



Research Paper

## Real-time Drilling Event Detection Based on Mud-logging data Using Long Short Term Memory Neural Networks

Mehran Badinloo<sup>1</sup>, Jamal Farashiani<sup>1</sup>, Davood Khoozan<sup>1,\*</sup>, Mehdi Mansoori<sup>2</sup>, Vahid Daneshkhah<sup>2</sup><sup>1</sup> Faculty of Chemical Engineering, Tarbiat Modares University, Tehran, Iran.<sup>2</sup> Dana Energy Company, Tehran, Iran.

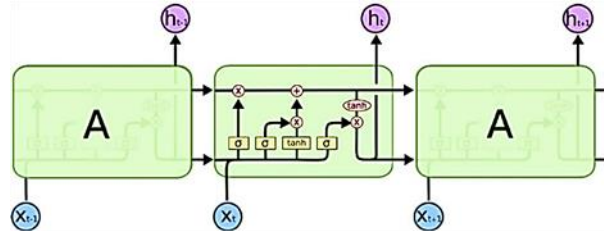
## ARTICLE INFO

Received 2024-08-12  
 Accepted 2024-11-19  
 Available online 2025-03-15  
 ISSN: 2588-5316  
 Online ISSN: 2588-5324

## Keywords:

Mud logging  
 Long-short term memory  
 Kick  
 Stuck pipe  
 SMOTE

## GRAPHICAL ABSTRACT



## ABSTRACT

**Research Subject:** Drilling operations frequently encounter numerous challenges that can lead to significant financial, human, and environmental losses. Therefore, predicting potential problems before they occur and implementing necessary preventive measures is crucial to minimizing risks. In this context, this study investigates the impact of employing artificial intelligence (AI) algorithms to forecast drilling complications using real-time mud logging data collected from existing wells in an Iranian oilfield.

**Research Approach:** A hybrid architecture combining Long Short-Term Memory (LSTM) and Fully Connected neural networks was developed for the identification and detection of anomalies such as kicks and stuck pipe. Given the scarcity of these anomalies in the dataset, which could adversely affect model accuracy and performance, the Synthetic Minority Oversampling Technique (SMOTE) was applied to balance class distribution and enhance the overall effectiveness of the network. Furthermore, the influence of varying hyperparameters on reducing network error was systematically analyzed.

**Main Results:** Various network architectures and structures were examined. The experimental results indicated that the optimal model achieved an accuracy of 94.45% on the testing dataset with the following hyperparameters: a lookback of 7, a learning rate of 0.001, a dropout rate of 0.2, a batch size of 32, and a four-layer network architecture with 512, 256, and 256 units in the first, second, and third hidden layers, respectively. This configuration yielded higher accuracy and fewer false alarms in anomaly detection compared to other tested models. Based on the obtained results, this approach demonstrates significant potential for real-time anomaly detection in drilling operations.

\* Corresponding author: [khoozan@modares.ac.ir](mailto:khoozan@modares.ac.ir)



مقاله تحقیقاتی

## پیش‌بینی برخط رخداد‌های حفاری بر اساس نمودارگیری گل با استفاده از شبکه عصبی حافظه کوتاه‌مدت طولانی

مهران بادینلو<sup>۱</sup>، جمال فراشینی<sup>۱</sup>، داود خوزان<sup>۱\*</sup>، مهدی منصوری<sup>۲</sup>، وحید دانشخواه<sup>۲</sup>  
<sup>۱</sup> گروه مهندسی نفت، دانشکده مهندسی شیمی، دانشگاه تربیت مدرس، تهران، ایران.  
<sup>۲</sup> شرکت انرژی دانا، تهران، ایران.

### چکیده

**موضوع تحقیق:** عملیات حفاری همواره دچار مشکلات فراوانی بوده است. این مشکلات می‌توانند خسارات مالی، جانی یا حتی محیط زیستی فراوانی ایجاد کنند. به همین دلیل به دنبال راه‌حلی باید بود که این مشکلات را کاهش دهد و قبل از وارد شدن خسارات جانی و مالی، این اتفاقات را پیش‌بینی کند و اقدامات لازم را برای رفع و کاهش خسارات عملی سازد. در این پژوهش، تأثیر استفاده از الگوریتم‌های هوش مصنوعی برای پیش‌بینی رخداد‌های حفاری مورد بررسی قرار می‌گیرد. برای این منظور از داده‌های نمودارگیری از گل حفاری (Mud Logging) که به صورت برخط اندازه‌گیری شده و از چاه‌های موجود در یکی از میادین نفتی ایران جمع‌آوری شده‌اند، استفاده می‌شود.

**روش تحقیق:** معماری ترکیبی شامل شبکه عصبی حافظه کوتاه مدت طولانی (LSTM) و شبکه عصبی تماماً متصل برای شناسایی و تشخیص ناهنجاری‌هایی مانند سیلان چاه (Kick) و گیرکردن لوله حفاری (Stuck Pipe) به کار گرفته شد. به دلیل کمبود نمونه‌های این ناهنجاری‌ها در مجموعه داده‌ها که می‌تواند به‌طور قابل توجهی دقت و عملکرد مدل را کاهش دهد، از روش نمونه‌برداری دسته اقلیت (SMOTE) برای تعادل توزیع دسته‌ها و بهبود عملکرد کلی شبکه استفاده شد. علاوه بر این، تأثیر تغییرات ابرمتغیرها بر کاهش خطای شبکه مورد بررسی قرار گرفت.

**نتایج اصلی:** در شبکه‌هایی که با ساختارها و معماری‌های مختلف مورد بررسی قرار گرفتند، نتایج تجربی نشان داد که شبکه عصبی با دقت ۹۴/۴۵ درصد بر روی مجموعه داده‌های آزمایشی عمل کرد. این عملکرد با تنظیم ابرمتغیرها به صورت زیر حاصل شد: پنجره نگاه به عقب ۷، نرخ یادگیری ۰/۰۱، نرخ حذف ۰/۲، اندازه دسته ۳۲ و معماری شبکه چهار لایه با تعداد واحدهای {۵۱۲، ۲۵۶، ۲۵۶} در لایه‌های پنهان اول، دوم و سوم. این پیