب است. بنابراین باید خصوصیات گل را طوری تغییر داد که از گیر اختلاف فشاری جلوگیری شود. می‌توان گفت گیر اختلاف فشاری تا زمانی که رشته‌ی حفاری می‌چرخد، رخ نمی‌دهد.

طبق تعریف به تمامی ترکیباتی که به منظور کمک به عملیات حفاری و خروج کنده‌ها و خرده‌های حفاری از چاه مورد استفاده قرار می‌گیرد، سیال حفاری اطلاق می‌گردد. زمانی که بخش عمده‌ی یک سیال حفاری، مایع باشد (آب یا روغن) و سوسپانسیونی از جامدات در این مایع وجود داشته باشد، عبارت گل حفاری به جای سیال حفاری استفاده می‌شود. در واقع گل حفاری یک سیستم کلوئیدی می‌باشد که از دو فاز مایع و جامد تشکیل شده است [۹ و ۱۰]. گل‌های حفاری حاوی یک سیال اصلی و ترکیبی از افزودنی‌های شیمیایی و معدنی برای انجام وظایف گوناگون در طی عملیات حفاری هستند. تکنولوژی سیال حفاری در ارتباط تنگاتنگ با علوم زمین‌شناسی، شیمی و فیزیک بوده و هدف آن اتمام عملیات حفاری و تکمیل چاه با حداقل مخارج و خسارت به چاه می‌باشد [۱۰].

۱-۱-۳ مهم‌ترین وظایف سیالات حفاری

سیال حفاری، یکی از بخش‌های مهم و پرهزینه‌ی حفاری است که وظایف مهم و حیاتی را در فرآیند حفر چاه به عهده دارد. برخی از مهم‌ترین وظایف سیال حفاری در ادامه ذکر می‌شود.

- تمیز کردن ته چاه و انتقال کنده‌های حفاری به سطح زمین [۹ و ۱۰]
- کنترل فشار زیرسطحی [۹ و ۱۰]
- روان‌کاری و خنک نمودن مته و لوله‌های حفاری [۹ و ۱۰]
- ایجاد اندود گل مناسب جهت جلوگیری از هرزروی گل و ناپایداری چاه [۹ و ۱۰]
- معلق نگه داشتن کنده‌های حفاری و مواد وزن افزا در زمان قطع جریان گردش سیال [۱۰]
- جداکردن مواد حفاری شده پس از رسیدن به سطح زمین [۱۱]
- انتقال توان هیدرولیکی پمپ‌ها به مته [۹ و ۱۰]
- کاهش آسیب به سازند [۹ و ۱۰]
- ارائه اطلاعات زمین‌شناسی مناسب از سازندها [۱۰]
- تحمل بخشی از وزن لوله‌های حفاری و لوله‌های جداری با استفاده از اثر نیروی شناوری [۱۰]
- جلوگیری از خوردگی لوله‌های حفاری [۱۰]
- به حداقل رساندن صدمات وارد بر محیط زیست [۹ و ۱۰]

۱-۱-۴ پارامترهای مؤثر گل حفاری بر گیر اختلاف فشاری

گل حفاری به‌عنوان یکی از مهم‌ترین بخش‌های حفاری، می‌تواند بر وقوع یا عدم وقوع پدیده‌ی گیر اختلاف فشاری مؤثر باشد. برخی از پارامترهای مؤثر سیال حفاری، در ادامه ذکر می‌گردد.

- وزن گل فراتعادلی (فشار بیش‌تر از فشار سازند)
- محتوای جامد گل حفاری
- نوع گل
- هرزروی گل
- فرمولاسیون گل که شامل وجود افزودنی‌هایی مثل چرب‌کننده‌ها و ذرات متشکله فیلتره گل می‌شود.
- کیفیت اندودگل شامل ضخامت، چرب‌بودن و مقاومت [۴]

۱-۱-۵ فاکتورهای افزایش دهنده‌ی نیروی گیر اختلاف فشاری

فاکتورهای زیر می‌توانند نیروی اختلاف فشاری و در نتیجه احتمال و شدت وقوع گیر را افزایش دهند:

- اندودهای گل ضخیم و تراکم‌پذیر
- افزایش قطر لوله برای یک چاه با اندازه‌ی مشخص
- افزایش سطح تماس لوله با دیواره‌ی چاه [۹]

۱-۱-۶ روش‌های تعیین نقطه آزادسازی

دو روش برای تعیین نقطه آزادسازی لوله حفاری وجود دارد: روش کشش و استفاده از ابزار تعیین.

۱-۶-۱-۱ روش کشش

با استفاده از کشش رشته‌ی حفاری به بالا و استفاده از جداول موجود، می‌توان نقطه‌ی گیر را تعیین کرد. این روش فقط یک مقدار تقریبی برای نقطه‌ی گیر را به ما می‌دهد که باید فقط برای تعیین فواصل پمپاژ پیل و یا استفاده از ابزار تعیین نقطه آزاد مورد استفاده قرار گیرد.

$$L = \frac{735294 \times W \times \Delta L}{\Delta F} \quad (1-1)$$

که در آن، L عمق نقطه آزاد بر حسب فوت (ft)، ΔL افزایش طول لوله بر اثر کشش بر حسب اینچ (in)، ΔF میزان حداکثر کشش وارد بر لوله بر حسب پوند (lbs) و W وزن لوله حفاری بر حسب پوند بر فوت (lb/ft) می‌باشد.

۱-۶-۱-۲ ابزار تعیین نقطه آزاد (FPIT)

این ابزار با استفاده از سیستم کابلی، جهت بازیابی رشته‌ی حفاری به عمیق‌ترین نقطه‌ی چاه فرستاده می‌شود. با اعمال کشش و گشتاور به لوله، ابزار تعیین نقطه آزاد می‌تواند پیچش و کشیدگی را در هر عمقی تعیین نماید. سپس با استفاده از نمودار عمق بر حسب درصد نیروی کشش و گشتاور اعمال شده به پایین رشته حفاری، عمق نقطه آزاد رشته‌ی حفاری تعیین می‌گردد [۱۲ و ۱۳].

۲-۱ تاریخچه تحقیق

تحقیقات در مورد گیر لوله‌ها از سال ۱۹۵۰ شروع شد. در سال ۱۹۸۵، کینگزبورو^۱ و همپکینز^۲ با انجام تحلیل استاتیکی گیر لوله‌ها بر مبنای پارامترهای حفاری، به این مهم پرداختند [۱۱]. این کار با مقایسه‌ی چاه‌هایی که در آنها گیر لوله‌ها اتفاق افتاده بود و چاه‌های بدون گیر لوله انجام شد. کینگزبورو و همپکینز، ۲۲۱ پارامتر را در ۱۳۱ مورد گیر در چاه‌های مکزیک مورد مطالعه قرار دادند و احتمال گیر لوله‌های حفاری را در چاه‌های اطراف پیش‌بینی کردند. در سال ۱۹۹۴، بایگلر^۳ و کاهن^۴ با ایجاد یک بانک اطلاعاتی برای ۲۲ پارامتر حفاری در ۷۳ چاه بدون گیر خلیج مکزیک و ۵۴ چاه دارای پدیده‌ی گیر لوله‌ها به تحلیل این موضوع پرداختند [۱۴]. این دو، سازنده تحلیل‌های اولیه‌ی مقایسه‌ای بودند. آن‌ها توسط تلفیق متغیرهای حفاری به این تحلیل پرداختند. آن‌ها نه تنها احتمال گیر لوله‌ها را پیش‌بینی کردند، بلکه توانستند مکانیسم آن را نیز شناسایی کنند. در سال ۱۹۹۴، هاوارد^۵ و گلاور^۶ با استفاده از تکنیک‌های آماری توانستند مدل‌های پیش‌بینی گیر لوله‌ها را ترقی بخشند [۱۵]. آن‌ها این کار را با آزمایش در ۱۰۰ چاه، در خلیج مکزیک انجام دادند. این مدل‌ها در جلوگیری از گیر لوله‌ها یا عملیات آزادسازی استفاده می‌شد. اخیراً نیز کاربردی از شبکه‌های عصبی برای پیش‌بینی گیر لوله‌ها به صورت اختلاف فشاری در خلیج مکزیک توسط شرکت هالیبرتون به چاپ رسیده است [۱۶].

مسائل گیر لوله به طور کلی مربوط به هزینه‌های اضافی چاه و زمان اضافی به عنوان یک زمان بدون بهره (به علت توقف عملیات حفاری) و تلاش جهت آزادسازی لوله‌های گیر افتاده می‌باشد. پس از گیر افتادن

1 Kingsborough

2 Hempkins

3 Biegler

4 Kuhn

5 Howard

6 Glover

رشته‌ی حفاری، می‌بایست به سرعت هرگونه تلاشی برای آزادسازی آن صورت گیرد. زیرا احتمال آزادسازی موفق، با گذشت زمان به سرعت کاهش می‌یابد [۱۷و۱]. این تلاش‌ها ممکن است بی‌نتیجه بماند و عدم رهاسازی لوله‌های حفاری عوارض عمده‌ای همانند از دست رفتن رشته‌ی حفاری یا از دست رفتن کامل چاه را در پی خواهد داشت [۱]. درصد زیادی از گیرلوله‌ها در نهایت به حفاری چاه انحرافی در نزدیک لوله گیرکرده که مانده نامیده می‌شوند، منتهی می‌گردد و حفاری مجدد را به همراه دارند [۱]. هر یک از حالت‌های ذکر شده اتلاف زمان و هزینه‌های مضاعف را در پی خواهد داشت. بنابراین، کلید واقعی موفقیت، اجتناب از عوامل منجر به وقوع گیر لوله می‌باشد. اگر این عوامل به طور دقیق مشخص شوند، می‌توان به گونه‌ای عمل نمود که احتمال گیر لوله کاهش یابد [۵ و ۱۸].

تشخیص سریع مؤثرترین عامل احتمالی منجر به گیر، بسیار تعیین‌کننده است. چون برای هر علت گیر، باید اقدامات متفاوت صورت گیرد. یک واکنش نامناسب در مقابل مسأله گیر می‌تواند به سادگی باعث بدتر شدن شرایط گردد. بررسی و ارزیابی وقوع گیر، روشی مناسب برای پیش‌بینی این پدیده و راهنمایی برای پیشگیری‌های صحیح و مفید می‌باشد [۱]. مسأله گیر اختلاف فشاری تا قبل از شناختن مکانیسم آن، با به‌کارگیری برخی روش‌های استاندارد و تکنیک‌ها بعد از حادث شدن، برطرف می‌شوند.

در گذشته از آزمایش گیر لوله شبیه‌سازی شده با سیالات حفاری گوناگون و تکنیک‌های آنالیز آماری چند متغیره [۱۴] برای بررسی شرایط مختلف گیر اختلاف فشاری، استفاده شده است. مکانیسم گیر اختلاف فشاری لوله‌های حفاری توسط هلمیک^۱ و لانگلی^۲ در سال ۱۹۵۷ و اوتمنز^۳ در سال ۱۹۵۸ [۱۹] بر اساس تست‌های آزمایشگاهی بیان شد. آن‌ها تشریح کردند که گیر اختلاف فشاری زمانی رخ می‌دهد که لوله در مقابل یک لایه با نفوذپذیری بالا متوقف می‌گردد، اندود گل حفاری قسمتی از سطح لوله را می‌پوشاند و

1 Helmick

2 Longley

3 Outmans

لوله توسط اختلاف فشار بین دیواره‌ی چاه و سازند نگه داشته می‌شود. آزمایش‌های آن‌ها همچنین نشان داد نیروی لازم برای حرکت یک لوله فولادی از میان اندود گل، با افزایش اختلاف فشار و زمان، افزایش می‌یابد [۴ و ۲۰]. در مورد آزمایش‌های انجام شده برای شبیه‌سازی، هزینه و زمان زیادی صرف می‌شد. همچنین تطابق شرایط آزمایشگاهی با شرایط واقعی بسیار مشکل بود. از طرف دیگر در بیش‌تر روابط آماری نیاز به جداسازی انواع مختلف گیر، حذف بسیاری از داده‌ها به علت پراکندگی و همچنین جستجوی بسیار به دنبال روش مناسب، وجود داشت [۴].

گیر اختلاف فشاری متأثر از بسیاری از پارامترها همانند خصوصیات سیال حفاری و خصوصیات سازند می‌باشد. برای به حساب آوردن همه‌ی جنبه‌های گیر و پیش‌بینی رفتارهای غیرخطی متغیرهای وابسته، نیاز به یک روش جدید به‌عنوان ابزار پیش‌بینی کننده، احساس می‌شود [۴، ۵ و ۲۱]. در سال‌های اخیر سیستم‌های هوشمند به‌عنوان یک روش جایگزین مؤثر، در حل بسیاری از مسائل بالادستی مهندسی نفت استفاده شده‌اند. در میان تکنیک‌های محاسبات هوشمند، به روش‌هایی همانند شبکه‌های عصبی مصنوعی^۱ (ANN) [۲۲، ۲۳، ۲۴ و ۲۵]، الگوریتم‌های ژنتیک^۲ (GA) [۲۶]، منطق فازی^۳ (FL) [۲۱ و ۲۲] و ماشین‌های بردار پشتیبان^۴ (SVM) [۲۶ و ۲۷] بیش‌تر پرداخته شده است. به علت این که آن‌ها توانایی نتیجه‌گیری معنی‌دار از داده‌های پیچیده و غیردقیق را دارند و می‌توانند برای بیرون کشیدن طرح‌ها و راه حل‌هایی استفاده شوند که دیگر تکنیک‌های انسانی یا کامپیوتری در آن ناتوانند. در میان این روش‌های هوشمند، شبکه عصبی مصنوعی رایج‌ترین روش می‌باشد و نتایج در خور توجه و مناسب داشته است [۲۷]. شبکه‌های عصبی نسبت به روش‌های آماری دارای مزایای متعددی همانند محاسبات سریع، انعطاف‌پذیری

1 Artificial Neural Networks

2 Genetic Algorithms

3 Fuzzy Logic

4 Support Vector Machines

نسبت به خطا و همچنین کار با داده‌ها بدون در نظر گرفتن مدل خاص، می‌باشند [۴]. علی‌رغم این مزایای متعدد به هر حال معایب ذاتی همانند مستعد بودن برای رفتار بی‌نظم، محاسبات زیاد در مرحله‌ی آموزش، کمینه محلی، برازش بیش از حد، مسأله تعیین ساختار و غیره در شبکه عصبی وجود دارد.

بعضی از روش‌ها زمانی مؤثر هستند که گیر لوله اتفاق افتاده باشد. برخی از کارهای مهم که بعد از گیر لوله، صرف‌نظر از نوع گیر انجام می‌شود، استفاده از ضربه‌کوب حفاری می‌باشد. ضربه‌کوب بر حسب نوع عوامل گیر لوله حفاری، می‌تواند بر دو نوع ضربه به بالا یا به پایین انتخاب شود.

گونزالس و همکارانش در سال ۲۰۰۷ [۲۸] اظهار کردند که عملیات ضربه زدن، ممکن است روزها یا هفته‌ها به طول انجامد که باعث هدر رفتن زمان می‌شود، لذا لرزش رزونانسی می‌تواند به جای ضربه‌کوب حفاری استفاده شود. در پژوهش آن‌ها بیان گردید که این روش زمانی که بلافاصله بعد از رخ دادن گیر لوله به کار برده شود، مؤثرتر خواهد بود. این روش و تکنیک سریعاً لوله را آزاد می‌کند و می‌توان به راحتی آن را به کار برد.

در سال ۲۰۰۸ شاراد و ناوین اگروال [۲۹] طراحی جدیدی از لوله‌های وزنه ارائه کردند که سطح تماس رشته‌ی حفاری را با اندود حفاری کاهش می‌دهد، در نتیجه نیروی لازم جهت آزادسازی کاهش می‌یابد. همپکینز و همکارانش در سال ۱۹۸۷ [۱۱] ایده‌ی استفاده از روش‌های آماری چند متغیره (MSA) جهت پیش‌بینی گیر لوله حفاری را ارائه کردند. در این روش از ۱۳۱ مورد گیر لوله و ۲۰ متغیر حفاری استفاده شد که ارتباط بین متغیر مستقل و وابسته‌ی حفاری را طوری مشخص می‌کرد که گیر حفاری را تعیین شود.

در سال ۲۰۱۰ شادی‌زاده و همکارانش [۳۰] تحقیقاتی را انجام دادند که در آن با استفاده از شبکه عصبی مصنوعی، احتمال گیر لوله در یکی از میادین نفتی ایران تعیین گردید. نتیجه‌ی تحقیقات آن‌ها

احتمال ۹۰٪ دقت را در پیش‌بینی گیر لوله در میدان مذکور نشان می‌داد. گیر لوله‌ها به دو نوع دینامیک و استاتیک تقسیم شدند. به‌صورتی که نوع دینامیک مربوط به زمانی است که گل حفاری در جریان است و حالت استاتیک مربوط به زمان سکون گل حفاری بود.

۳-۱ رویکرد این پایان‌نامه در حل مسأله

هدف اصلی این پژوهش ساخت مدلی است که در آن وقوع گیر رشته حفاری پیش‌بینی شود. در این تحقیق ابتدا با استفاده از داده‌های واقعی به‌دست آمده از صنعت حفاری و به‌کارگیری ابزارهای دسته‌بندی مانند الگوریتم ماشین بردار پشتیبان و الگوریتم K نزدیک‌ترین همسایه، مجموعه داده‌ها بر حسب الگوی گیر و عدم گیر دسته‌بندی می‌شوند. دسته‌بندی مطلوب داده‌ها به دسته‌های گیر و عدم گیر، سبب اطمینان از امکان پیش‌بینی به‌وسیله مدل‌های پیش‌بینی کننده، با استفاده از داده‌های موجود، می‌گردد. در ادامه کارایی شبکه عصبی مصنوعی و سری زمانی به‌عنوان روش‌های نوین در پیش‌بینی پدیده گیر رشته حفاری، مورد مقایسه و تحلیل قرار می‌گیرد. سری زمانی به‌عنوان مدلی دینامیکی (بر اساس شرایط و رخداد‌های روزهای گذشته) جهت پیش‌بینی گیر رشته حفاری، به‌کار رفته است. در این پایان‌نامه از شبکه عصبی پرسپترون چند لایه^۱ (MLP) و سری‌های زمانی تحت شبکه عصبی^۲، جهت پیش‌بینی استفاده گردید.

1 Multi Layer Perceptron

2 neural network time series

۱-۴ ساختار پایان نامه

در فصل اول کلیات پایان نامه شامل تعریف مسأله، تاریخچه تحقیق و رویکرد این پایان نامه در حل مسأله بیان گردید. در فصل دوم این پایان نامه، داده ها و پارامترهایی که در ساخت مدل ها به کار رفته اند، معرفی می شوند. این پارامترها شامل پارامترهای حفاری (مانند پارامترهای عملیاتی و پارامترهای مسیر چاه)، پارامترهای سیال حفاری و پارامتر محیطی مرتبط با زمین شناسی منطقه می باشند. در فصل سوم این پایان نامه، روش تحقیق بیان و ابزارها و مدل های به کار رفته در پژوهش معرفی می گردد. در فصل چهارم نتایج به دست آمده از این پژوهش به تفصیل مورد بحث قرار می گیرد و در فصل پنجم نتیجه گیری و پیشنهاداتی که پس از انجام این تحقیق قابل ارائه است، بیان خواهد شد.

فصل دوم: داده‌های تحقیق

۱-۲ داده‌ها و پارامترهای ورودی

۲-۲ پارامترهای عملیات و سیال حفاری

۳-۲ سازندها و لیتولوژی

۱-۲ داده‌ها و پارامترهای ورودی

به منظور به‌دست آوردن داده‌های مورد نیاز این پژوهش، گزارشات روزانه حفاری و نمودارگیری گل مربوط به حفر چاه‌های یکی از میادین جنوب غربی ایران مطالعه گردید. با توجه به اهمیت زیاد داده‌های ورودی، این مطالعات با دقت نظر فراوان و صرف زمان صورت پذیرفت. اکثر چاه‌های این میدان در دسته‌ی چاه‌های عمیق قرار دارند. در جدول ۱-۲ تقسیم‌بندی انواع چاه‌ها بر اساس عمق حفاری تا رسیدن به هدف^۱ مورد نظر، بیان گردیده است [۳۱].

جدول ۱-۲ طبقه‌بندی چاه‌ها بر اساس عمق حفاری شده نهایی

نام دسته چاه	عمق
کم عمق	کمتر از ۲۰۰۰ متر
متعارف ^۲	بین ۲۰۰۰ تا ۳۵۰۰ متر
عمیق	بین ۳۵۰۰ تا ۵۰۰۰ متر
خیلی عمیق	بیش از ۵۰۰۰ متر

طی مطالعه‌ی گزارشات ثبت شده عملیات و سیال حفاری در صنعت، ۲۰ پارامتر به عنوان پارامترهای ورودی به‌دست آمد که این پارامترها در ادامه معرفی می‌شوند. همچنین با مطالعه‌ی گزارشات زمین‌شناس، اطلاعات سازند نیز به عنوان پارامتری مجزا و به‌صورت معادل عددی لیتولوژی تعیین گردید. در مجموع ۹۷۳ مجموعه داده که هر کدام شامل ۲۱ پارامتر است، به‌دست آمد. از این تعداد مجموعه داده، ۱۲۶ مجموعه مربوط به موارد رخداد گیر و ۸۴۷ مجموعه مربوط به عدم رخداد گیر در چاه‌های حفاری شده بودند.

1 target

2 conventional

۲-۲ معرفی پارامترهای عملیات و سیال حفاری

در این قسمت پارامترهای ورودی معرفی می‌گردد. این پارامترها شامل پارامترهای عملیات حفاری، پارامترهای حفاری مرتبط با مسیر چاه و پارامترهای سیال حفاری می‌باشند.

وزن روی مته (WOB) : بار یا وزن روی مته یکی از پارامترهای اساسی عملیات حفاری است که بر روی نرخ حفاری تأثیر بسزایی دارد. این پارامتر معمولاً با واحد تن یا کیلو پوند بیان می‌گردد. اثر وزن روی مته بر سرعت حفاری در سنگ‌های مختلف متفاوت است. گاهی در سازندهای بسیار نرم بدون اعمال وزن روی مته، حفاری فقط با انرژی هیدرولیکی گل صورت می‌گیرد. اعمال بیش از حد وزن روی مته نه تنها سبب افزایش نرخ حفاری نخواهد شد، بلکه ممکن است باعث آسیب به مته شود [۳۲].

سرعت چرخش مته (RPM) : بیانگر تعداد دفعات چرخش مته‌ی حفاری در هر دقیقه است که به طور مستقیم در میزان حفاری مؤثر می‌باشد.

سرعت چرخش موتور درون چاهی (RPM Turbin) : گاهی علاوه بر چرخش کل رشته حفاری، با قرار دادن موتور درون چاهی نوعی چرخش به مته و تجهیزات انتهایی رشته حفاری اعمال می‌گردد. این موتور در جهت‌دهی به حرکت رشته حفاری در مسیر از پیش تعیین شده، مؤثر خواهد بود.

نرخ پمپ شدن سیال^۱: بیانگر حجم گل حفاری است که توسط پمپ گل در هر دقیقه به درون چاه فرستاده میشود. این پارامتر بر حسب گالن بر دقیقه بیان می‌شود.

1 flow rate

نرخ نفوذ حفاری^۱ (ROP): میزان پیشروی متهی حفاری درون سنگ‌های تحت حفاری به عنوان نرخ نفوذ حفاری تعریف و با واحدهای فوت بر دقیقه یا متر بر ساعت شناخته می‌شود [۳۱].

زاویه‌ی انحراف مته: نشان دهنده‌ی میزان انحراف یا کج شدگی متهی حفاری از خط قائم می‌باشد و با واحد درجه بیان می‌گردد.

آزیموت مته: میزان اختلاف جهت حرکت مته و رشته حفاری نسبت به شمال جغرافیایی است که با واحد درجه بیان می‌گردد.

میزان انحراف چاه از نقطه شروع: با شروع حفاری و با توجه به انحراف و آزیموت مته، چاه از حالت کاملاً عمود خارج می‌شود و در هر عمق، چاه نسبت به نقطه شروع مقداری انحراف جانبی بر حسب متر یا فوت خواهد داشت.

چگالی گل: وزن گل یا چگالی گل جهت کنترل فشارهای زیرسطحی و پایداری چاه به کار می‌رود. این پارامتر با واحدهای پوند بر گالن (ppg) یا پوند بر فوت مکعب (pcf) بیان می‌گردد [۳۳].

میزان اسیدی یا بازی بودن (PH) گل: اندازه‌گیری و تنظیم این پارامتر به دلیل فعل و انفعالات رس و حل شوندگی افزودنی‌ها، از اهمیت فراوان در کنترل سیال حفاری برخوردار است.

میزان نمک (کلر) گل: میزان کلر موجود در گل، به خوبی حفاری در لایه‌های نمک و ورود جریان آب شور را بیان می‌کند و کیفیت گل را از این نظر مورد سنجش قرار می‌دهد.

میزان یون Ca^{2+} موجود در گل: کاتیون‌های دو ارزشی^۲ می‌توانند بر روی خصوصیات گل اثر بگذارند. یکی از این کاتیون‌ها کلسیم است که به صورت مجزا از سایر کاتیون‌ها بررسی می‌شود [۳۳].

1 rate of penetration

2 divalent cations

قطر چاه: قطر دهانه‌ی چاه به دلیل اثراتی که می‌تواند بر تجمع یا تخلیه خرده‌های حفاری و یا تماس اندود گل و رشته حفاری داشته باشد، یکی از پارامترهای مهمی است که باید در نظر گرفته شود.

فشار پمپ (PP): این فشار، فشار عملیاتی است و تعیین‌کننده فشار سیال در چاه و مقابل سازند خواهد بود. حالت فراتعادلی زیاد موجب افزایش افت صافی و ضخیم‌تر شدن اندود گل حفاری می‌گردد. بنابراین افزایش آن احتمال گیر اختلاف فشاری را افزایش می‌دهد.

گران‌روی قیف مارش گل حفاری (FV): این پارامتر جهت بیان درجه غلظت گل به کار می‌رود. گران‌روی قیف مارش، میزان زمان لازم بر حسب ثانیه جهت خروج ۰/۹۴۶ لیتر گل حفاری از قیف مارش است [۳۳]. هنگامی که نفوذپذیری اندود گل کم و مقدار گران‌روی صافاب گل پایین باشد، ضخامت اندود گل افزایش می‌یابد و در پی آن گیر لوله حفاری اتفاق می‌افتد.

گران‌روی پلاستیکی (PV): گران‌روی پلاستیکی بر تمیزکاری چاه و فشار پمپ مؤثر است. با افزایش مقدار جامدات در گل حفاری، ضریب اصطکاک میان ذرات و تنش‌های برشی (گران‌روی پلاستیکی) افزایش می‌یابد. بنابراین خطر وقوع گیر اختلاف فشاری بالا می‌رود. روش‌هایی مانند رقیق‌سازی با آب، به‌کارگیری تجهیزات جداکننده جامدات و استفاده از مواد شیمیایی برای کاهش گران‌روی پلاستیکی مؤثر می‌باشند.

نقطه‌ی واروی (YP): نقطه‌ی واروی یک معیار برای مقاومت داخلی سیال در شرایط اولیه جریان می‌باشد. این مقاومت به‌وسیله نیروهای جاذبه الکتریکی ذرات کلوئیدی به‌وجود می‌آید. این پارامتر در جابه‌جایی ذرات جامد و کنده‌های حفاری مؤثر است [۱].

استحکام ژلی اولیه^۱ یا ۱۰ ثانیه (IGS) و استحکام ژلی بعد از ۱۰ دقیقه (10 min GS): استحکام ژلی گل حفاری به‌صورت توانایی گل حفاری در معلق نگه‌داشتن ذرات جامد تعیین می‌گردد. استحکام ژلی تابعی از

1 Initial Gel Strength

مقدار و نوع جامدات معلق، زمان، دما و مواد شیمیایی موجود در گل می‌باشد. استحکام بندش نامناسب منجر به ته‌نشینی کنده‌های حفاری و دیگر جامدات گل می‌گردد که این موارد در نهایت وقوع گیر را در پی خواهد داشت.

عمق چاه (MD): منظور از عمق چاه، عمق اندازه‌گیری شده در حین عملیات حفاری می‌باشد. گرچه عمق چاه را بر خلاف خصوصیات گل حفاری، نمی‌توان کنترل کرد اما اگر دمای گل شاخصی از دمای زمین در عمق حفاری باشد و به علت این که با افزایش عمق چاه، طول لوله حفاری افزایش می‌یابد، کنترل شرایط مشکل‌تر می‌گردد [۱]. در نتیجه امکان وقوع گیر نیز افزایش می‌یابد. لذا این پارامتر نیز به‌عنوان ورودی در نظر گرفته شده است.

۲-۳ سازندها و لیتولوژی

سازندهای موجود در میدان مورد مطالعه در این پژوهش متعلق به ناحیه‌ی زاگرس می‌باشند.

۲-۳-۱ ناحیه‌ی زاگرس

این ناحیه، ناحیه‌ی جنوب و جنوب غرب ایران در بر می‌گیرد. حد غربی آن راندگی اصلی زاگرس و حد شرقی آن گسل میناب است. ناحیه زاگرس خود به سه ناحیه کوچک‌تر به نام‌های دشت خوزستان، زاگرس چین‌خورده و زاگرس رورانده تقسیم می‌شود.

۲-۳-۱-۱ دشت خوزستان

این قسمت غربی‌ترین ناحیه‌ی زاگرس را شامل می‌شود و به‌صورت دشت تقریباً مسطحی است که قسمت زیادی از بیرون‌زدگی آن را رسوبات آبرفتی عهد حاضر و یا نئوژن تشکیل می‌دهد. ضخامت این پوشش رسوبی از اهواز به طرف آبادان و سواحل خلیج فارس به‌تدریج زیاد می‌شود.

۲-۳-۱-۲ زاگرس چین‌خورده

این واحد ساختمانی پهنایی حدود ۱۵۰ تا ۲۵۰ کیلومتر دارد که در حاشیه‌ی غربی ایران و به موازات راندگی زاگرس قرار دارد و به علت چین خوردگی رسوبات آن، به آن زاگرس چین‌خورده می‌گویند. قسمت اعظم ذخایر نفتی کشور ما در این بخش از زاگرس قرار دارد. ضخامت رسوبات در منطقه زاگرس چین‌خورده بیش از ۷۰۰۰ متر است. این رسوبات اساساً از جنس دولومیتی، تبخیری، آهکی، ماسه‌سنگی و تخریبی است و نشان دهنده‌ی دریای کم‌عمق در طی تکوین این بخش از زاگرس است.

۲-۳-۱-۳ زاگرس رورانده

محققین زمین‌شناس این بخش را زاگرس داخلی، زاگرس مرتفع و ناحیه خردشده هم گفته‌اند. این ناحیه به‌صورت نوار باریک و طولی در قسمت شرق زاگرس چین‌خورده و به موازات آن قرار دارد [۳۴]. در ادامه سازندهای مهم ناحیه‌ی زاگرس معرفی می‌گردند.

۲-۳-۲ سازندهای مهم ناحیهی زاگرس

تمامی سازندهایی که حفاری چاههای این میدان در آن صورت پذیرفته است، متعلق به ناحیهی زاگرس می‌باشند. لیتولوژی و سنگ‌های هر سازند و عمق میانگین که هر سازند در چاههای مختلف میدان شروع می‌شود، در جدول ۲-۲ ارائه گردیده است.

جدول ۲-۲ لیتولوژی هر سازند و عمق شروع سازند^۱ در میدان

لیتولوژی غالب	نام سازند یا گروه	عمق شروع میانگین بر حسب متر
انیدریت، مارل، سنگ آهک، دولومیت	گروه فارس	۹۹/۱۳
سنگ آهک و دولومیت	آسماری	۵۷۰
دولومیت و سنگ آهک	چهرم	۶۳۵
سنگ آهک	ایلام	۱۰۶۴
شیل، گل‌سنگ، سنگ آهک	لافان	۱۱۳۵
سنگ آهک	سروک	۱۱۶۱
شیل، گل‌سنگ، ماسه‌سنگ	کزدمی	۱۲۱۳
سنگ آهک	داریان بالایی	۱۲۶۲
مارل، شیل	هوار	۱۳۱۹
سنگ آهک	داریان پایینی	۱۳۵۴
سنگ آهک رسی	گدون بالایی	۱۳۸۸
سنگ آهک	عضو خلیج	۱۴۱۸
سنگ آهک رسی	گدون پایینی	۱۴۳۸

¹ formation top

ادامه جدول ۱-۲ نمایش لیتولوژی هر سازند و عمق شروع سازند در میدان

۱۴۸۰	فهلپان	سنگ آهک، دولومیت، چرت
۱۷۰۶	هیث	انیدریت و دولومیت
۱۸۲۰	سورمه بالایی	دولومیت و انیدریت
۱۹۲۰	دولومیت بالایی	دولومیت
۱۹۸۶	سنگ آهک بالایی	سنگ آهک
۲۱۸۸	ناحیه چرتی	دولومیت و سنگ آهک
۲۳۰۰	سنگ آهک میانی	سنگ آهک
۲۵۷۶	عضو مند	سنگ آهک و دولومیت
۲۷۴۴	سنگ آهک پایینی	سنگ آهک
۳۰۲۰	شیل سورمه پایینی	سنگ آهک رسی و گل سنگ
۳۱۱۲	عضو لیتوتیس	سنگ آهک، دولومیت و گل سنگ
۳۱۷۸	نیریز	دولومیت، سنگ آهک، گل سنگ
۳۲۱۰	دشتک بالایی	دولومیت و سنگ آهک
۳۳۹۲	تبخیری	انیدریت
۳۵۶۶	سودیر بالایی	گل سنگ، سنگ آهک رسی
۳۶۶۸	انیدریت حجیم	انیدریت و گل سنگ
۳۷۲۸	S8	شیل و گل سنگ
۳۷۹۵	S7	گل سنگ
۳۸۴۷	S6	دولومیت
۳۸۸۰	S5	دولومیت و سنگ آهک
۳۹۳۴	S4	دولومیت و سنگ آهک

ادامه جدول ۱-۲ نمایش لیتولوژی هر سازند و عمق شروع سازند در میدان

۴۰۱۱	S3	دولومیت
۴۰۲۶	S2	دولومیت
۴۱۴۹	شیل آغار	گل سنگ و شیل
۴۲۰۶	کنگان بالایی	دولومیت، گل سنگ، انیدریت
۴۲۳۱	K1	دولومیت، سنگ آهک، انیدریت
۴۴۵۸	K2	دولومیت و سنگ آهک
۴۵۴۸	K3	دولومیت، سنگ آهک، انیدریت
۴۸۶۵	K4	دولومیت و سنگ آهک

فصل سوم: روش شناسی

تحقیق

۳-۱ مقدمه

۳-۲ الگوریتم K نزدیک ترین همسایه

۳-۳ ماشین بردار پشتیبان

۳-۴ معیار ارزیابی روش های دسته بندی

۳-۵ تحلیل مؤلفه های اصلی (PCA)

۳-۶ شبکه عصبی مصنوعی

۳-۷ سری های زمانی

۳-۸ پیاده سازی مدل های شبکه عصبی و سری زمانی تحت شبکه عصبی، در نرم-

افزار

۳-۱ مقدمه

در این فصل روش‌ها و ابزارهای به‌کار رفته در این پژوهش معرفی می‌شوند. این روش‌ها شامل الگوریتم‌های دسته‌بندی K نزدیک‌ترین همسایه و ماشین بردار پشتیبان، همچنین روش‌های آماری مانند تحلیل مؤلفه‌های اصلی و روش‌های پیش‌بینی مانند شبکه عصبی و سری‌های زمانی تحت شبکه عصبی می‌باشند. در این پایان‌نامه جهت به‌کارگیری روش‌ها و الگوریتم‌های ذکر شده، از نرم‌افزار متلب^۱ استفاده گردید. پس از بیان مفاهیم روش‌های پیش‌بینی، پیاده‌سازی این مدل‌ها در نرم‌افزار تشریح می‌گردد.

۳-۲ الگوریتم K نزدیک‌ترین همسایه^۲

این الگوریتم یک الگوریتم آموزش با سرپرستی است. در حالت کلی از این الگوریتم به منظور تخمین تابع چگالی داده‌های آموزش و دسته‌بندی داده‌های آزمون بر اساس الگوهای یادگیری استفاده می‌شود.

برای تخمین $p(x)$ از روی n نمونه‌ی آموزش توسط الگوریتم K نزدیک‌ترین همسایه، می‌توان یک سلول به مرکزیت x ایجاد کرد و اجازه داد شعاع این سلول تا حدی گسترش پیدا کند که Kn نمونه‌ی آموزش را دربر گیرد. این نمونه‌ها Kn نزدیک‌ترین همسایه‌های x هستند. در حالت کلی K به‌صورت Kn در نظر گرفته می‌شود، که Kn تابعی تعریف شده از n است.

1 MATLAB

2 K Nearest Neighbor (KNN)

اگر چگالی نقاط تعلیم اطراف x زیاد باشد، سلول کوچک می‌شود و بنابراین نتیجه‌ی به‌دست آمده نتیجه‌ی بهتری است. در صورتی که چگالی نقاط آموزش اطراف x کم باشد، سلول بزرگ می‌شود. در حالت کلی چگالی توزیع به ازای هر نقطه x توسط رابطه‌ی زیر محاسبه می‌شود:

$$P_n(x) = \frac{K_n/n}{V_n} \quad (1-3)$$

اگر با افزایش n ، K_n نیز افزایش پیدا کند به‌طوری که با رفتن n به سمت بی‌نهایت K_n نیز به بی‌نهایت میل کند، آن‌گاه می‌توان مطمئن بود که K_n/n تخمین خوبی از این احتمال است که یک نقطه در یک سلول به حجم V_n قرار گیرد. بنابراین دو شرط زیر شروط لازم و کافی برای این است که $P_n(x)$ در نهایت به $P(x)$ همگرا شود:

$$\lim_{n \rightarrow \infty} K_n = \infty \quad (2-3)$$

$$\lim_{n \rightarrow \infty} K_n/n = 0 \quad (3-3)$$

به عنوان مثال اگر K_n ، $\sqrt{n}$ در نظر گرفته شود، آن‌گاه:

$$K_n = \sqrt{n} \quad (4-3)$$

$$V_n \approx \frac{1}{\sqrt{n}P(x)} \quad (5-3)$$

$$V_n \approx \frac{V_1}{\sqrt{n}} \rightarrow 0, n \rightarrow \infty \quad (6-3)$$

بنابراین شروط فوق برای k_n داده شده، برقرار بوده و با افزایش n به سمت بی‌نهایت، $P_n(x)$ به $P(x)$ میل می‌کند.

۳-۲-۱ قانون نزدیک‌ترین همسایه

در تعریف قانون نزدیک‌ترین همسایه، k مساوی یک در نظر گرفته شده است. فرض می‌کنیم $D_n = \{x_1, \dots, x_n\}$ مجموعه‌ای از n الگوی ورودی باشد و $x' \in D_n$ نزدیک‌ترین الگوی ورودی به نقطه‌ی x باشد. قانون نزدیک‌ترین همسایه برای طبقه‌بندی x ، آن را در دسته‌ای مشابه با دسته x' قرار می‌دهد. قانون نزدیک‌ترین همسایه یک روال کمتر از حد مطلوب^۱ است. بنابراین نرخ خطای آن معمولاً بیش‌تر از حداقل نرخ خطای ممکن، یعنی نرخ خطای الگوریتم بیز است. در صورت استفاده از تعداد نامحدودی از الگوهای ورودی، نرخ خطا در بدترین حالت بیش‌تر از دو برابر نرخ خطای الگوریتم بیز نخواهد شد.

قانون K نزدیک‌ترین همسایه، گسترشی از قانون نزدیک‌ترین همسایه است. همان‌طور که واضح است این قانون x را در دسته‌ای قرار می‌دهد که بیش‌ترین تکرار را در بین K نزدیک‌ترین همسایه‌ی x دارد. در این روش از فاصله‌ی اقلیدسی جهت معیار فواصل بین همسایگی‌ها استفاده می‌شود. فاصله اقلیدسی فاصله‌ی معمولی دو نقطه است که توسط قضیه‌ی فیثاغورس به‌دست می‌آید.

اگر $p = (p_1, p_2, \dots, p_n)$ و $q = (q_1, q_2, \dots, q_n)$ دو نقطه در فضای n بعدی و pq پاره‌خطی باشد که آن‌ها را به هم وصل کند و معادل فاصله‌ی آن دو نقطه باشد، آنگاه این فاصله به صورت زیر تعریف می‌شود:

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (۷-۳)$$

که می‌توان آن را به صورت زیر بیان کرد [۳۵ و ۳۶]:

$$D(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (۸-۳)$$

¹ suboptimal

۳-۳ ماشین بردار پشتیبان^۱

برای دسته‌بندی و تشخیص الگوی نمونه‌های ورودی، به الگوریتمی نیاز است که بتواند یک مجموعه بزرگ از داده‌های آموزشی با ابعاد زیاد را حمایت کند. ماشین بردار پشتیبان (SVM) یکی از الگوریتم‌های مناسب برای تحقق این هدف است. ماشین بردار پشتیبان یک روش برای دسته‌بندی و تشخیص الگوی داده‌های ورودی است که در سال ۱۹۹۲ توسط باسر^۲، گویان^۳ و وپنیک^۴ معرفی گردید. ایشان ماشین بردار پشتیبان را براساس نظریه یادگیری آماری که در دهه‌ی ۶۰ توسط وپنیک و چروننکیس^۵ بیان شده بود، ارائه دادند. ماشین بردار پشتیبان از نظر تجربی دارای عملکرد خوبی بوده و کاربرد موفقی در بسیاری از زمینه‌ها داشته است [۳۷]. عملکرد ماشین بردار پشتیبان بدین صورت است که هنگام آموزش، مرز تصمیم‌گیری به گونه‌ای انتخاب می‌گردد که فاصله آن با هر یک از دسته‌های مورد نظر حداکثر باشد. در این صورت تصمیم‌گیری، شرایط نویزی را به خوبی تحمل کرده و پاسخ‌دهی مطلوبی خواهد داشت. این نحوه انتخاب مرز بر اساس نقاطی به نام بردارهای پشتیبان انجام می‌پذیرد [۳۸]. روش‌های دسته‌بندی خطی، با ایجاد یک ابرصفحه^۶ که عبارت از یک معادله خطی است، داده‌ها را از هم شناسایی و تفکیک می‌کنند. ماشین بردار پشتیبان نیز به عنوان یک دسته‌بندی کننده خطی، بهترین ابرصفحه‌ای را پیدا می‌کند که با حداکثر حاشیه^۷، داده‌های دو دسته را از هم تفکیک کند. در شکل ۳-۱ حالت ساده‌ای از این عمل نشان داده شده است [۳۷]. در این

^۱ Support Vector Machine (SVM)

^۲ Boser

^۳ Guyon

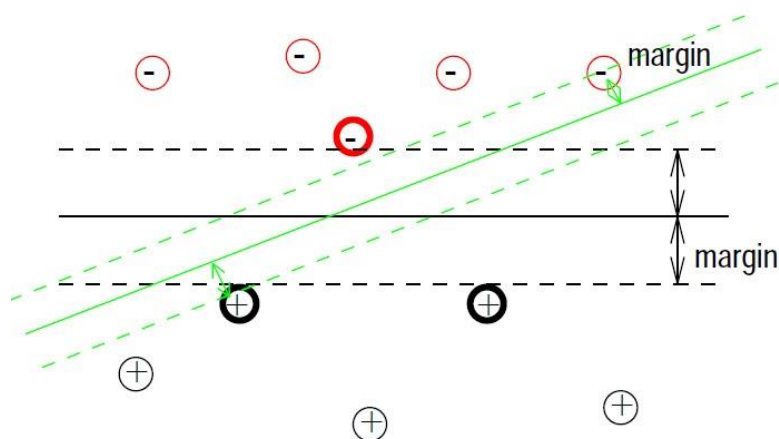
^۴ Vapnik

^۵ Chervonenkis

^۶ hyper plane

^۷ margin

شکل نمونه‌های قرار گرفته بر روی مرزها که با دوایر ضخیم‌تر نمایش داده شده‌اند، بردارهای پشتیبان نام دارند.

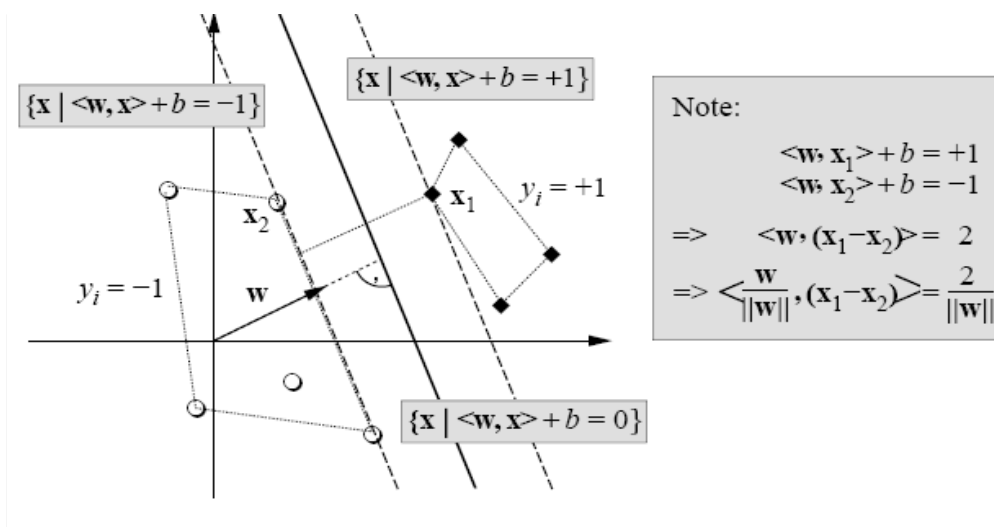


شکل ۱-۳ ابرصفحه با حداکثر مرز جداکننده (خطوط) به همراه مرزهای جداکننده (خط‌چین‌ها)

۱-۳-۳ روش عملکرد ماشین بردار پشتیبان در دسته‌بندی

فرض می‌شود مجموعه نقاط داده $\{(x_1, c_1), (x_2, c_2), \dots, (x_n, c_n)\}$ در اختیار و تفکیک آن‌ها به دو دسته $c_i = \{-1, 1\}$ مطلوب است. هر یک بردار p بعدی از اعداد حقیقی است که در واقع همان متغیرهای بیان‌گر رفتار هستند.

روش دسته‌بندی ماشین بردار پشتیبان بهترین ابرصفحه‌ای را پیدا می‌کند که با حداکثر فاصله^۱، داده‌های مربوط به دو طبقه را از هم تفکیک کند. نحوه‌ی تشکیل ابرصفحه جداکننده توسط ماشین بردار پشتیبان در شکل ۲-۳ نشان داده شده است.



شکل ۲-۳ نحوه ساخت ابرسطح جداکننده بین دو طبقه داده در فضای دو بعدی [۳۷]

در شکل ۲-۳ در اطراف نقاط مربوط به کلاس ۱- و نقاط مربوط به کلاس ۱+ پوسته محدب^۲ رسم شده است. ایده‌ی اصلی این است که یک جداکننده مناسب انتخاب شود. منظور جداکننده‌ای است که بیش‌ترین فاصله را با نقاط همسایه از هر دو طبقه دارد. این جواب در واقع بیش‌ترین مرز را با نقاط مربوط به دو دسته‌ی مختلف دارد و می‌تواند با دو ابرصفحه موازی که حداقل از یکی از نقاط دو دسته عبور می‌کنند، کران‌دار شود. این بردارها، بردارهای پشتیبان نام دارند. فاصله‌ی بیشتر بین دو ابرصفحه موازی، خطای دسته‌بندی را کاهش خواهد داد [۳۹].

^۱ maximum margin

^۲ convex

فرمول ریاضی این دو ابرصفحه موازی که مرز جداکننده را تشکیل می‌دهند، در عبارات (۳-۹) و (۳-۳۰)

(۱۰) نشان داده شده است:

$$w \cdot x - b = 1 \quad (۳-۹)$$

$$w \cdot x - b = -1 \quad (۳-۱۰)$$

که b عرض از مبدأ برای ابرصفحه با حداکثر مرز جداکننده است. نکته‌ی قابل توجه این است که اگر داده‌های تعلیمی به صورت خطی تفکیک‌پذیر باشند، می‌توان دو ابرصفحه را به گونه‌ای انتخاب کرد که هیچ داده‌ای بین آن‌ها نباشد و سپس فاصله‌ی بین این دو ابرصفحه موازی را به حداکثر رساند. با به‌کارگیری قضایای هندسی، فاصله این دو ابرصفحه عبارت است از $2/|w|$ ، پس باید $|w|$ را به حداقل رساند. همچنین باید از قرار گرفتن نقاط داده در ناحیه‌ی درون مرز جلوگیری کرد. برای این کار یک محدودیت ریاضی به تعریف اضافه می‌گردد. برای هر i ، با اعمال محدودیت‌های زیر، اطمینان حاصل می‌شود که هیچ نقطه‌ای در مرز قرار نمی‌گیرد:

$$w \cdot x_i - b \geq 1 \quad \text{برای داده‌های مربوط به طبقه اول} \quad (۳-۱۱)$$

$$w \cdot x_i - b \leq -1 \quad \text{برای داده‌های مربوط به طبقه دوم} \quad (۳-۱۲)$$

می‌توان این محدودیت‌ها را به صورت زیر بیان کرد:

$$c_i (w \cdot x_i - b) \geq 1, \quad 1 \leq i \leq n \quad (۳-۱۳)$$

از لحاظ ریاضی، حل یک مسأله غیرمحدب بسیار دشوارتر از حل یک مسأله محدب است. می توان با جایگزین نمودن $\frac{1}{2} \|w\|^2$ به جای $|w|$ بدون تغییر در پاسخ، مسأله را حل کرد. این مسأله از نوع مسائل درجه دوم^۱ است.

در روش ماشین بردار پشتیبان، بردارهای ورودی به یک فضای چندبعدی نگاشت می شوند. پس از آن، یک ابرصفحه ساخته خواهد شد که با حداکثر فاصله ممکن، بردارهای ورودی را از هم جدا خواهد کرد. به این ابرصفحه، ابرصفحه با حداکثر مرز جداکننده گفته می شود. به طوری که در شکل ۳-۲ نشان داده شده است، دو ابرصفحه موازی در دو سمت ابرصفحه با حداکثر مرز جداکننده ساخته خواهد شد که داده های مربوط به دو طبقه را به گونه ای از هم مجزا می کنند که هیچ داده ای در مرز بین این دو ابرصفحه قرار نمی گیرد. ابرصفحه با حداکثر مرز جداکننده، ابرصفحه ای است که فاصله بین دو ابرصفحه موازی را به حداکثر می رساند. فرض بر این است که اگر مرز جداکننده یا در واقع، فاصله ی بین دو ابرسطح موازی بیش تر باشد، خطای طبقه بندی کم تر خواهد بود.

۳-۴ معیار ارزیابی روش های دسته بندی

میزان خطای رگرسیون و میزان خطا بر اساس اعتبارسنجی متقاطع^۲ را می توان به عنوان معیارهای رایج در ارزیابی مدل هایی مانند الگوریتم K نزدیک ترین همسایه و ماشین بردار پشتیبان نام برد. داده ها به سه دسته آموزش، اعتبارسنجی و آزمون تقسیم می شوند. اعتبارسنجی یک روش ارزیابی است که مشخص می کند نتایج یک تحلیل آماری بر روی یک مجموعه داده تا چه اندازه قابل تعمیم و مستقل از داده های آموزشی

^۱ quadratic

^۲ cross validation

است. این تکنیک به‌طور ویژه در کاربردهای پیش‌بینی مورد استفاده قرار می‌گیرد تا مشخص شود مدل مورد نظر تا چه اندازه در عمل مفید خواهد بود [۴۰].

اعتبارسنجی K-Fold: در این نوع اعتبارسنجی، داده‌ها به K زیرمجموعه افراز می‌شوند. از این زیر-مجموعه‌ها، هر بار یکی برای اعتبارسنجی و مابقی جهت آموزش به‌کار می‌روند. این روال K بار تکرار می‌شود و همه‌ی داده‌ها دقیقاً یک بار برای آموزش و یک بار برای اعتبارسنجی به‌کار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود.

۳-۵ تحلیل مؤلفه‌های اصلی^۱ (PCA)

با توجه به تنوع پارامترهای ورودی به مدل و ارتباط و همبستگی بین برخی از این پارامترها، استفاده از برخی تحلیل‌های آماری می‌تواند منجر به کاهش پراکندگی داده‌ها و بهبود شرایط مدل‌سازی گردد. تحلیل مؤلفه‌ی اصلی یک روش آماری برای کاهش بعد داده‌ها می‌باشد. در این روش یک سری پارامترهای جدید به نام مؤلفه‌های اصلی ایجاد می‌گردد که ترکیبی خطی از متغیرهای موجود در مسأله است. در واقع به وسیله‌ی این روش وابستگی خطی بین متغیرها پیدا می‌شود و می‌توان در فضای چندبعدی داده‌ها، جهاتی را یافت که در راستای این جهات داده‌ها بیش‌ترین واریانس را دارند. بنابراین می‌توان با تصویر کردن داده‌ها بر روی این جهات، ابعاد آن را کاهش داد بدون این‌که اطلاعات قابل توجهی حذف شود. تحلیل مؤلفه‌ی اصلی فضایی تولید می‌کند که در آن داده‌ها براساس مقدار واریانس مرتب می‌شوند. داده‌ها با بیش‌ترین

¹ Principle Component Analysis

واریانس به بُعد اول منتقل می‌شود و بدین ترتیب آخرین مؤلفه‌ی اصلی کم‌ترین واریانس را دارد. انتخاب تعداد مؤلفه‌ی اصلی به اختلاف بیش‌ترین و کم‌ترین مقدار بستگی دارد [۴۱].

تحلیل مؤلفه‌ی اصلی یکی از انواع روش‌های تحلیل داده‌های چند متغیره است که هدف اصلی آن تقلیل بُعد مسأله مورد مطالعه است. یکی از کاربردهای مهم تحلیل مؤلفه‌ی اصلی، در رگرسیون است. با استفاده از تحلیل مؤلفه‌ی اصلی، می‌توان تعداد زیادی متغیر را با تعداد محدودی متغیر مستقل جدید که مؤلفه‌های اصلی نامیده می‌شوند، جایگزین نمود. به این ترتیب بعد مسأله تقلیل می‌یابد [۴۲].

۳-۶ شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی، نوعی شبیه‌سازی ساختار مغز انسان در قالب یک سیستم مصنوعی هستند. این شبکه‌ها از تعداد نسبتاً زیادی عضوهای پردازش‌گر (نرون‌های مصنوعی) تشکیل و به شکلی خاص و منظم طراحی شده‌اند. مجموعه‌ی این نرون‌های مصنوعی، توانایی آموزش دیدن، یادگیری، به خاطر سپردن و تعمیم دادن مجموعه‌ای از داده‌ها را دارد [۴۳].

پردازش در مغز انسان موازی انجام می‌شود. پیاده کردن این ایده در کامپیوترهای دیجیتال موجب شد که بحث شبکه‌های عصبی امروزی مطرح شود [۴۴]. دلایل به‌کارگیری شبکه‌های عصبی مصنوعی را می‌توان به صورت زیر خلاصه کرد:

- می‌توانند اطلاعات نامناسب و همراه با اغتشاش (نویز) را پردازش کنند.
- می‌توانند با سعی و خطا شرایط پیش‌بینی نشده را از دیگر حدود بشناسند و در این شرایط جدید استنتاج‌های قابل قبولی ارائه کنند.

- می‌توانند داده‌های ورودی فراوانی را پردازش کنند و به سرعت ارتباط آن‌ها را به‌دست آورند.
- می‌توانند روابط عوامل ورودی را آموزش ببینند.
- انعطاف‌پذیر هستند، یعنی همانند مغز انسان، می‌توانند با شرایط جدید سازگار شوند.

با توجه به ویژگی‌های منحصر به فرد شبکه‌های عصبی مصنوعی، دامنه‌ی کاربرد این روش‌ها بسیار گسترده است [۴۳]. امروزه قدرت شبکه‌های عصبی به حدی رسیده است که در حل بسیاری از مسائل پیچیده مانند تخمین، تشخیص الگو، طبقه‌بندی و خوشه‌بندی به‌کار می‌روند [۴۴]. باید توجه داشت که روش شبکه‌های عصبی نیز همانند همه‌ی روش‌ها کاستی‌هایی دارد که مهمترین آن‌ها عبارت‌اند از: نیاز به مجموعه‌ای از داده‌ها برای آموزش و نیز آزمودن شبکه، عدم ارائه توضیح و نحوه رسیدن به پاسخ و امکان ارائه یک جواب زودرس و نامناسب.

۳-۶-۱ ساختار شبکه‌های عصبی

ساختار شبکه عصبی، متغیرهایی را مانند تعداد لایه در یک شبکه، تابع تبدیل هر لایه، تعداد نرون‌ها در هر لایه و اتصالات بین نرون‌ها تعریف می‌کند. اجزای یک شبکه عصبی عبارتند از:

الف) نرون مصنوعی: مهم‌ترین عنصر یک شبکه عصبی مصنوعی است که به‌عنوان پرسپترون یا عنصر پردازش‌گر نیز شناخته می‌شود و اجزایی شبیه به نرون‌های طبیعی دارد. یک نرون مصنوعی شامل یک عمل‌گر مجموع و یک عمل‌گر غیر خطی است که تابع تحریک نامیده می‌شود.

ب) لایه‌ها: نرون‌ها در قالب لایه‌ها سازماندهی می‌شوند. یک شبکه عصبی مصنوعی از سه نوع لایه تشکیل شده است و هر لایه مسئول یک کار خاص است.

- لایه ورودی: این لایه اطلاعات را از محیط دریافت و آن‌ها را به لایه میانی منتقل می‌کند. هر نرون موجود در لایه ورودی با همه نرون‌های لایه میانی مرتبط است.

- لایه میانی یا پنهان: این لایه اطلاعات وارد شده از محیط را به شبکه عصبی، پردازش و تجزیه و تحلیل می‌کند.

- لایه خروجی: نتیجه تجزیه و تحلیل اطلاعات توسط لایه میانی را دریافت و به شکلی معنی‌دار تبدیل می‌کند و دوباره به محیط بر می‌گرداند.

تعداد نرون‌های لایه ورودی برابر با تعداد متغیرهایی است که به‌عنوان ورودی شبکه، تعریف می‌شوند. نرون‌های لایه میانی مهم‌ترین پردازش‌گرهای شبکه عصبی هستند و در واقع ورودی‌های شبکه را به خروجی معین تبدیل می‌کنند. تعداد نرون‌های لایه خروجی نیز از همان قانون تعیین نرون‌های لایه ورودی تبعیت می‌کند و برابر با تعداد متغیرهای خروجی است [۴۳].

تعیین تعداد لایه‌های پنهان، تعداد نرون‌های هر لایه پنهان و همچنین نوع توابع انتقال یا ترکیب، یکی از مهم‌ترین مراحل طراحی شبکه است که به پیچیدگی شبکه و کاربرد آن بستگی دارد. برای تعیین بهینه تعداد لایه‌ها، نرون‌ها و نوع توابع انتقال، از روش حدس و خطا استفاده می‌شود. بدین ترتیب که ساختارهای مختلف در نظر گرفته می‌شوند و بهترین مدل بر اساس مبانی آماری مانند ضریب همبستگی^۱ و میانگین مربعات خطا^۲ برای داده‌های آموزش و آزمون، انتخاب می‌گردد.

ضریب همبستگی: نسبت انحراف داده‌ها از مدل را بیان می‌کند و هرچه مقدار آن به یک نزدیک‌تر باشد، انطباق داده‌ها با مدل مورد نظر بیش‌تر خواهد بود. این نوع همبستگی به رگرسیون نیز معروف است.

1 correlation coefficient

2 Mean Squared Errors (MSE)

$$R^2 = 1 - \frac{\sum_{i=1}^N (obs - cal)^2}{\sum_{i=1}^N (obs - calave)^2} \quad (14-3)$$

که در آن N تعداد داده‌ها، obs مقدار خروجی واقعی مشاهده شده، cal مقدار خروجی محاسبه شده با مدل و $calave$ میانگین مقادیر خروجی محاسبه شده با مدل می‌باشد.

میانگین مربعات خطا: هر چه مقدار آن به سمت صفر میل کند، نشانه‌ی موفقیت بیش‌تر مدل در پیش‌بینی نتایج می‌باشد [۴۴].

$$MSE = \frac{1}{N} \sum_{i=1}^N (obs - cal)^2 \quad (15-3)$$

همبستگی و وابستگی بین متغیرها: ضریب همبستگی، یکی از معیارهای مورد استفاده در تعیین همبستگی دو متغیر است. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین ۱- و ۱ متغیر است و در صورت عدم وجود رابطه بین دو متغیر، برابر صفر می‌باشد.

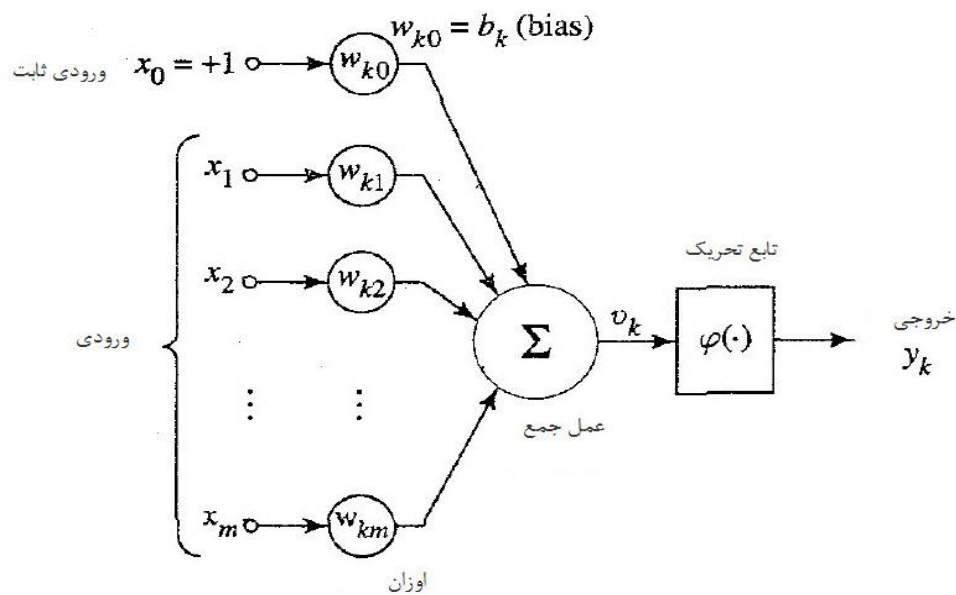
۳-۶-۲ مدل ریاضی یک نرون در شبکه عصبی مصنوعی

مدل ریاضی نرون (سلول عصبی) از دو بخش اساسی تابع ترکیب و تابع تحریک تشکیل شده است.

۳-۶-۲-۱ تابع ترکیب

ورودی‌های یک نرون از نرون‌های مجاور و یا منابع خارجی سرچشمه می‌گیرند. نخست، هر یک از ورودی‌ها در وزن مخصوص به خود، ضرب می‌شوند و آن‌گاه، به کمک یک عمل‌گر مجموع، مجموع وزن‌دار شده

ورودی‌ها به‌دست می‌آیند. این وزن‌های مربوط به هر ورودی، شبیه به قدرت سیناپس‌ها در نرون‌های طبیعی عمل و مقدار تأثیر هر ورودی را در خروجی نرون تعیین می‌کنند. وزن‌ها ابتدا دلخواه و اغلب به‌صورت تصادفی مقداردهی و بعد در مرحله آموزش شبکه اصلاح و بهینه می‌شوند. طرح کلی یک نرون در شکل ۳-۳ نمایش داده شده است.



شکل ۳-۳ نمایش طرح کلی یک نرون [۴۵]

تابع ترکیب مجموع وزن‌دار ورودی‌های یک نرون است که معمولاً برای در نظر گرفتن دیگر اثرات مستقل از ورودی‌ها یک مقدار آستانه^۱ نیز به آن اضافه می‌شود.

$$f = \sum (W_i X_i) - \theta \quad (۱۶-۳)$$

W_i و X_i به‌ترتیب ورودی نام و وزن مربوط به آن است. θ نیز همان مقدار آستانه است.

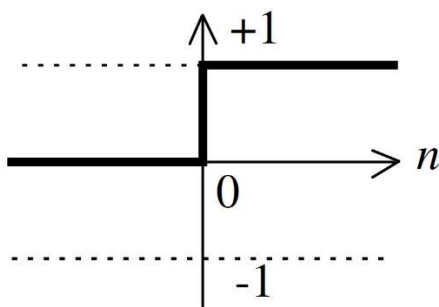
^۱ bias

۳-۶-۲-۲ تابع تحریک

تابع تحریک (تابع انتقال یا فعال سازی) فرایند تحریک در یک نرون را کنترل و مشخص می کند که خروجی یک نرون، فعال باشد یا نباشد. بیش تر وقت ها، فعال بودن خروجی یک نرون به مثبت یا منفی بودن ورودی تابع تحریک وابسته است. ورودی تابع تحریک همان خروجی تابع ترکیب است. بنابراین اگر تابع تحریک با a و خروجی نرون با Y نشان داده شود، آن گاه [۴۳]:

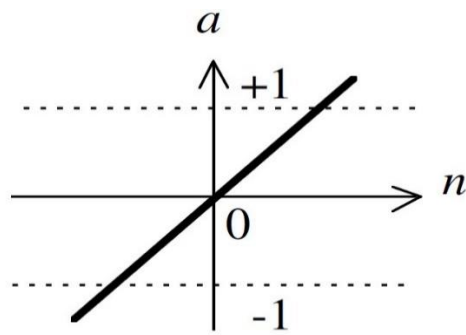
$$Y=a(f)=a(\sum(W_iX_i) - \theta) \quad (۱۷-۳)$$

توابع تحریک پله ای، خطی، سیگموئیدی و تانژانت هایپربولیک به ترتیب در شکل های ۳-۴ تا ۳-۷ نشان داده شده است [۴۰].



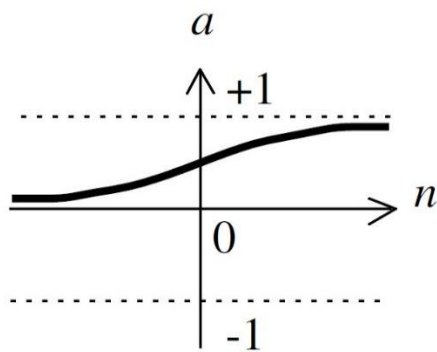
$$a = \text{hardlim}(n)$$

شکل ۳-۴ نمایش تابع تحریک پله ای



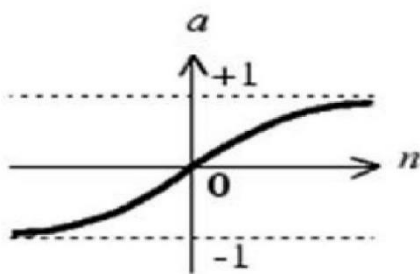
$$a = \text{purelin}(n)$$

شکل ۳-۵ نمایش تابع تحریک خطی



$$a = \text{logsig}(n)$$

شکل ۳-۶ نمایش تابع تحریک سیگموئیدی



$$a = \text{tansig}(n)$$

شکل ۳-۷ نمایش تابع تحریک تانژانت هایپربولیک

غیر خطی بودن تابع تحریک کمک می کند تا خروجی یک شبکه عصبی بتواند بر حسب هر تابع به شدت غیر خطی از مجموعه متغیرهای ورودی، بیان شود.

۳-۶-۳ الگوریتم یادگیری

قانون یا الگوریتم یادگیری، به معنی روش اصلاح مقادیر وزن ها و بایاس های شبکه است. هدف قانون یادگیری، آموزش شبکه برای انجام یک کار خاص است. انواع گوناگونی از الگوریتم های یادگیری وجود دارند که در یکی از سه گروه زیر قرار می گیرند.

۳-۶-۳-۱ الگوریتم یادگیری نظارتی^۱ (با سرپرست)

در یادگیری به کمک مثال یا یادگیری نظارتی، مثال هایی از رفتار ایده آل شبکه (که مجموعه آموزشی نامیده می شود) به قانون یادگیری اعمال می شود. در این الگوریتم با اعمال ورودی ها به شبکه، خروجی های شبکه با خروجی های مطلوب مقایسه می شوند. سپس، قانون یادگیری برای تنظیم متغیرهای شبکه به کار می رود، به گونه ای که خروجی های شبکه را هر چه بیش تر به اهداف مطلوب نزدیک کند. هنگامی که به یک شبکه ای آموزش ندیده الگویی عرضه شود، خروجی های تصادفی تولید می کند. سپس تابع خطایی تعریف می شود که تفاوت خروجی واقعی و خروجی مطلوب را نشان دهد. در هر بار ارائه ی ورودی ها به شبکه، وزن های شبکه بر اساس تابع خطا اصلاح می شوند. چون خروجی مطلوب از قبل معلوم است، این نوع یادگیری را یادگیری نظارتی می نامند.

^۱ supervised

۳-۶-۱-۱-۱ الگوریتم پس انتشار خطا^۱

الگوریتم پس انتشار خطا یک روش یادگیری نظارت شده است که از آن برای آموزش شبکه‌های عصبی رو به جلو، مانند شبکه‌های پرسپترون چند لایه^۲ استفاده می‌شود. این الگوریتم از دو مسیر مختلف موسوم به مسیرهای رفت و برگشت، تشکیل شده است. در مسیر رفت، ابتدا بردار ورودی به شبکه اعمال و با در نظر گرفتن مقادیر کاملاً تصادفی برای اوزان و بایاس‌های شبکه، تأثیرات آن از لایه‌های میانی به لایه‌های خروجی منتشر می‌شود. بردار تشکیل شده در لایه‌ی خروجی، پاسخ شبکه را بیان می‌کند. در مسیر برگشت، ابتدا خطای بین جواب‌های متناظر شبکه با جواب‌های آموزشی با استفاده از رابطه‌ی موردنظر کاربر که به تابع عملکرد نیز مشهور است، محاسبه می‌شود. سپس با استفاده از روش‌های موجود، مشخصه‌های شبکه برای کاهش خطای محاسبه‌شده اصلاح می‌شوند.

روش‌های مختلفی برای الگوریتم پس انتشار خطا مطرح هستند. در تمام این روش‌ها از گرادیان تابع عملکرد برای کاهش خطای شبکه‌ی موردنظر و در نتیجه تنظیم مشخصه‌های شبکه استفاده می‌شود. برخی از روش‌های مطرح در این زمینه، در ادامه بیان می‌گردد.

- روش گرادیان نزولی^۳

- روش گرادیان نزولی با مومنتم^۴

- روش گرادیان مزدوج^۵

۱ back propagation

۲ Multi Layer Perceptron (MLP)

۳ gradient descent

۴ gradient descent with momentum

۵ conjugate gradient

- روش شبه نیوتن^۱

- روش لونبرگ-مارکوارت^۲

سرعت عملکرد دو روش اول کم است، در حالی که دیگر روش‌ها از دو روش اول کارا تر هستند. از آن‌جا که روش لونبرگ-مارکوارت سریع‌ترین روش در بین این روش‌هاست، لذا معمولاً از آن برای تنظیم مشخصه‌های مدل استفاده می‌شود [۴۴].

۳-۶-۳-۲ الگوریتم یادگیری تقویت‌شده یا تشدید

یادگیری تقویت‌شده مشابه یادگیری نظارتی است، با این تفاوت که به جای دادن خروجی واقعی مربوط به هر ورودی به قانون یادگیری، تنها یک عدد یا درجه که معرف مقدار عملکرد شبکه روی مجموعه‌ای از ورودی‌هاست، به شبکه داده می‌شود. قانون یادگیری، متغیرهای شبکه را برای بیشینه کردن این عدد تنظیم می‌کند.

۳-۶-۳-۳ الگوریتم یادگیری غیرنظارتی^۳ (بدون سرپرست)

در بسیاری از کاربردها باید از شبکه خواست که داده‌های آموزشی را به تنهایی طبقه‌بندی کند. برای این کار دو شرط اساسی باید در نظر گرفته شوند: عضویت در یک طبقه به طور کلی به معنای داشتن ویژگی‌های مشترک است و شبکه می‌تواند این ویژگی‌های مشترک را در گستره‌ی داده‌های ورودی تمیز دهد. شبکه‌های

^۱ Quasi-Newton

^۲ Levenberg-Marquardt

^۳ unsupervised

خودسازمان ده از این نوع شبکه‌ها هستند و از آموزش بدون نظارت بهره می‌گیرند. در این یادگیری خروجی - های مطلوب در دسترس نیستند و متغیرهای شبکه به کمک پاسخ خود شبکه اصلاح می‌شوند. با استفاده از این نوع الگوریتم‌ها، شبکه ویژگی‌هایی از ورودی‌ها را که باعث طبقه‌بندی می‌شوند حدس می‌زند و نوعی عملیات خوشه‌بندی را انجام می‌دهند [۴۳].

۳-۶-۴ انتخاب داده‌های آموزشی

انتخاب یک مجموعه آموزشی مناسب، بسیار مهم است و باید دامنه‌ای وسیع از داده‌های فوق را در برگیرد. زیرا کیفیت جواب نهایی شبکه، به مثال‌های آموزشی مرحله‌ی آموزش وابسته است.

۳-۶-۵ اهمیت نسبی ورودی‌های شبکه

تجسم تصویری واضح از چگونگی کارکرد شبکه عصبی در ارائه جواب، غیرممکن است. در واقع ویژگی غیرخطی نرون‌های شبکه که آن‌ها را قدرتمندتر می‌کند، از تولید قوانین ساده جلوگیری می‌نماید. روش تحلیل حساسیت می‌تواند برای شناخت چگونگی کار مدل‌ها به کار رود. این روش مقدار حساسیت خروجی شبکه را به هر یک از ورودی‌های آن تعیین می‌نماید. شبکه به بعضی از ورودی‌ها چندان حساس نیست و برعکس ورودی‌هایی نیز وجود دارند که در حالت‌های مختلف تأثیری متفاوت بر خروجی می‌گذارند.

۳-۶-۶ معیارهای پایان آموزش شبکه

چند روش برای پایان دادن به روند آموزش شبکه وجود دارند که انتخاب هر یک، به مسأله و هدف از طراحی شبکه بستگی دارد. مهم‌ترین این روش‌ها در ادامه بیان می‌گردد.

- هنگامی که میانگین مربعات خطا در هر دوره^۱ کمتر از مقداری مشخص شود، آموزش شبکه متوقف می‌شود. باید توجه کرد که هر دوره شامل یک بار خوراندن همه‌ی نمونه‌های یادگیری، به شبکه است.
- تغییرات تابع خطا در دوره‌های متوالی، از یک مقدار از پیش تعیین شده کوچک‌تر می‌شود. به عبارتی، هنگامی که تغییرات در میانگین مربعات خطا (یا هر نوع تابع خطای دیگر) در دوره‌های متوالی به اندازه کافی کوچک (معمولاً کمتر از ۰/۱ تا ۰/۰۱ درصد) باشند، آموزش شبکه متوقف می‌شود و تنظیم و اصلاح وزن‌ها پایان می‌یابد.
- روش اعتبارسنجی متقاطع^۲: در این روش درصدی از داده‌های آموزش با نام داده‌های اعتبارسنجی جدا می‌شود. آزمایش شبکه با داده‌های اعتبارسنجی هم‌زمان با آموزش، در هر دوره انجام می‌شود و زمانی که خطا در داده‌های اعتبارسنجی شروع به بالا رفتن می‌کند، آموزش شبکه متوقف می‌گردد. در حقیقت این روش از بیش‌برازش شدن^۳ شبکه جلوگیری می‌کند [۴۳].

۳-۷ سری‌های زمانی

تحلیل سری‌های زمانی به‌طور نظری و عملی از سال‌های ۱۹۷۰ به بعد، برای پیش‌بینی و کنترل، به سرعت توسعه پیدا کرده است. این تحلیل معمولاً به داده‌هایی مربوط می‌شود که مستقل نبوده و به‌طور متوالی به هم وابسته‌اند. همین وابستگی بین مشاهدات متوالی است که مورد توجه قرار می‌گیرد و بیش‌تر کاربرد آن،

^۱ epoch

^۲ cross validation

^۳ over fitting

در پیش‌بینی حوادث آینده بر اساس رخداد‌های گذشته خواهد بود. سری زمانی مجموعه‌ای از مشاهدات است که بر حسب زمان مرتب شده باشد. معمولاً آن را به صورت زیر نشان می‌دهند:

$$X_{t1}, X_{t2}, \dots, X_{tn} \quad (3-18)$$

۳-۷-۱ انواع سری زمانی

سری‌های زمانی را معمولاً به صورت گسسته یا پیوسته بررسی می‌کنند. اگر مشاهدات به صورت پیوسته بر حسب زمان در نظر گرفته شوند، سری زمانی حاصل را پیوسته و اگر مشاهدات به طور منظم و در فاصله‌های مساوی ثبت گردند، سری زمانی را گسسته می‌نامند.

۳-۷-۲ اجزای تشکیل دهنده سری زمانی

با دقت در سری زمانی و با توجه به نمودار آن، اجزای تشکیل دهنده سری زمانی را می‌توان شناخت و اندازه‌گیری کرد. اجزای اصلی سری زمانی به این شرح هستند.

۳-۷-۲-۱ روند

روند عبارت از تغییرات دراز مدت در میانگین سری زمانی است. در این صورت افت و خیزهای سری زمانی نادیده گرفته شده و نمای کلی آن مورد توجه قرار می‌گیرد. از مطالعه داده‌ها در یک دوره‌ی طولانی، می‌توان یک ایده‌ی کلی نسبت به رفتار پدیده‌ی مورد بررسی به دست آورد که در پیش‌بینی آینده به ما کمک کند.

۳-۷-۲-۲ تغییرات فصلی

تغییرات فصلی تغییراتی هستند که در دوره‌های تناوبی کوتاه پیش می‌آیند. این تغییرات مربوط به عواملی هستند که به طریقی منظم و چرخه‌ای، روی یک دوره‌ی کم‌تر از یک سال عمل می‌کنند.

۳-۷-۲-۳ تغییرات دوره‌ای

حرکات نوسانی در یک سری زمانی با دوره‌ی نوسان بیش‌تر از یک سال را تغییرات دوره‌ای می‌نامند.

۳-۷-۲-۴ تغییرات نامنظم

در هر سری زمانی عامل دیگری وجود دارد که آن را تغییرات نامنظم یا تصادفی می‌نامند. این تغییرات کاملاً تصادفی بوده و نتیجه‌ی عواملی غیر قابل پیش‌بینی هستند که به طریقی نامنظم عمل می‌کنند.

با توجه به عوامل فوق، الگوی سری زمانی را می‌توان به‌صورت زیر در نظر گرفت:

$$X_t = T. S. C. I \quad (۳-۱۹)$$

که در آن X_t مقدار مشاهده شده در لحظه t ، T روند، S تغییرات فصلی، C تغییرات دوره‌ای و I تغییرات نامنظم است.

۳-۷-۳ الگوسازی برای یک سری زمانی

در تجزیه و تحلیل یک سری زمانی، شناختن و ساختن یک الگو بر مبنای داده‌های سری زمانی اهمیت ویژه‌ای دارد. مراحل الگوسازی یک سری زمانی عبارتند از شناخت، برآورد و بررسی درستی تشخیص.

در تحلیل سری زمانی، اولین مرحله رسم نمودار داده‌ها می‌باشد. با امتحان و بررسی دقیق نمودار سری زمانی می‌توان ایده‌ی خوبی در مورد این که روند، نوسانات فصلی، نقاط پرت و واریانس غیرثابت وجود دارند یا خیر، به دست آورد. برای برآورد کردن پارامترهای یک الگوی سری زمانی، می‌توان از روش‌های متداول برآورد نظیر روش گشتاورها، روش درست نمایی ماکزیمم و روش کم‌ترین مربعات استفاده کرد.

یکی از اهداف ساختن یک الگو برای یک سری زمانی این است که بتوان مقدار آینده سری را با توجه به گذشته‌ی آن پیش‌بینی نمود. بر مبنای گذشته سری تا زمان t ، یعنی $X_1, X_2, \dots, X_{t-1}, X_t$ ، برای پیش‌بینی مقدار X_{t+1} یعنی مقدار سری در l واحد زمانی بعد، t را مبدأ زمان و l زمان تقدم در نظر گرفته می‌شود. بنابراین پیش‌بینی l مرحله بعد $X_t(l)$ با X_t نشان داده می‌شود و به صورت زیر تعریف می‌گردد [۴۶]:

$$X_t(l) = E(X_{t+l} \mid X_t, X_{t-1}, \dots, X_1) \quad (۳-۲۰)$$

۳-۸ پیاده‌سازی مدل‌های شبکه عصبی و سری‌های زمانی تحت

شبکه عصبی، در نرم افزار

جهت پیاده‌سازی مدل‌های شبکه عصبی و سری‌های زمانی در نرم افزار متلب گام‌های متفاوتی نیاز به کدنویسی دارد :

- بارگذاری داده‌ها

- تقسیم داده‌ها
- نرمال سازی داده‌ها
- ساختار شبکه
- آموزش شبکه
- ارزیابی شبکه

۳-۸-۱ بارگذاری داده‌ها

در بارگذاری داده‌ها، یک فایل اکسل از داده‌های در دسترس جمع آوری می‌شود. نرم افزار این فایل را به صورت یک ماتریس شناسایی می‌کند. ۲۱ ستون اول ماتریس به عنوان ورودی و ستون آخر به عنوان خروجی واقعی، تعریف می‌شوند. با فراخوانی این فایل در محیط متلب، داده‌های ورودی و خروجی برای اعمال به مدل جداسازی می‌شوند.

۳-۸-۲ تقسیم داده‌ها

دسترسی به داده، یکی از مهم‌ترین فاکتورها در آموزش سیستم‌های هوشمند می‌باشد. داده‌ها باید تمامی حالات مسأله را مدنظر قرار داده و همچنین دارای الگوهای کافی جهت آموزش و ارزیابی باشند. داده‌ها به سه دسته‌ی آموزش، اعتبارسنجی و آزمون تقسیم می‌شوند. داده‌های هر سه دسته جهت امکان بهره گیری

از تمام فضای داده‌ها، به صورت تصادفی و پراکنده انتخاب می‌گردند. در این پایان نامه ۵۵ درصد داده‌ها به آموزش شبکه، ۱۵ درصد به اعتبارسنجی و ۳۰ درصد به آزمون، اختصاص داده شده است.

۳-۸-۳ نرمال سازی داده‌ها

اختلاف در دامنه تغییرات و ابعاد پارامترهای مورد استفاده در مدل، باعث ایجاد محدودیت‌هایی در فرآیند آموزش شبکه خواهد شد. برای رفع این مشکل، از روش‌های نرمال سازی استفاده می‌شود. توابع نرمال سازی، داده‌ها را در یک بازه مشخص همگن می‌کنند. در این پژوهش رابطه‌ی زیر جهت نرمال سازی داده‌ها، مورد استفاده قرار گرفت.

$$X = 2 * \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) - 1 \quad (۲۱-۳)$$

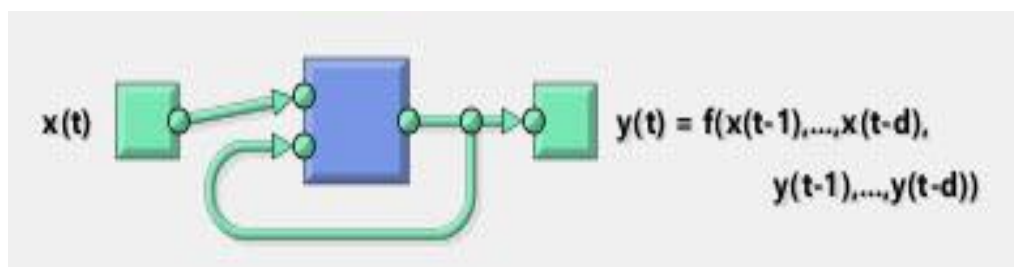
در رابطه فوق، x مقدار میدانی هر یک از ورودی‌ها و x_{min} و x_{max} کم‌ترین و بیش‌ترین مقادیر هر ستون از داده‌ها، که مربوط به یک ورودی یا خروجی می‌باشد، را مشخص می‌کند. بنابراین تمام داده‌های موجود در بازه $[-۱, ۱]$ همگن یا نرمال می‌شوند [۴۰].

۳-۸-۴ ساختار شبکه

ساختار شبکه عصبی (همچنین در مورد سری‌های زمانی تحت شبکه عصبی) در موفقیت آن بسیار مؤثر است. در این مدل‌ها برای شناسایی الگوهای گیر و عدم گیر، از یک لایه میانی در ساختار شبکه استفاده گردید. به وسیله آزمون و خطا، تعداد نرون لایه‌ی میانی بهینه برای هر مدل تعیین گردید.

۳-۸-۵ آموزش شبکه

برای آموزش هر دو مدل شبکه عصبی و سری‌های زمانی، از روش لونیگ-مارکوارت استفاده گردید. تفاوت سری زمانی در استفاده از خروجی‌ها و ورودی‌های قبلی، بر اساس تأخیر زمانی است که کاربر برای آن تعیین می‌نماید. این امر به علت دینامیک بودن مدل سری‌های زمانی می‌باشد. نمایی کلی از مدل سری زمانی تحت شبکه عصبی در شکل ۳-۸ نشان داده شده است.



شکل ۳-۸ نمایش مدل سری زمانی تحت شبکه عصبی، در نرم افزار متلب. در این شکل، x ورودی، y خروجی و d تأخیر زمانی می‌باشد.

۳-۸-۶ ارزیابی شبکه

شبکه با داده‌های آزمون، عملکرد خود را ارزیابی می‌نماید. در این پژوهش به دلیل وجود داده‌های کافی، ۳۰ درصد از کل داده‌ها جهت آزمون شبکه به کار رفت. این عمل قابلیت تعمیم مدل را افزایش می‌دهد.

فصل چهارم:

نتایج شناسایی و پیش‌بینی

گیر رشته‌ی حفاری

۴-۱ مقدمه

۴-۲ پردازش بر روی داده‌ها

۴-۳ نتایج الگوریتم K نزدیک‌ترین همسایه

۴-۴ نتایج الگوریتم ماشین بردار پشتیبان

۴-۵ نتایج الگوریتم‌های دسته‌بندی به همراه تحلیل مؤلفه اصلی

۴-۶ تشخیص الگو و پیش‌بینی به وسیله‌ی شبکه عصبی

۴-۷ پیش‌بینی به وسیله‌ی سری زمانی

۴-۸ بررسی پارامترهای ورودی به مدل

۴-۱ مقدمه

در این فصل ابتدا پردازش و محاسبات آماری بر روی داده‌های مربوط به پارامترهای عملیات و سیال حفاری انجام و نمایش داده می‌شود. این پردازش شامل محاسبه حداقل، حداکثر و میانگین مقادیر مربوط به هر پارامتر می‌باشد. سپس نمودار هیستوگرام هر پارامتر رسم می‌گردد. پس از بیان اطلاعات داده‌ها، توسط الگوریتم‌های دسته‌بندی مانند ماشین بردار پشتیبان و K نزدیک‌ترین همسایه، مجموعه داده‌های ورودی بر اساس الگوهای گیر و یا عدم گیر دسته‌بندی می‌شوند. جهت بهبود عملکرد این الگوریتم‌ها از تحلیل مؤلفه‌ی اصلی استفاده و نتایج جدید بیان می‌گردد. سپس با ساخت مدل‌های شبکه عصبی و سری زمانی، پدیده گیر لوله توسط این ابزارها، پیش‌بینی می‌گردد. در این پژوهش از شبکه عصبی پرسپترون چند لایه و سری‌های زمانی تحت شبکه عصبی استفاده شده است.

۴-۲ پردازش بر روی داده‌ها

پردازش بر روی داده‌های مرتبط با هر پارامتر ورودی انجام و حداقل، حداکثر و میانگین داده‌ها محاسبه شد. پارامترهای ورودی شامل پارامترهای عملیات، پارامترهای مرتبط با مسیر چاه و پارامترهای سیال حفاری می‌باشند. نتایج این محاسبات در جدول ۴-۱ بیان شده است.

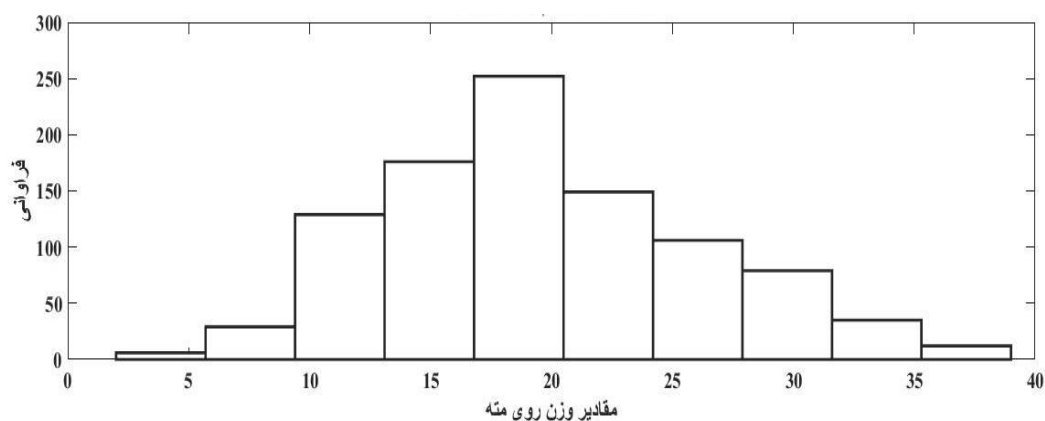
جدول ۴-۱ مقادیر حداقل، حداکثر و میانگین پارامترهای ورودی

نام پارامتر	حداقل	حداکثر	میانگین
وزن روی مته	۲	۳۹	۱۹/۶۵

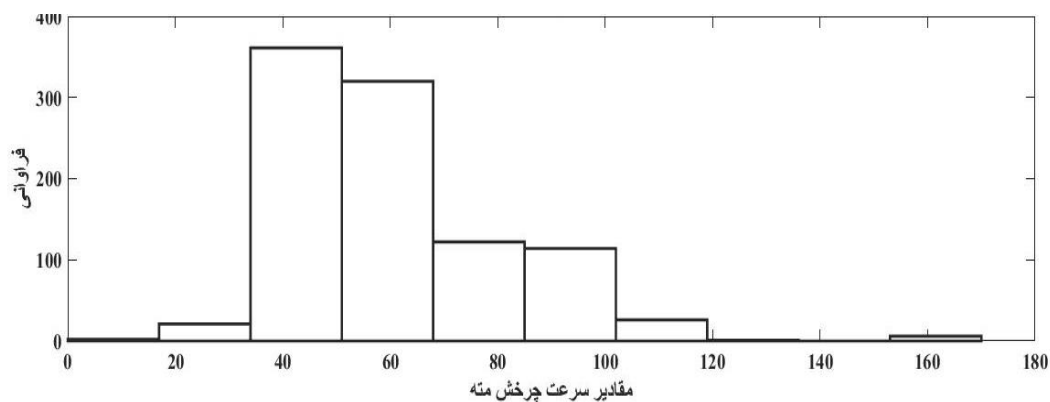
ادامه جدول ۴-۱ مقادیر حداقل، حداکثر و میانگین پارامترهای ورودی

۶۱/۶۷۲	۱۷۰	۰	سرعت چرخش مته
۵۸/۱۷۵	۱۹۰	۰	سرعت چرخش موتور درون چاهی
۸۴۰/۰۲	۱۴۰۰	۱۴۲	نرخ پمپ شدن سیال
۲۴۶۵/۳	۴۵۰۶	۳۱۸	فشار پمپ
۵/۹۸	۱۸	۰/۰۲	نرخ نفوذ حفاری
۴۹/۷۱	۸۰/۶۸	۲/۹	زاویه انحراف مته
۲۰۸/۵۸	۳۵۳/۹	۱/۹۱	آزیموت مته
۱۴۷۷	۳۵/۱۵	۵/۸۱	میزان انحراف چاه از نقطه شروع
۲۷۵۶/۶	۵۱۸۶	۳۱۸	عمق چاه
۱۰/۹۱	۱۷/۲۷	۸/۵۵	چگالی گل
۵۰/۲۶	۱۷۹	۲۶	گرانروی قیف مارش گل
۲۰/۷۷	۹۷	۳/۶	گرانروی پلاستیکی
۲۴/۴۱	۷۶	۱	نقطه‌ی واروی
۷/۰۷	۷۳	۱	استحکام ژلی اولیه
۱۰/۸۳	۱۴۳	۲	استحکام ژلی پس از ۱۰ دقیقه
۹/۵۹	۱۱/۵	۷/۵	میزان اسیدی یا بازی بودن گل
۹۴/۴۹	۴۰۰	۱۰/۲	میزان نمک (کلر) گل
۶۳۷/۴۴	۷۲۴۰	۶۰	میزان یون Ca^{2+}
۱۲/۵۷	۲۳/۵	۸/۵	قطر چاه

جهت نمایش فروانی داده‌های مرتبط با هر پارامتر در مقادیر مشخص، نمودار مستطیلی (هیستوگرام) پارامترهای ورودی رسم گردیدند. نمودار هیستوگرام پارامترهای عملیات حفاری در شکل ۴-۱ نشان داده شده است. این شکل شامل نمودارهای هیستوگرام پارامترهای وزن روی مته، سرعت چرخش مته، سرعت چرخش موتور درون چاهی، نرخ پمپ شدن سیال، فشار پمپ، نرخ نفوذ حفاری، عمق چاه و قطر چاه می‌باشد.

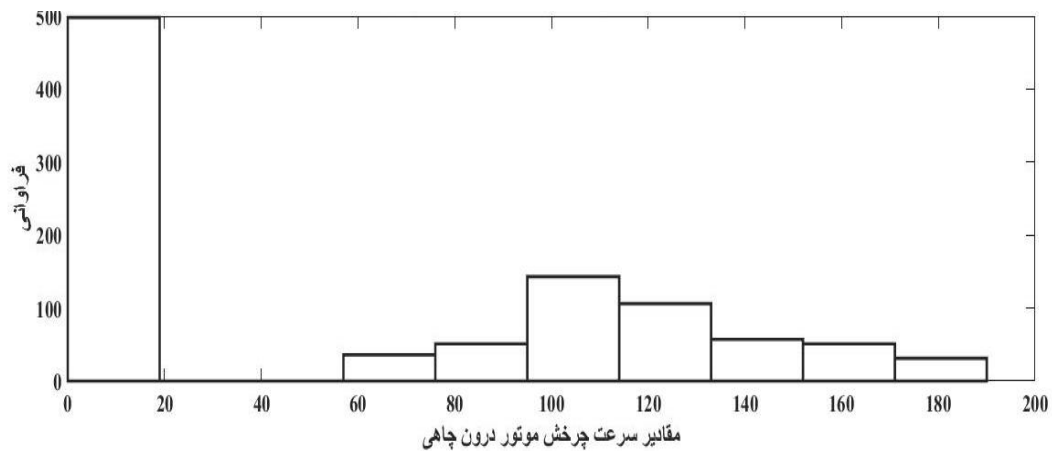


الف

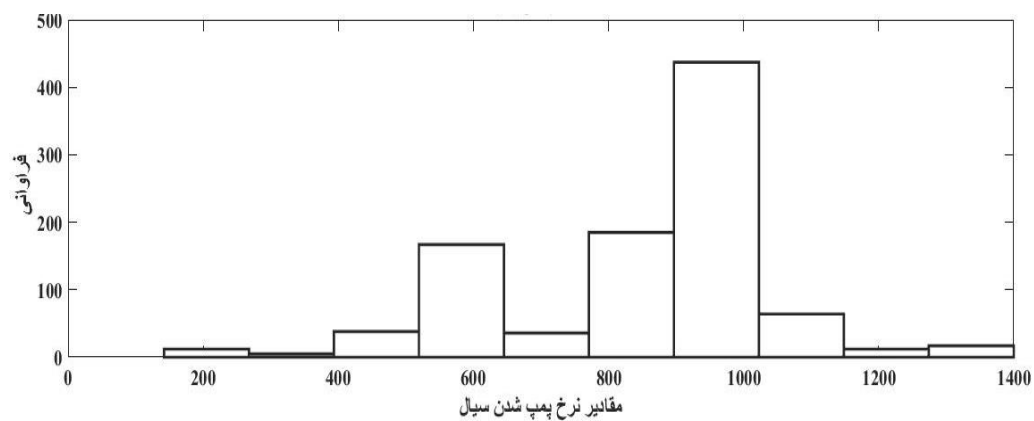


ب

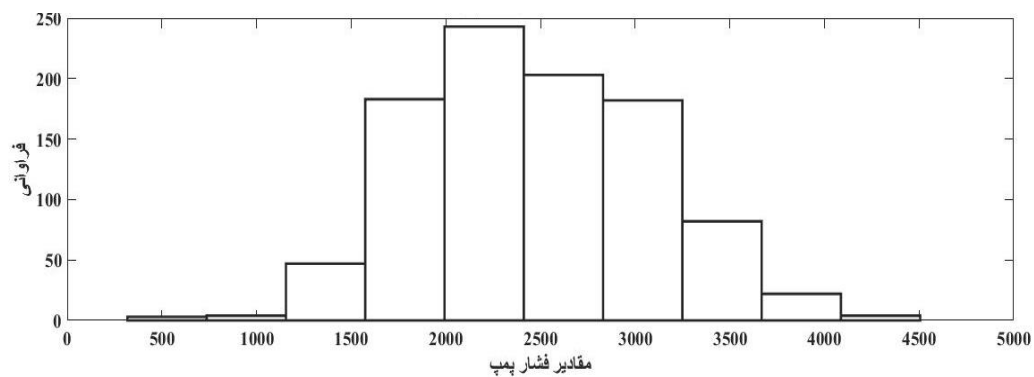
شکل ۴-۱ نمودار هیستوگرام پارامترهای عملیات حفاری. الف- نمودار هیستوگرام وزن روی مته بر حسب کیلو پوند، ب- نمودار هیستوگرام سرعت چرخش مته بر حسب دور در دقیقه



ج



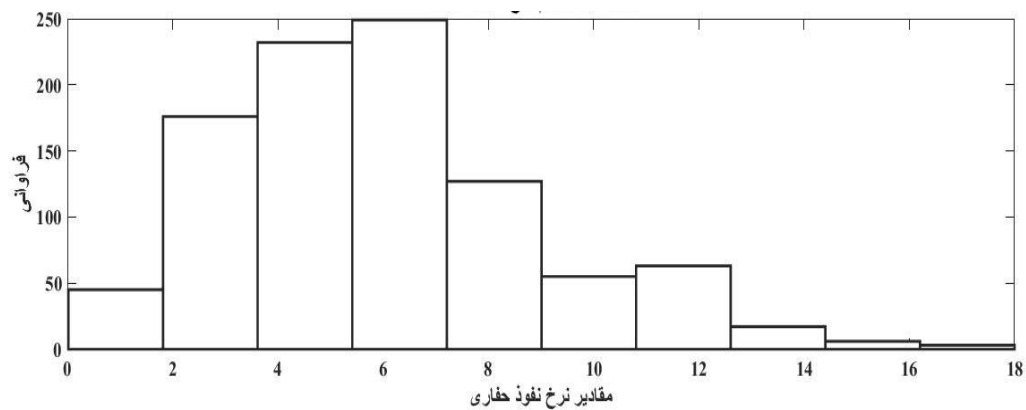
د



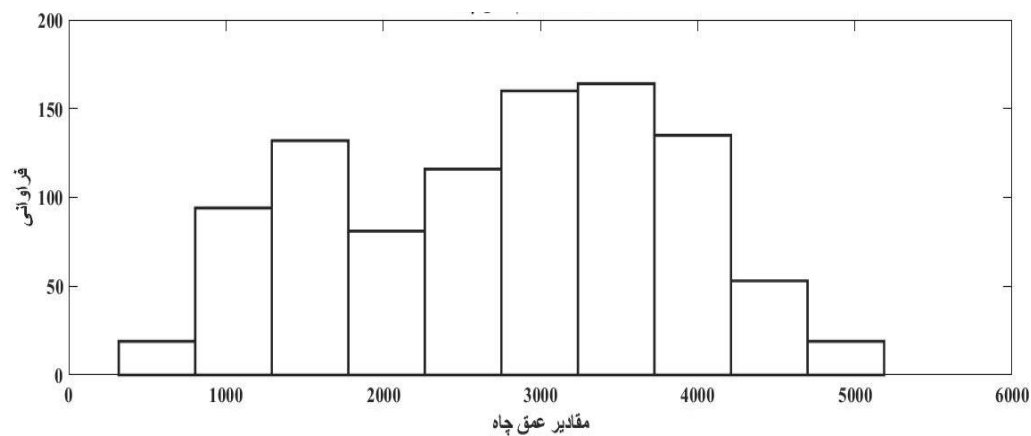
ه

ادامه‌ی شکل ۴-۱ نمودار هیستوگرام پارامترهای عملیات حفاری. ج- نمودار هیستوگرام سرعت چرخش موتور درون چاهی، د-

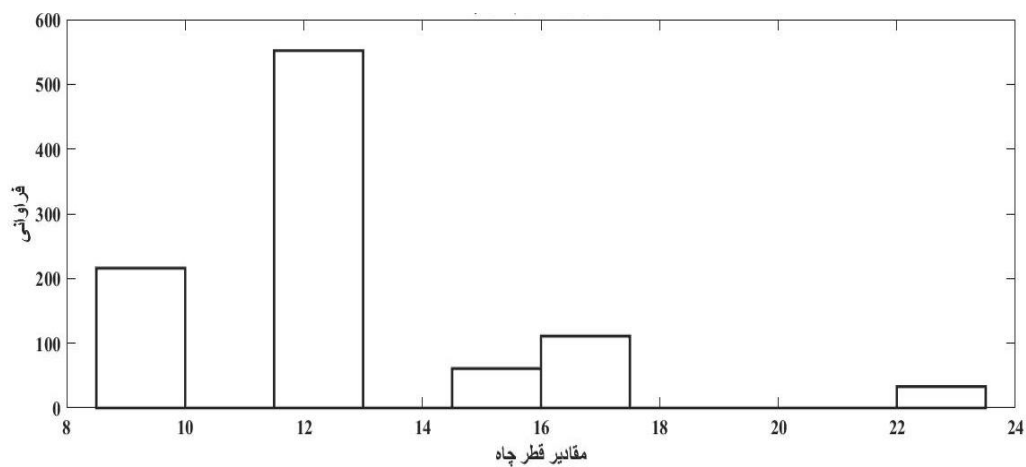
نمودار هیستوگرام نرخ پمپ شدن سیال بر حسب گالن در دقیقه، ه- نمودار هیستوگرام فشار پمپ بر حسب psi



9



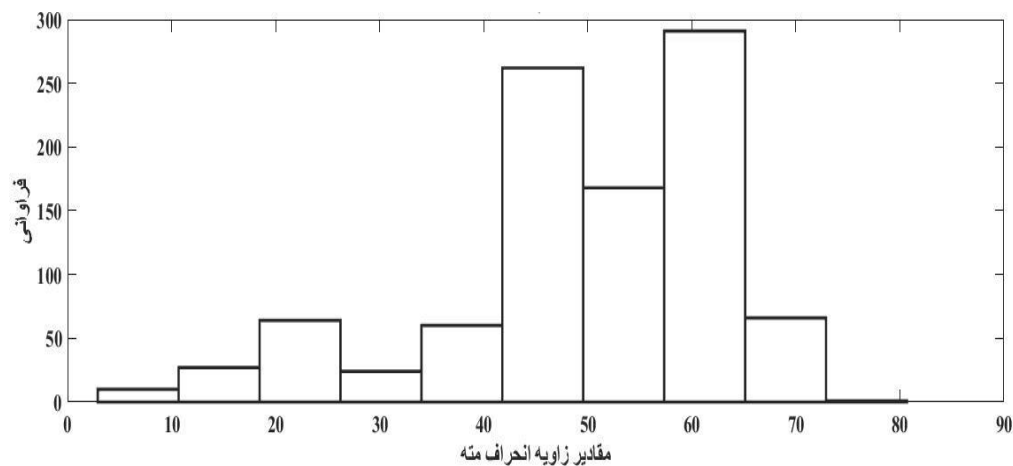
ز



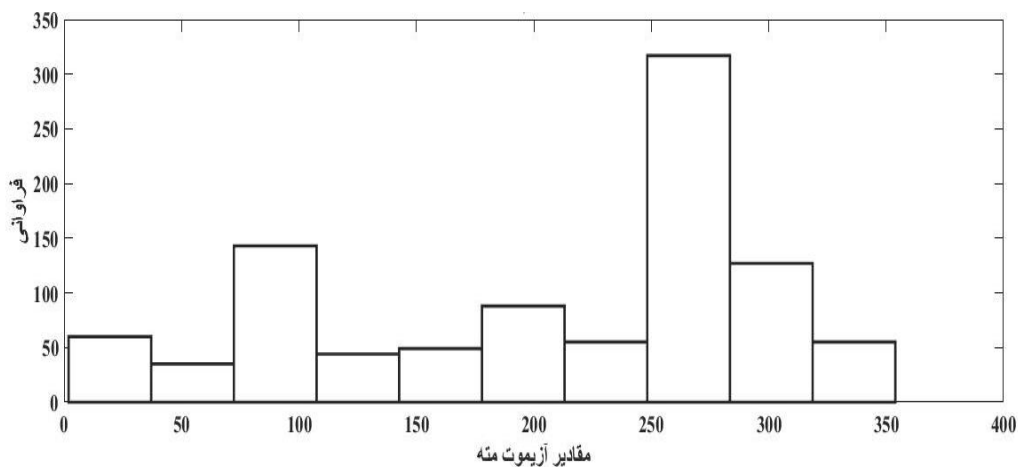
ح

ادامه‌ی شکل ۴-۱ نمودار هیستوگرام پارامترهای عملیات حفاری. و- نمودار هیستوگرام نرخ نفوذ حفاری بر حسب متر بر دقیقه،
 ز- نمودار هیستوگرام عمق چاه بر حسب متر، ح- نمودار هیستوگرام قطر چاه بر حسب اینچ

نمودار هیستوگرام پارامترهای حفاری مرتبط با مسیر چاه در شکل ۴-۲ نمایش داده شده است. این شکل شامل نمودار پارامترهای زاویه انحراف مته، آزیموت مته و میزان انحراف چاه از نقطه شروع، می‌باشد.

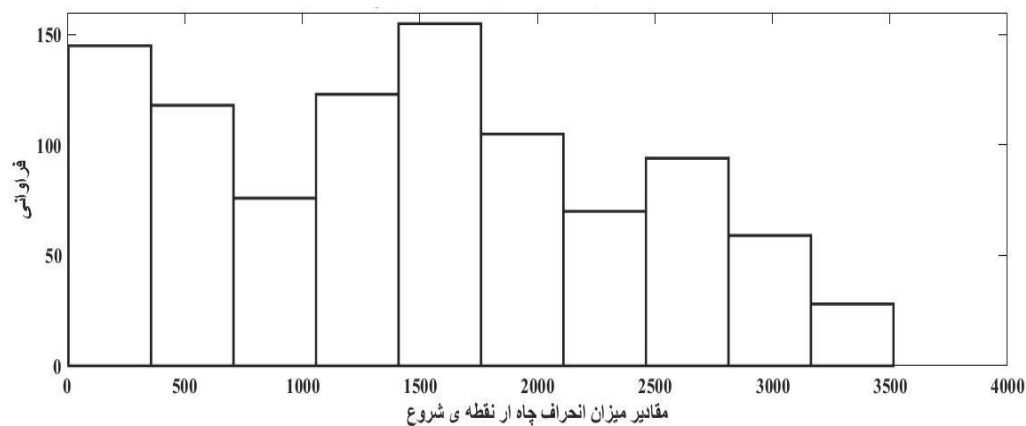


الف



ب

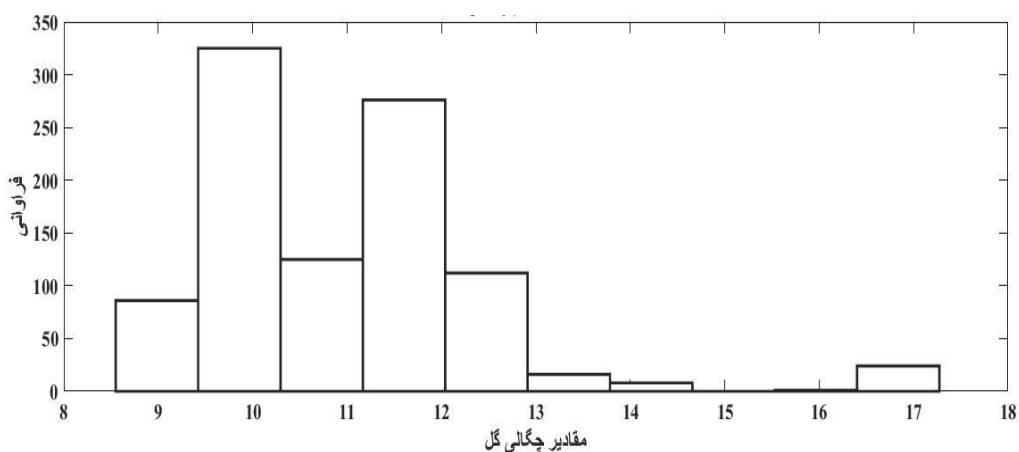
شکل ۴-۲ نمودار هیستوگرام پارامترهای حفاری مرتبط با مسیر چاه. الف- نمودار هیستوگرام زاویه انحراف مته بر حسب درجه، ب- نمودار هیستوگرام آزیموت مته بر حسب درجه



ج

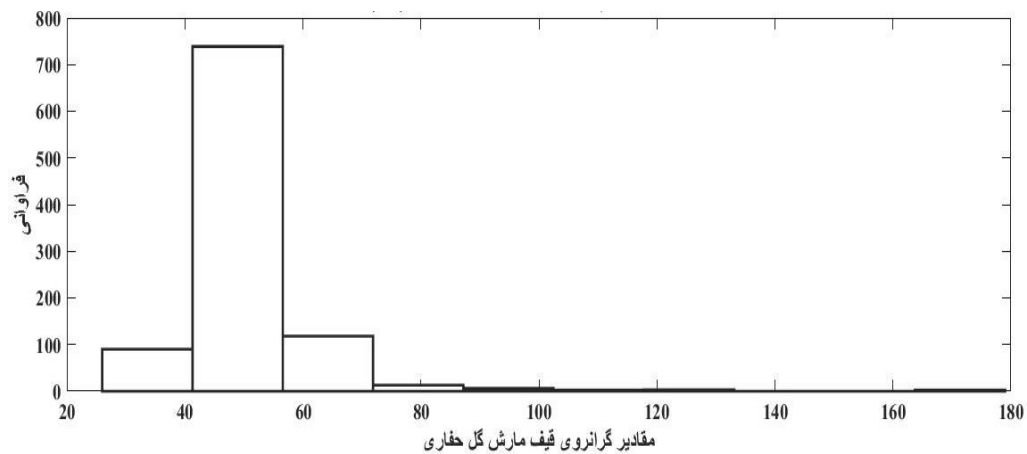
ادامه ی شکل ۲-۴ نمودار هیستوگرام پارامترهای حفاری مرتبط با مسیر چاه. ج- نمودار هیستوگرام میزان انحراف چاه از نقطه شروع بر حسب متر

نمودار هیستوگرام پارامترهای سیال حفاری در شکل ۳-۴ نمایش داده شده است. این شکل شامل نمودار پارامترهای چگالی گل، گرانروی قیف مارش، گرانروی پلاستیکی، نقطه ی واروی، استحکام ژلی اولیه، استحکام ژلی پس از ۱۰ دقیقه، میزان اسیدی یا بازی بودن، میزان نمک (کلر) موجود در گل و میزان یون کلسیم موجود در گل، می باشد.

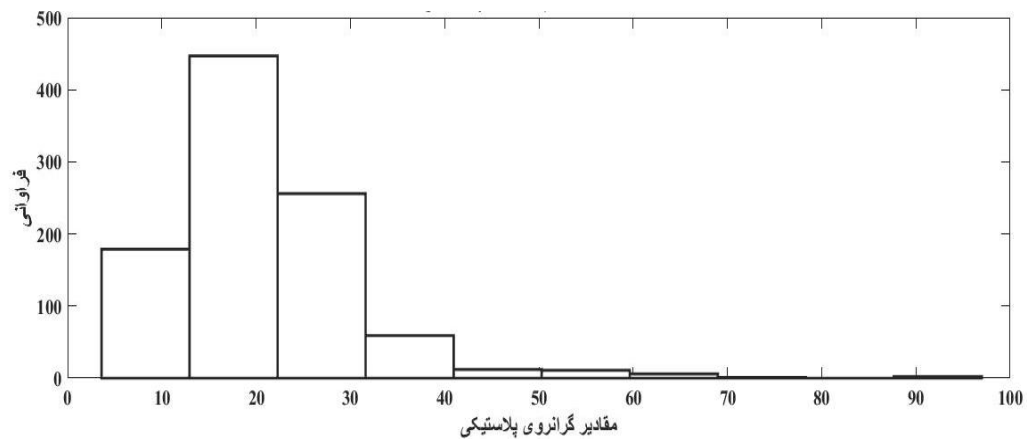


الف

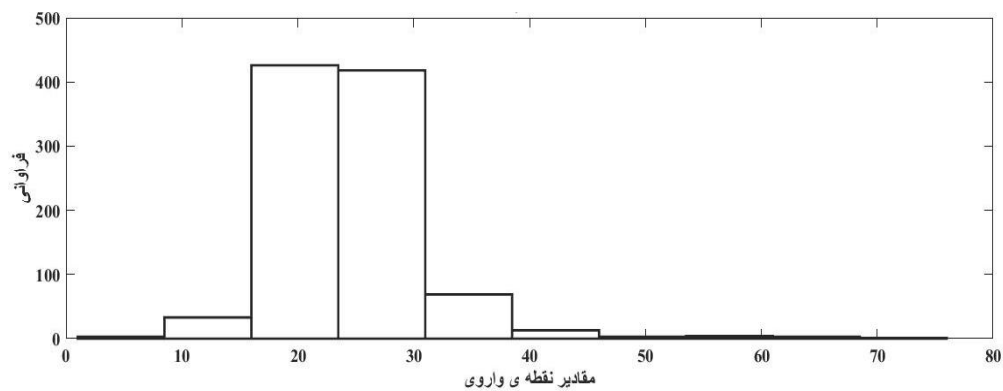
شکل ۳-۴ نمودار هیستوگرام پارامترهای سیال حفاری. الف- نمودار هیستوگرام چگالی گل بر حسب ppG



ب

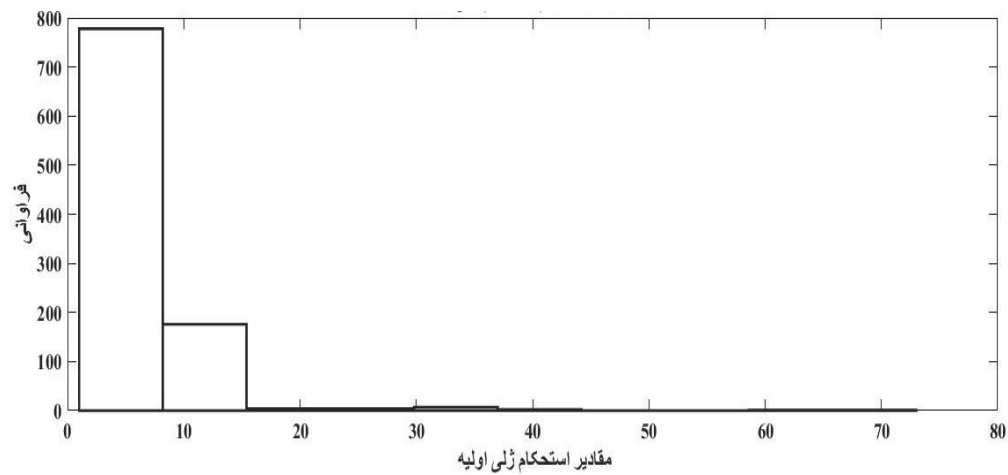


ج

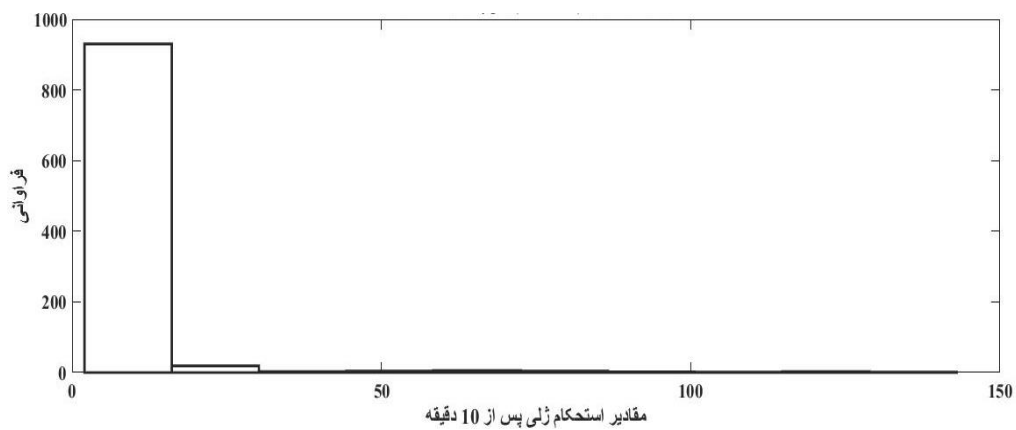


د

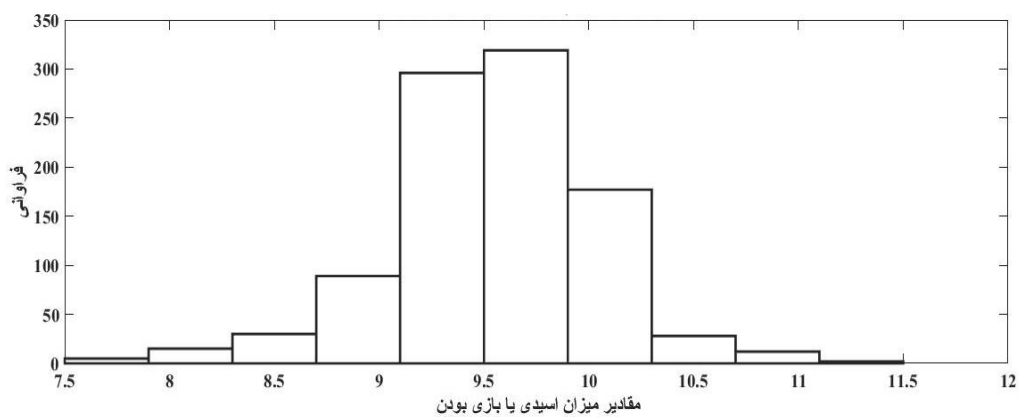
ادامه‌ی شکل ۳-۴- نمودار هیستوگرام پارامترهای سیال حفاری. ب- نمودار هیستوگرام گرانروی قیف مارش گل حفاری بر حسب ثانیه، ج- نمودار هیستوگرام گرانروی پلاستیکی بر حسب cp ، د- نمودار هیستوگرام نقطه واروی (پوند بر ۱۰۰ فوت مربع)



ه

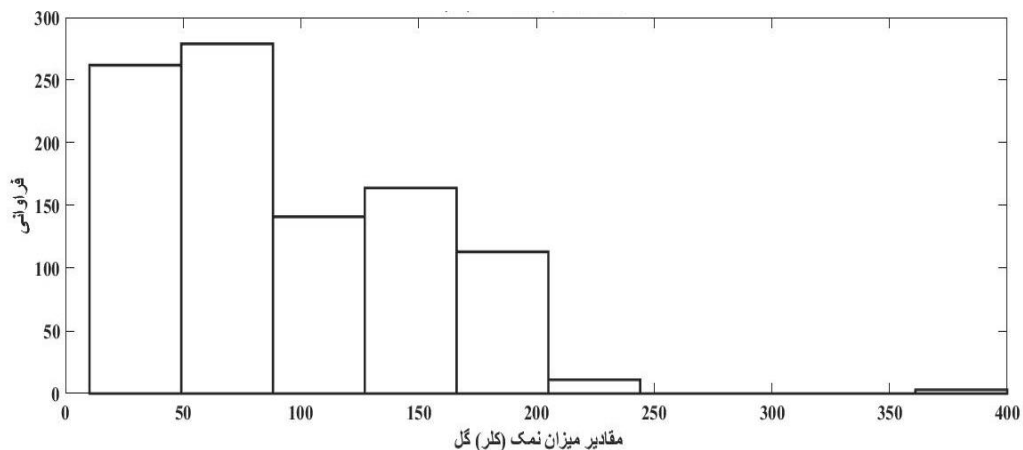


و

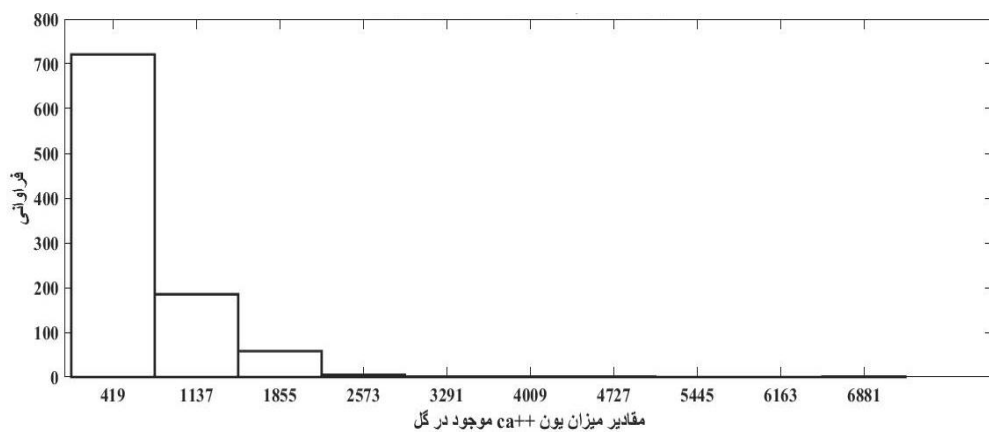


ز

ادامه‌ی شکل ۳-۴ نمودار هیستوگرام پارامترهای سیال حفاری. ه- نمودار هیستوگرام استحکام ژلی اولیه بر حسب پوند بر ۱۰۰ فوت مربع، و- نمودار هیستوگرام استحکام ژلی پس از ده دقیقه، ز- نمودار هیستوگرام میزان اسیدی یا بازی بودن گل.



ح



ط

ادامه‌ی شکل ۴-۳ نمودار هیستوگرام پارامترهای سیال حفاری. ح- نمودار هیستوگرام میزان نمک (کلر) گل بر حسب گرم بر لیتر ط- نمودار هیستوگرام میزان یون Ca^{2+} موجود در گل بر حسب ppm

در ابتدا از روش‌های دسته‌بندی جهت شناسایی پدیده گیر استفاده گردید. این کار علاوه بر این که می‌تواند الگوهای گیر و عدم گیر را تشخیص دهد، می‌تواند نشان‌دهنده‌ی صحت داده‌ها و پارامترهای ورودی باشد. اگر روش‌های دسته‌بندی بر اساس داده‌ها و پارامترهای ورودی قادر به شناسایی الگوها از یکدیگر باشند، می‌توان با اطمینان از آن داده‌ها جهت پیش‌بینی وقوع یا عدم وقوع گیر رشته‌ی حفاری، در مدل‌های پیش‌بینی کننده، استفاده نمود.

۳-۴ نتایج الگوریتم K نزدیک ترین همسایه

در این پژوهش به عنوان اولین روش، از الگوریتم K نزدیک ترین همسایه، جهت دسته بندی داده های گیر و عدم گیر استفاده گردید. اطلاعات ساختار این الگوریتم در جدول ۲-۴ آمده است. جهت این الگوریتم، تمامی مجموعه داده ها شامل ۱۲۶ مورد وقوع گیر و ۸۴۷ مورد عدم وقوع گیر، مورد استفاده قرار گرفت. هر مجموعه داده شامل ۲۰ پارامتر عملیات حفاری، مسیر چاه و سیال حفاری می باشد.

جدول ۲-۴ مشخصات و ساختار الگوریتم K نزدیک ترین همسایه

ویژگی	شرح
تعداد نزدیک ترین همسایه (K)	۳
نوع اعتبارسنجی	5-Fold cross validation
مقیاس فاصله	فاصله ی اقلیدسی
تعداد پیش بینی کننده	۲۰
تعداد دسته ها در خروجی	۲

نتایج حاصل از دسته بندی توسط الگوریتم K نزدیک ترین همسایه، به صورت ماتریس طبقه بندی^۱ در شکل ۴-۴ نمایش داده شده است. این الگوریتم در حالت مجموع ۹۲/۵ درصد در شناسایی دسته های گیر و عدم گیر موفق بود. این الگوریتم توانست از ۸۴۷ داده ی مرتبط با عدم وقوع گیر رشته حفاری، ۸۲۲ مورد را در

¹ Confusion matrix

دسته صحیح، طبقه‌بندی نماید. همچنین در مورد داده‌های وقوع گیر لوله، ۷۸ مورد از مجموع ۱۲۶ مجموعه داده‌ی وقوع گیر رشته حفاری، به‌درستی طبقه‌بندی گردید.

		طبقه خروجی مدل		
		عدم گیر	گیر	
طبقه خروجی واقعی	عدم گیر	822	25	97%
	گیر	48	78	62%

شکل ۴-۴ ماتریس طبقه‌بندی نتایج KNN

۴-۴ نتایج الگوریتم ماشین بردار پشتیبان

سپس از الگوریتم ماشین بردار پشتیبان جهت دسته‌بندی استفاده گردید. اطلاعات ساختار این الگوریتم در جدول ۳-۴ ارائه شده است. داده‌ها و پارامترهای ورودی این الگوریتم، مشابه الگوریتم KNN می‌باشد.

جدول ۳-۴ اطلاعات ساختار الگوریتم ماشین بردار پشتیبان

ویژگی	شرح
نوع اعتبارسنجی	5-Fold cross validation
تابع کرنل ^۱	Linear
تعداد پیش‌بینی کننده	۲۰
تعداد دسته‌ها در خروجی	۲ (گیر و عدم گیر)

نتایج حاصل از دسته‌بندی توسط این الگوریتم، به صورت ماتریس طبقه‌بندی در شکل ۴-۵ نمایش داده شده است. این الگوریتم در مجموع، ۹۶/۹ درصد در شناسایی دسته‌های گیر و عدم گیر رشته‌ی حفاری موفق بود. این الگوریتم توانست ۹۷ مورد از ۱۲۶ مجموعه داده‌ی مرتبط با گیر و ۸۳۵ مورد از ۸۴۷ مجموعه داده‌ی مربوط به عدم گیر رشته حفاری را به درستی دسته‌بندی نماید.

		طبقه خروجی مدل		
		عدم گیر	گیر	
طبقه خروجی واقعی	عدم گیر	835	12	98/6%
	گیر	29	97	77%

شکل ۴-۵ ماتریس طبقه‌بندی نتایج SVM

۴-۵ نتایج الگوریتم‌های دسته‌بندی به همراه تحلیل مؤلفه اصلی

جهت بهبود عملکرد این روش‌های دسته‌بندی از تحلیل مؤلفه‌ی اصلی (PCA) استفاده گردید. تعداد ۳ مؤلفه جهت ارتقاء این الگوریتم‌ها به کار رفت. نتایج حاصل از مدل ارتقاء یافته‌ی K نزدیک‌ترین همسایه با تحلیل مؤلفه‌ی اصلی، در شکل ۴-۶ نمایش داده شده است. الگوریتم ارتقاء یافته توانست ۹۳/۶ درصد دسته‌های گیر و عدم گیر را به‌درستی شناسایی نماید.

		طبقه خروجی مدل		
		عدم گیر	گیر	
طبقه خروجی واقعی	عدم گیر	827	20	97/6%
	گیر	42	84	66/7%

شکل ۴-۶ ماتریس طبقه‌بندی نتایج KNN ارتقاء یافته با PCA

نتایج حاصل از مدل ارتقاء یافته‌ی ماشین بردار پشتیبان با تحلیل مؤلفه‌ی اصلی، در شکل ۴-۷ نمایش داده شده است. این الگوریتم در مجموع، ۹۷ درصد در شناسایی دسته‌های گیر و عدم گیر رشته‌ی حفاری موفق بود.

		طبقه خروجی مدل		
		عدم گیر	گیر	
طبقه خروجی واقعی	عدم گیر	837	10	98/8%
	گیر	19	107	84/9%

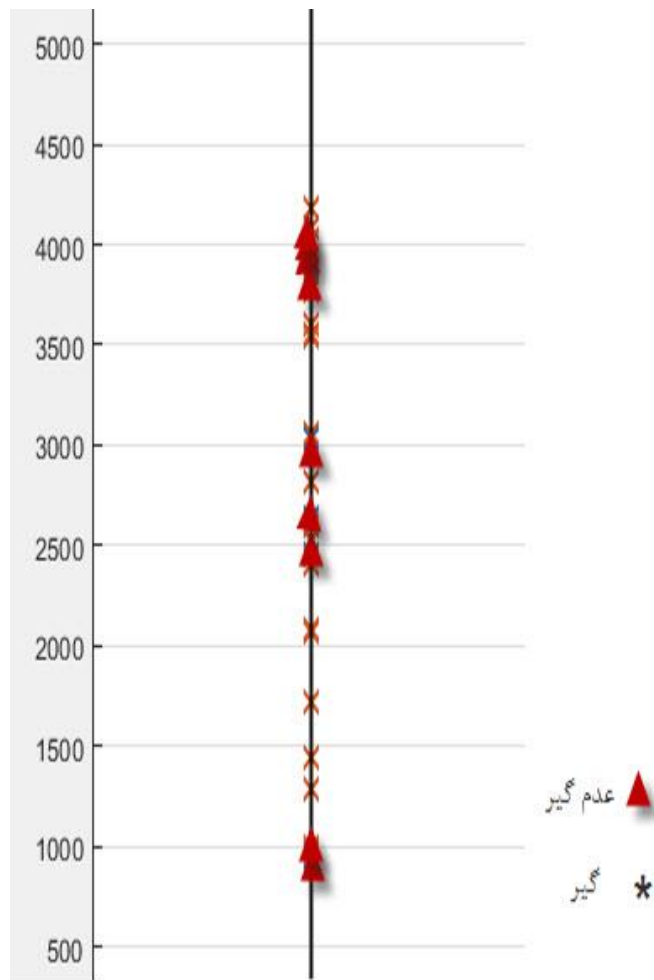
شکل ۴-۷ ماتریس طبقه‌بندی نتایج SVM ارتقاء یافته با PCA

به کارگیری تحلیل مؤلفه‌ی اصلی در الگوریتم‌های دسته‌بندی، بهبود عملکرد خصوصاً در دسته‌ی داده‌های گیر رشته حفاری، که برای ما از اهمیت بیشتری برخوردار است، را به همراه داشت. مقایسه‌ی بین نتایج روش ماشین بردار پشتیبان (SVM) و k نزدیک‌ترین همسایه (KNN)، در جدول ۴-۴ ارائه شده است. در این جدول، عملکرد الگوریتم‌ها در شناسایی دسته‌های گیر، عدم گیر و مجموع دو دسته بیان گردیده است.

جدول ۴-۴ مقایسه نتایج الگوریتم‌های KNN و SVM

الگوریتم	دسته گیر	دسته عدم گیر	مجموع دو دسته
KNN	۶۲٪	۹۷٪	۹۲/۵٪
KNN ارتقاء یافته با PCA	۶۶/۷٪	۹۷/۶٪	۹۳/۶٪
SVM	۷۷٪	۹۸/۶٪	۹۶/۹٪
SVM ارتقاء یافته با PCA	۸۴/۹٪	۹۸/۸٪	۹۷٪

الگوریتم بردار پشتیبان ارتقاء یافته با تحلیل مؤلفه‌ی اصلی، در شناسایی دسته‌های گیر و عدم گیر، بهترین عملکرد را داشت. عملکرد این الگوریتم در عمق‌های مختلف بررسی شد و مواردی که الگوریتم در شناسایی دسته صحیح ناموفق بود، بر حسب عمق چاه مشخص گردید. عمق داده‌هایی که الگوریتم در شناسایی آن‌ها ناموفق بود، در شکل ۴-۸ نشان داده شده است. در این شکل افزایش محور عمودی نشانه‌ی افزایش عمق است.



شکل ۴-۸ نمایش مواردی که الگوریتم SVM در دسته بندی آن‌ها ناموفق بود، در برابر افزایش عمق

بررسی لیتولوژی منطقه نشان دهنده این نکته بود که اعماقی که تراکم خطا در آن‌ها بیش‌تر است، غالباً دارای لیتولوژی سنگ آهک به‌همراه ناخالصی‌هایی از شیل و انیدریت می‌باشد.

نتایج روش‌های KNN و SVM و پس از آن نتایج حاصل از مدل‌های ارتقاء یافته‌ی این الگوریتم‌ها نشان داد که با داده‌ها و پارامترهای موجود، امکان شناسایی مطلوب الگوهای گیر و عدم‌گیر وجود دارد.

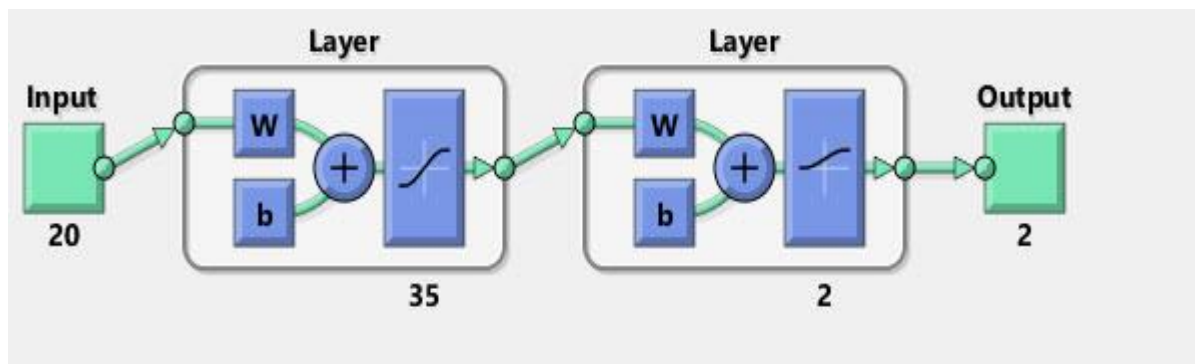
۴-۶ تشخیص الگو و پیش‌بینی به‌وسیله‌ی شبکه عصبی

برای شناسایی و پیش‌بینی پدیده‌ی گیر رشته‌ی حفاری از شبکه‌ی عصبی مصنوعی پرسپترون چند لایه (MLP) استفاده گردید. ساختار و ویژگی‌های این مدل در جدول ۴-۵ بیان گردیده است.

جدول ۴-۵ ساختار و ویژگی‌های مدل شبکه‌ی عصبی MLP

ویژگی	شرح
نوع شبکه عصبی	پرسپترون چند لایه (MLP)
تعداد لایه‌ی میانی	۱
تعداد نرون لایه‌ی میانی	۳۵
تابع الگوریتم آموزش	لونیگ-مارکوارت (trainlm)
درصد داده‌های آموزش	۵۵ درصد
درصد داده‌های آزمون	۳۰ درصد
درصد داده‌های اعتبارسنجی	۱۵ درصد
تابع تحریک لایه‌ی میانی	Tansig
تابع تحریک لایه‌ی خروجی	Logsig

مدل شبکه عصبی پرسپترون چند لایه در شکل ۴-۹ نمایش داده شده است.



شکل ۴-۹ مدل شبکه عصبی پرسپترون چند لایه

نتایج مدل شبکه عصبی به صورت ماتریس طبقه بندی، برای داده های آموزش در شکل ۴-۱۰ نشان داده شده است. برای این داده ها میانگین مربعات خطا برابر ۰/۰۲۵۵ بود. مدل شبکه عصبی توانست تمامی موارد عدم گیر را به درستی شناسایی نماید. همچنین در رابطه با داده های گیر، ۷۰ مورد از ۷۸ مجموعه داده ای این دسته، به طور صحیح شناسایی گردید.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	457	8	98/3%
	گیر	0	70	100%
		100%	89/7%	98/5%

شکل ۴-۱۰ ماتریس طبقه بندی نتایج مدل شبکه عصبی MLP برای داده های آموزش

همچنین نتایج مدل به صورت ماتریس طبقه‌بندی، برای داده‌های آزمون در شکل ۴-۱۱ نمایش داده شده است. برای این داده‌ها میانگین مربعات خطا برابر ۰/۰۶۵۸ بود. مدل شبکه عصبی توانست ۲۴۱ مورد از ۲۵۳ داده‌ی عدم گیر و ۳۰ مورد از ۳۹ داده‌ی وقوع گیر را به صورت صحیح شناسایی نماید.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	241	9	96/4%
	گیر	12	30	71/4%
		95/3%	76/9%	92/8%

شکل ۴-۱۱ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP برای داده‌های آزمون

۴-۶-۱ عملکرد شبکه عصبی با افزودن پارامتر لیتولوژی

علاوه بر عوامل عملیاتی و قابل تغییر مانند پارامترهای حفاری و گل، عوامل محیطی و پارامترهای مرتبط با سنگ و زمین نیز می‌توانند در وقوع گیر یا عدم وقوع گیر مؤثر باشند. در این پژوهش با توجه به خصوصیات سنگ‌ها، برای هر سنگ یک معادل عددی در نظر گرفته شد و با توجه به ترکیب و توالی سنگ‌ها، برای هر ترکیب سنگی عددی تعیین گردید. این اعداد به صورت ورودی مجزا علاوه بر ورودی‌های قبلی، به مدل وارد

گردید. معادل عددی هر سنگ در جدول ۴-۶ بیان شده است. این اعداد بر اساس ویژگی‌های سنگ، خصوصاً ویژگی‌های مخزنی مانند تخلخل و تراوایی، تعیین گردیده است.

جدول ۴-۶ معادل عددی تعیین شده برای هر سنگ

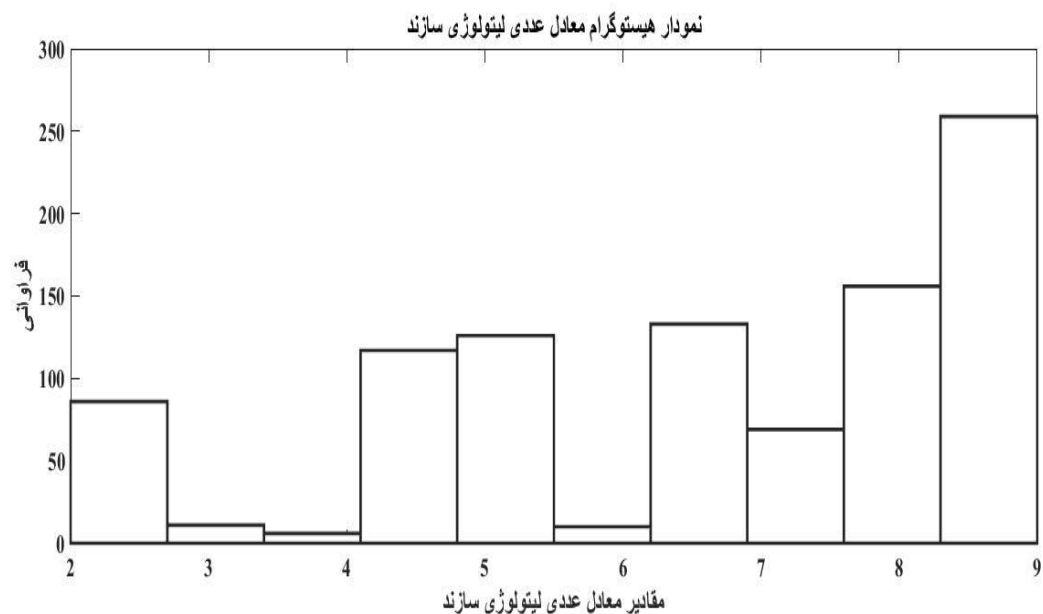
معادل عددی	نوع سنگ
۱	مارن
۲	انیدریت
۳	گل سنگ
۴	چرت
۵	ماسه سنگ
۶	شیل
۷	سنگ آهک رسی
۸	سنگ آهک
۹	دولومیت

معادل عددی برای ترکیبات سنگی مختلف در جدول ۴-۷ ارائه شده است.

جدول ۴-۷ معادل عددی برای ترکیبات سنگی مختلف

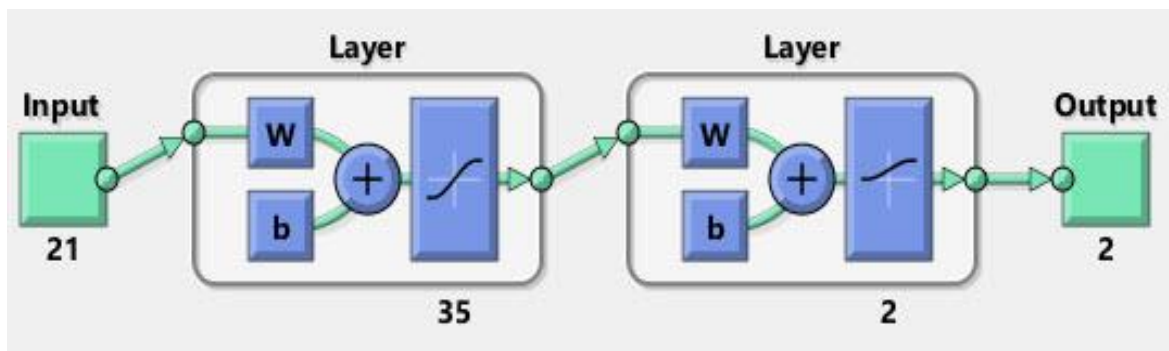
ترکیب سنگی	معادل عددی
انیدریت، مارن، سنگ آهک، دولومیت	۵
سنگ آهک و دولومیت	۸/۵
سنگ آهک	۸
شیل، گل سنگ، سنگ آهک	۵/۶۶
شیل، گل سنگ، ماسه سنگ	۴/۶۶
مارن و شیل	۳/۵
سنگ آهک رسی	۷
سنگ آهک، دولومیت، چرت	۷
انیدریت، دولومیت	۵/۵
دولومیت	۹
گل سنگ، سنگ آهک رسی	۵
سنگ آهک، دولومیت، گل سنگ	۶/۶۶
انیدریت	۲
انیدریت، گل سنگ	۲/۵
شیل، گل سنگ	۴/۵
گل سنگ	۳
دولومیت، گل سنگ، انیدریت	۴/۶۶
دولومیت، سنگ آهک، انیدریت	۶/۳۳

نمودار هیستوگرام معادل عددی لیتولوژی برای داده‌های این پژوهش در شکل ۴-۱۲ نشان داده شده است. حداقل، حداکثر و میانگین این پارامتر به ترتیب برابر با ۲، ۹ و ۶/۴۷ محاسبه گردید.



شکل ۴-۱۲ نمودار هیستوگرام معادل عددی لیتولوژی سازندهای میدان مورد مطالعه

این اطلاعات لیتولوژی برای هر کدام از مجموعه داده‌ها، بر اساس عمق تعیین گردید و به‌عنوان پارامتر مجزا به ورودی‌های مدل اضافه گردید. مدل با ورودی‌های جدید ساخته شد. مدل شبکه عصبی MLP ساخته شده با ۲۱ پارامتر در شکل ۴-۱۳ نشان داده شده است.



شکل ۴-۱۳ نمایش مدل شبکه عصبی MLP با افزودن پارامتر معادل عددی لیتولوژی

نتایج مدل مربوط به داده‌های آموزش در حالت افزودن پارامتر لیتولوژی به سایر پارامترها، به صورت ماتریس طبقه‌بندی در شکل ۴-۱۴ آمده است. میانگین مربعات خطا برای این مدل برای این داده‌ها، برابر با ۰/۰۰۸۸ بود. مدل شبکه عصبی در این حالت توانست تمامی موارد عدم گیر و ۶۹ مورد از ۷۱ مجموعه داده‌ی گیر را به صورت صحیح شناسایی کند.

طبقه خروجی واقعی			
طبقه خروجی مدل	گیر	عدم گیر	
	464	2	99/6%
	0	69	100%
	100%	97/2%	99/6%

شکل ۴-۱۴ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP با ۲۱ پارامتر ورودی برای داده‌های آموزش

نتایج مدل مربوط به داده‌های آزمون، در حالت افزودن پارامتر لیتولوژی به سایر پارامترها، به صورت ماتریس طبقه‌بندی در شکل ۴-۱۵ آمده است. میانگین مربعات خطا برای این مدل در این دسته از داده‌ها، برابر با ۰/۰۴۶۱ بود. مدل شبکه عصبی توانست در این حالت، ۲۴۳ مورد از ۲۵۰ مجموعه داده‌ی عدم گیر و ۳۴ مورد از ۴۲ مجموعه داده‌ی گیر را به درستی شناسایی کند.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	243	8	96/8%
	گیر	7	34	82/9%
		97/2%	81%	94/9%

شکل ۴-۱۵ ماتریس طبقه‌بندی نتایج مدل شبکه عصبی MLP با ۲۱ پارامتر ورودی برای داده‌های آزمون

به کارگیری پارامتر لیتولوژی به عنوان یکی از پارامترهای محیطی، سبب پیش‌بینی و دسته‌بندی بهتر داده‌های گیر و عدم گیر توسط مدل شبکه عصبی پرسپترون چند لایه گردید. به دلیل وجود مجموعه داده‌های بسیار در این پژوهش، ۵۵ درصد کل داده‌ها جهت آموزش شبکه کافی به نظر رسید. در نتیجه ۳۰ درصد داده‌ها جهت آزمون شبکه به کار رفت، تا از عملکرد مدل و قابلیت تعمیم آن اطمینان حاصل گردد.

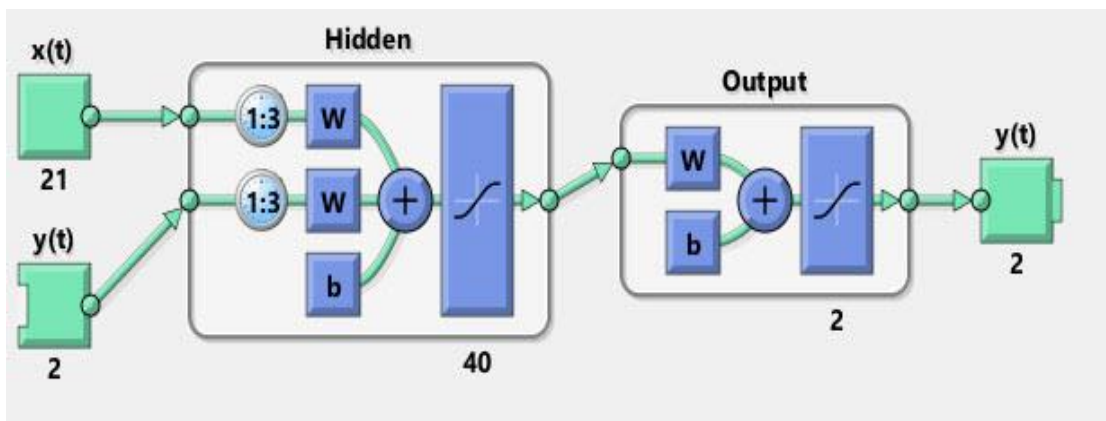
۷-۴ پیش‌بینی به‌وسیله‌ی سری زمانی

جهت پیش‌بینی پدیده‌ی گیر لوله به‌صورت دینامیکی، از مدل سری‌های زمانی تحت شبکه عصبی استفاده گردید. مشخصات و ساختار مدل سری زمانی به‌کار رفته در این پژوهش، در جدول ۸-۴ ارائه گردیده است.

جدول ۸-۴ مشخصات مدل سری زمانی

ویژگی	شرح
تعداد لایه‌ی میانی	۱
تعداد نرون در لایه‌ی میانی	۴۰
تأخیر زمانی	۳
تابع تحریک لایه‌ی میانی	Tansig
تابع تحریک لایه‌ی خروجی	Tansig

مدل سری زمانی ساخته شده در شکل ۱۶-۴ نمایش داده شده است.



شکل ۴-۱۶ نمایش مدل سری زمانی ساخته شده

نتایج خروجی مدل سری زمانی به صورت ماتریس طبقه بندی برای داده های آموزش در شکل ۴-۱۷ نشان داده شده است. میانگین مربعات خطا برای این داده ها ۰/۰۳۱۸ به دست آمد. این مدل توانست ۴۵۱ مورد از ۴۶۳ مجموعه ی داده ی عدم گیر و ۵۹ مورد از ۷۱ مجموعه داده ی گیر را به درستی تشخیص دهد.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	451	12	97/4%
	گیر	12	59	83/1%
		97/4%	83/1%	95/5%

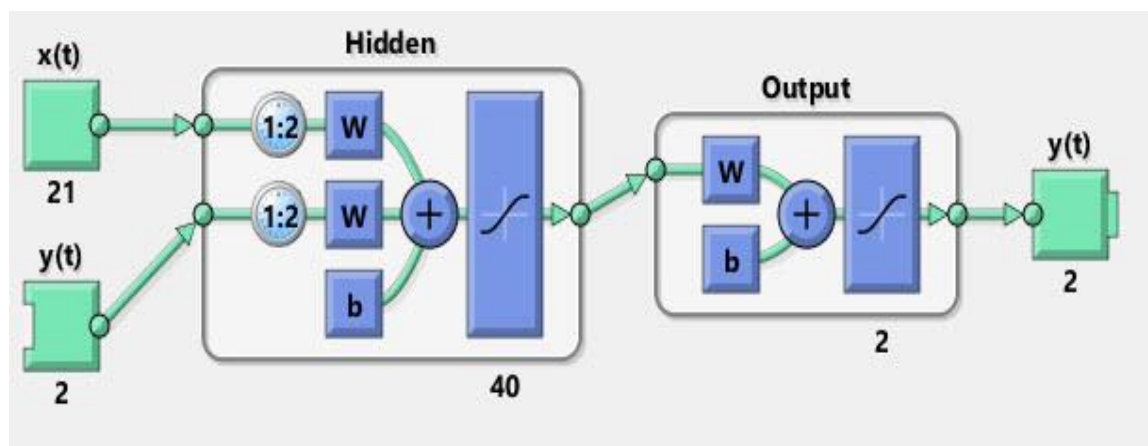
شکل ۴-۱۷ ماتریس طبقه بندی نتایج مدل سری زمانی برای داده های آموزش

نتایج خروجی مدل سری زمانی به صورت ماتریس طبقه‌بندی برای داده‌های آزمون در شکل ۴-۱۸ نشان داده شده است. میانگین مربعات خطا برای این داده‌ها ۰/۰۴۸۱ به دست آمد. این مدل توانست ۲۴۴ مورد از ۲۵۲ مجموعه‌ی داده‌ی عدم گیر و ۲۹ مورد از ۳۹ مجموعه داده‌ی گیر را به درستی تشخیص دهد.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	244	10	96/1%
	گیر	8	29	78/4%
		96/8%	74/4%	93/8%

شکل ۴-۱۸ ماتریس طبقه‌بندی نتایج مدل سری زمانی برای داده‌های آزمون

جهت بهبود عملکرد مدل سری زمانی، تغییراتی در مدل اعمال گردید تا نتیجه بهینه به دست آید. تغییر میزان تأخیر زمانی از ۳ به ۲ روز، بهبود عملکرد مدل سری زمانی را در پی داشت. در این حالت علاوه بر کاهش زمان اجرای مدل در برنامه، بهبود در نتایج نیز حاصل گردید. تأخیر زمانی ۲ به عنوان میزان بهینه‌ی این عامل تعیین گردید. مدل سری زمانی جدید در شکل ۴-۱۹ نشان داده شده است.



شکل ۴-۱۹ نمایش مدل سری زمانی با تأخیر زمانی ۲

نتایج خروجی مدل سری زمانی با تأخیر زمانی ۲، به صورت ماتریس طبقه‌بندی برای داده‌های آموزش، در شکل ۴-۲۰ نشان داده شده است. میانگین مربعات خطا برای این داده‌ها ۰/۰۳۷۲ به دست آمد. این مدل توانست، ۴۴۹ مورد از ۴۵۸ مجموعه داده‌ی عدم گیر و ۶۵ مورد از ۷۶ مجموعه داده‌ی وقوع گیر را به صورت صحیح تشخیص دهد.

طبقه خروجی واقعی			
طبقه خروجی مدل		عدم گیر	گیر
	عدم گیر	449	11
	گیر	9	65
		98%	85/5%
		97/6%	87/8%
		96/3%	

شکل ۴-۲۰ ماتریس طبقه‌بندی نتایج مدل سری زمانی با تأخیر زمانی ۲، برای داده‌های آموزش

نتایج خروجی مدل سری زمانی با تأخیر زمانی ۲، به صورت ماتریس طبقه‌بندی برای داده‌های آزمون در شکل ۴-۲۱ نشان داده شده است. میانگین مربعات خطا برای این داده‌ها ۰/۰۵۵۱ به دست آمد. این مدل توانست، ۲۳۷ مورد از ۲۴۳ مجموعه داده‌ی عدم گیر و ۳۷ مورد از ۴۸ مجموعه داده‌ی وقوع گیر را به درستی شناسایی کند.

		طبقه خروجی واقعی		
		عدم گیر	گیر	
طبقه خروجی مدل	عدم گیر	237	11	95/6%
	گیر	6	37	86%
		97/5%	77/1%	94/2%

شکل ۴-۲۱ ماتریس طبقه‌بندی نتایج مدل سری زمانی با تأخیر زمانی ۲، برای داده‌های آزمون

نتایج مدل‌های شبکه عصبی پرسپترون چند لایه و سری زمانی در جدول ۴-۹ ارائه شده است. ۲۱ ورودی برای مدل‌های شبکه عصبی MLP و سری زمانی، شامل ۲۰ پارامتر عملیات و سیال حفاری و یک پارامتر معادل عددی لیتولوژی، می‌باشد. مدل سری زمانی در دو حالت با ۲۱ ورودی ساخته شد.

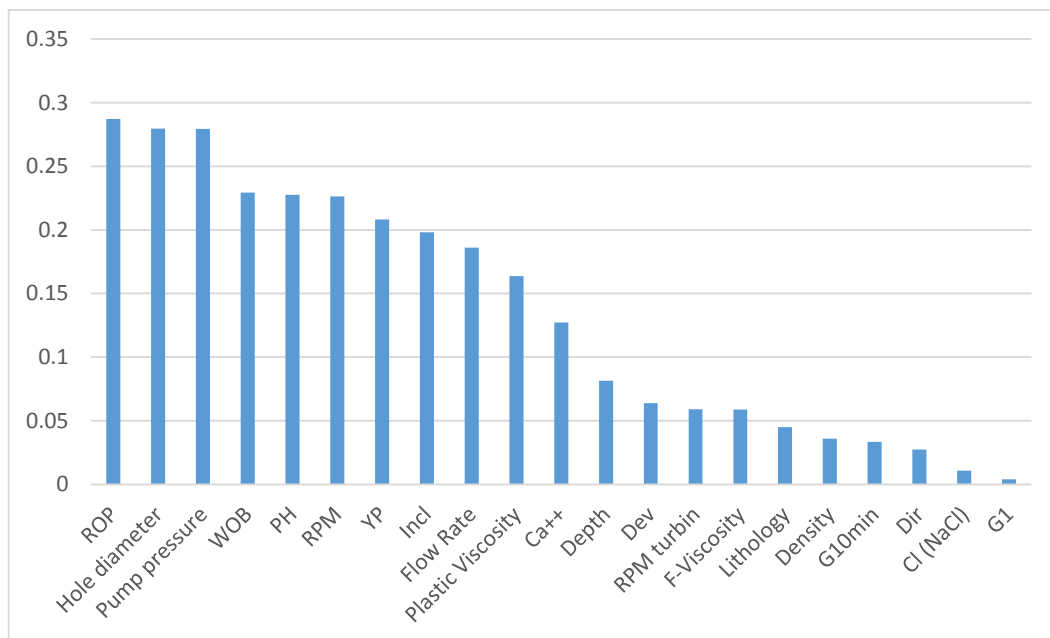
جدول ۴-۹ مقایسه نتایج مدل‌های شبکه عصبی MLP و سری زمانی

نتایج برای داده‌های آموزش		نتایج برای داده‌های آزمون		مدل
دسته گیر	دسته عدم گیر	دسته گیر	دسته عدم گیر	
۸۹/۷٪	۱۰۰٪	۷۶/۹٪	۹۵/۳٪	شبکه عصبی MLP با ۲۰ ورودی
۹۷/۳٪	۱۰۰٪	۸۱٪	۹۷/۲٪	شبکه عصبی MLP با ۲۱ ورودی
۸۳/۱٪	۹۷/۴٪	۷۴/۴٪	۹۶/۸٪	سری زمانی با تأخیر زمانی ۳
۸۵/۵٪	۹۸٪	۷۷/۱٪	۹۷/۵٪	سری زمانی با تأخیر زمانی ۲

مقایسه‌ی نتایج مدل‌های پیش‌بینی کننده، بیان‌گر عملکرد مطلوب هر دو مدل شبکه عصبی MLP و سری زمانی، در شناسایی و پیش‌بینی وقوع یا عدم وقوع گیر می‌باشد. تفاوت نسبی در دقت شناسایی دسته‌های گیر و عدم گیر رشته حفاری، ناشی از عدم تعادل در تعداد داده‌های ورودی مرتبط با این دو دسته می‌باشد. در این پژوهش به دلیل خودداری از انتخاب یا حذف داده‌های ورودی (جهت حفظ اعتبار و قابلیت تعمیم مدل)، تعادل بین داده‌های ورودی مرتبط با وقوع گیر رشته حفاری و داده‌های مرتبط با عدم وقوع این پدیده، برقرار نگردید.

۴-۸ بررسی پارامترهای ورودی به مدل

پارامترهای ورودی به مدل، بر اساس میزان همبستگی با خروجی (وقوع یا عدم وقوع گیر رشته حفاری)، در شکل ۴-۲۲ نمایش داده شده‌اند. در این شکل بزرگی مقادیر همبستگی پارامترهای مختلف، بدون در نظر گرفتن مثبت (مستقیم) یا منفی (غیرمستقیم) بودن اثر آن‌ها، نشان داده شده است.



شکل ۴-۲۲ نمایش میزان همبستگی پارامترهای ورودی با خروجی

شناسایی و پیش‌بینی پدیده‌ی گیر رشته حفاری می‌تواند به جلوگیری از وقوع آن کمک نماید. پرسنل حفاری با اطلاع از احتمال وقوع گیر، قادر خواهند بود تا از طریق تغییر پارامترهای غیرمحیطی و پارامترهای قابل تغییر، از وقوع آن جلوگیری نمایند.

فصل پنجم:

نتیجه‌گیری و پیشنهادات

۵-۱ نتیجه‌گیری

۵-۲ پیشنهادات

۵-۱ نتیجه گیری

- در این پژوهش از الگوریتم‌های دسته‌بندی مانند ماشین بردار پشتیبان و K نزدیک‌ترین همسایه، جهت شناسایی دسته‌های گیر و عدم‌گیر استفاده گردید که نتایج خروجی مدل بیان‌گر عملکرد مطلوب این الگوریتم‌ها در حل چنین مسائلی است. در حالت مجموع، الگوریتم ماشین بردار پشتیبان ۹۶/۹ درصد و الگوریتم K نزدیک‌ترین همسایه ۹۲/۵ درصد، در شناسایی صحیح دسته-های گیر و عدم‌گیر موفق بودند.
- به‌کارگیری تحلیل مؤلفه‌های اصلی در کنار الگوریتم‌های ماشین بردار پشتیبان و K نزدیک‌ترین همسایه، سبب بهبود عملکرد این الگوریتم‌ها در دسته‌بندی مجموعه داده‌ها، خصوصاً داده‌های گیر رشته حفاری گردید. شناسایی این دسته از داده‌ها که برای ما از اهمیت بیش‌تری برخوردارند، در این حالت از ۶۱/۹ درصد به ۶۶/۷ درصد برای الگوریتم KNN و از ۷۷ درصد به ۸۴/۹ درصد برای الگوریتم SVM، بهبود یافت.
- در مدل شبکه عصبی، علاوه بر داده‌های مربوط به پارامترهای حفاری و سیال حفاری، داده‌های مربوط به لیتولوژی نیز به عنوان یک ورودی وارد گردید. افزودن این ورودی سبب بهبود عملکرد مدل شبکه عصبی گردید. در این حالت عملکرد شبکه، برای داده‌های آزمون از ۹۲/۸ درصد به ۹۴/۹ درصد برای مجموع دو دسته، بهبود یافت. می‌توان نتیجه گرفت که عوامل محیطی و غیرقابل تغییر نیز در عملکرد مدل مؤثرند، لذا به کارگیری پارامترهایی از این دست به ما در کسب نتایج معتبرتر کمک می‌نماید.
- در این پژوهش از مدلی که قادر باشد رفتار دینامیکی شرایط وقوع گیر و عدم‌گیر را پیش‌بینی نماید، استفاده گردید. سری زمانی یکی از مدل‌های دینامیکی است که می‌تواند وقوع یا عدم وقوع

گیر رشته‌ی حفاری را با توجه به شرایط و عملکرد روزهای قبل حفاری، پیش‌بینی نماید. در این پژوهش از سری‌های زمانی تحت شبکه عصبی استفاده گردید. این مدل توانست داده‌های آموزش را ۹۶/۳ درصد و داده‌های آزمون را ۹۴/۲ درصد، به‌درستی شناسایی نماید.

۵-۲ پیشنهادات

- به‌کارگیری پارامتر لیتولوژی که از عوامل محیطی و مرتبط با خصوصیات زمین است، سبب بهبود نتایج در این پژوهش گردید، لذا پیشنهاد می‌گردد تا حد امکان از پارامترهای سنگ، به‌عنوان ورودی مدل در کنار سایر پارامترها استفاده شود. مقاومت تراکمی سنگ می‌تواند یکی از این پارامترها باشد.
- با توجه به اثری که ورود پارامتر لیتولوژی بر بهبود عملکرد شبکه عصبی داشت، پیشنهاد می‌گردد بر روی این پارامتر تحلیل حساسیت انجام گیرد.
- بررسی داده‌های این میدان بیان‌گر وقوع گیر عمدتاً در برابر لایه‌های تراوا و لایه‌های مخزنی بود. پیشنهاد می‌شود پدیده‌ی گیر لوله و شرایط وقوع آن، در چاه‌های افقی که سطح تماس چاه و رشته‌ی حفاری با لایه‌ی مخزنی بسیار زیاد است، بررسی و مدل‌سازی گردد.
- در مدل سری زمانی، درصدی از میزان خطا ناشی از نوع ورود داده‌ها به این مدل بود. در این پژوهش به دلیل مقایسه‌ی روش‌های مختلف، شرایط یکسانی برای ورودی‌ها در نظر گرفته شد. با توجه به نوع عملکرد سری‌های زمانی در پیش‌بینی، که استفاده از تأخیر زمانی می‌باشد، ترتیب ورود داده‌ها بر اساس تاریخ، مهم و مؤثر است. از سوی دیگر در حد فاصل بین ورود اطلاعات دو چاه مختلف، امکان بروز خطا وجود دارد. زیرا مدل جهت تأخیر زمانی برای ورودی‌های اول هر چاه،

از اطلاعات پایانی چاه قبل استفاده می‌کند که منجر به بالارفتن خطای مدل می‌گردد. در صورت عدم قیاس بین مدل‌های مختلف و تمایل جهت استفاده از این مدل به تنهایی، پیشنهاد می‌گردد که در مطالعات کاربردی نوع ورود داده به صورت چاه‌های جداگانه باشد و درصدی از چاه‌ها به عنوان داده‌های آموزش، اعتبارسنجی و آزمون مدل، در نظر گرفته شوند.

- پیشنهاد می‌شود جهت بهبود عملکرد مدل‌های پیش‌بینی، تعادل بین تعداد داده‌های مرتبط با هر دسته برقرار شود.

منابع

- [1] MI L. L. C., (1998), "**Drilling Fluid Engineering Manual**". MI Drilling Fluids Company.
- [2] Jardine, S. I., McCann, D. P., Barber, S. S., (1992), "An Advanced System for the Early Detection of Sticking Pipe," paper SPE 23915 presented at SPE Drilling Conference, New Orleans.
- [۳] شوراب، ح.، (۱۳۷۴)، "مانده یابی"، گزارش شماره ۴۶۱۵، انتشارات اداره کل حفاری.
- [4] Miri, R., Sampaio, J., Afshar, M. and Lourenco, A., (2007), "Development of Artificial Neural Networks to Predict Differential Pipe Sticking in Iranian Offshore Oil Fields", Paper SPE 138533 presented at the 2337 International oil conference and Exhibition in Mexico held in Veracruz, Mexico.
- [5] Bowes, C., Procter, R., (1997), "**Drillers Stuck pipe Handbook**", Procter & Collins Ltd.
- [6] Baker Hughes., (1995), "**Drilling Engineering Workbook**", Baker Hughes INTEQ.
- [7] Anadrill, (1997), "**Stuck-Pipe Prevention**", Self Learning Course. Anadrill.
- [8] Rabia, H., (1985), "**Oilwell Drilling Engineering: Principles and Practice**", Graham & Trotman, USA.
- [۹] عادل زاده، م. ر.، (۱۳۸۵)، "اصول مهندسی حفاری"، انتشارات ستایش. ۸۶۴ صفحه.
- [۱۰] مطوریان، ر.، ملایری، م. ر.، (۱۳۸۸)، "سیالات حفاری"، شرکت نفت فلات قاره ایران، روابط عمومی. ۲۱۶ صفحه.
- [11] Hemphkins, W. B., Kingsborough, R. H., Lohec, W. E., Nini, C. J., (1987), "Multivariate Statistical Analysis of Stuck drill pipe situations", SPE drilling engineering.
- [12] Kessler, C., Dorfffer, D., Crawford, D., Dehart, R. and Weiser, J., (2009), "Case Histories of a New Wireline Logging Tool for Determination of Free Point in Support of Drilling and

Pipe Recovery Operations”, SPE/IADC 119907, presentation at the SPE/IADC Drilling Conference and Exhibition held in Amsterdam, The Netherlands.

[13] Mohanna, A., Voghell, M., Alali, A., Bakir, R. and Al-Mousa, A., (2012), “A Step Change in Fishing Efficiency: Recovering Stuck Pipe Using the Fishing Agitation Tool”, SPE 160880, presentation at the SPE Saudi Arabia Section Technical Symposium and Exhibition held in Al-Khobar, Saudi Arabia.

[14] Biegler, m. w., Kuhn, G. R., (1994), “Advances in prediction of stuck pipe using multivariate statistical analysis”, Presented at the SPE/IADC conference in Dallas.

[15] Howard, J. A., Glover, S. S., (1994), “Tracking Stuck Probability While drilling”, Paper SPE 27528 Presented for presentation at the 1994 IADC/SPE Drilling Conference held in dallas, Texas.

[16] Siruvuri, C., Nagarakanti, S., Samuel, R., (2006), “Stuck Pipe Prediction and Avoidance: A Convolutional Neural Network Approach”, Paper SPE48378 Presented at the IADC/SPE Drilling.

[17] Halliday, W. S., Clapper, D. K., (1989), “Toxicity and performance testing of non-oil spotting fluid for differential stuck pipe”, SPE/IADC Drilling Conference. New Orleans, Louisiana, Paper SPE 18684, pp. 521-528.

[18] Reid, P. Santos, H., (2003), “Novel Drilling, Completion and Workover Fluids for Depleted Zones: Avoiding Losses, Formation Damage and Stuck Pipe”, SPE/IADC Middle East Drilling Technology Conference & Exhibition. Abu Dhabi, UAE, Paper IADC/SPE 85326, pp. 1-9.

[19] Outmans, H, D., (1958), “Mechanics of Differential Pressure Sticking of Drill Collars”, Trans, AIME 213,265.

[20] Reid, P. I., Meeten, G. H., Way, P. W., Clark, P., Chambers, B. D., Gilmour, A., Sandres, M. W., (2000) , **“Differential Sticking, mechanisms and a simple wellsite test for Monitoring and Optimizing Drilling Mud Properties”**.

- [21] Murillo, A., Neuman, J., Samuel, R., (2009), “Pipe sticking prediction and avoidance using adaptive fuzzy logic and neural network modeling”. SPE Production and Operation Symposium, Oklahoma, USA, Paper SPE 120128, pp. 1-15.
- [22] Ali, J. K., (1994), “Neural Networks: A new tool for the petroleum industry”, paper SPE 27561 presented at the European Petroleum Computer Conference held in Aberdeen, UK.
- [23] Batyrshin, I., Sheremetov, L., Markov, M., Panova, A., (2005), “Hybrid method for porosity classification in carbonate formations”, Journal of Petroleum Science and engineering, 35-50.
- [24] Fruhwirth, R. K., Thonhauser, G., Mathins, W., (2006), “hybrid simulation using neural network to predict drilling hydraulics in real time”, SPE Annual technical conference and exhibition. San Antonio, Texas, USA, Paper SPE 103217, pp. 1-8.
- [25] Malallah, A., Nashawi, I. S., (2005), “Estimating the fracture gradient coefficient using neural networks for a field in the Middle East”. Journal of petroleum science and engineering, 193-211.
- [26] Hou, J., Li, Zh. Q., Cao, X. I., Song, X. W., (2009), “integrating genetic algorithm and support vector machine for polymer flooding production performance prediction”. Journal of petroleum science and engineering, 29-39.
- [27] El-Sebakhy, A. A., (2009), “Forecasting pvt properties of crude oil systems based on support vector machines modeling scheme” Journal of petroleum science and engineering, 25-34.
- [28] Gonzalez, O., Bernat, H. and Moore, P., (2007) “The Extraction of Mud-Stuck Tubulars Using Vibratory Resonant Techniques”, Paper SPE 109530 presented at the SPE Annual Technical Conference and Exhibition, Anaheim, California, U. S. A.
- [29] Agarwal, S. and Agarwal, N., (2008), “Auto Release Drill Collars”, Paper SPE 112348 presented at the Indian Oil and Gas Technical Conference and Exhibition, Mumbai, India.

[30] Shadizadeh, S. R., Karimi, F. and Zoveidavianpoor, M., (2010), "Drilling Stuck Pipe Prediction in Iranian Oil Fields: An Artificial Neural Network Approach", Iranian Journal of Chemical Engineering Vol. 7, No. 4, IACHE.

[31] Dipl. I., Prassl. W. F., "**Drilling Engineering**", Curtin University of Technology.

[32] Tobenna, C. U., (2010), "**Hole Cleaning and Hydraulics**", University of Stavanger, Norway.

[33] Amoco Production Company, "**Drilling Fluids Manual**".

[۳۴] درویش‌زاده، ع.، (۱۳۸۹) "زمین‌شناسی ایران؛ چینه‌شناسی، تکتونیک، دگرگونی و ماگماتیسم"، مؤسسه انتشارات امیرکبیر. تهران. ۴۳۴ صفحه.

[35] Duda, R. O., Hart, P. E., Stork, D. G., "**Pattern Classification**", Second Edition, Wiley Inc.

[36] Webb, A. R., (2002), "**Statistical pattern recognition**", Second edition, John Wiley & Sons, Ltd.

[37] Weston, J., "Support Vector Machine and Statistical learning theory Tutorial", NEC Labs America.

[38] Murillo, A., Neuman, J., Samuel, R., (2009), "Pipe sticking prediction and avoidance using adaptive fuzzy logic and neural network modeling", SPE Production and Operation Symposium, Oklahoma, USA, Paper SPE 120128, pp. 1-15.

[39] Cortes, C., Vapnik, V., (1995), "Support-vector networks". Machine Learning 20, 273-297.

[40] Hagan, M.T., Demuth, H.B., Beale, M., (1996), "**Neural Network Design**". PWS Publishing Company.

[41] Qiu, J., Wang, H., Lu, J., Zhang, B. and Du, K. L., (2012) "Neural Network Implementations for PCA and Its Extensions". International Scholarly Research Network.

[42] Hyvarinen, A., Karhunen, J., Oja, E., "**Independent Component Analysis**", John Wiley & Sons, Inc.

[۴۳] کمالی، م. ر.، علی مددی، ف.، فخری، ا.، (۱۳۸۹)، "کاربرد روش‌های هوشمند در مهندسی نفت

و علوم زمین"، انتشارات پژوهشگاه صنعت نفت. تهران. ۱۲۶ صفحه.

[۴۴] صادق‌زاده اهری، ج.، سعدی، م.، بهجت، ی.، (۱۳۹۰)، "کاربرد نرم‌افزار MATLAB در مهندسی"،

انتشارات پژوهشگاه صنعت نفت. تهران. ۳۰۰ صفحه.

[45] Haykin, S., (1999), "Neural Networks, A Comprehensive Foundation", Pearson Education Inc.

[۴۶] نیرومند، ح. ع.، بزرگ‌نیا، ا.، (۱۳۸۴)، "سری‌های زمانی"، انتشارات دانشگاه پیام نور. تهران. ۲۷۶

صفحه.

Abstract

Pipe sticking is a conventional problem in drilling industry. This matter can cause several problems from time and cost lost, to losing the hole. In this study, tried to making a model to predict pipe sticking. For this task, drilling and fluid parameters of several wells of a giant field have used. 126 case of 973 data sets, that each one is included of drilling and fluid parameters, are related to stuck pipe and others are related to non-stuck pipe observations. In this thesis, at first classification algorithms such as K nearest neighbor and support vector machines, have used to classification of data. KNN 93.6 percent and SVM 97 percent were successful to this task. Then multi layer perceptron neural network and neural network time series have used to predict pipe sticking. For first time ever, time series as a dynamic method has used to predict of pipe sticking. Prediction accuracy was 94.9 percent for MLP neural network and 94.2 percent for time series. Being informed to pipe sticking probability could help to prevent it by changing effective parameters.

Keywords: Drilling pipe sticking, multi layer perceptron neural network, time series, K nearest neighbor algorithm, support vector machine.



Shahrood University of Technology
Faculty of Mining, Petroleum and Geophysics Engineering

MSc Thesis of Petroleum Engineering, Drilling and Production

**Prediction and Prevention of Drilling Pipe Sticking by Artificial Neural Network and
Time Series**

By

Sajad Eskandarian

Supervisors:

Dr. Behzad Tokhmechi

Dr. Vamegh Rasouli

June 2016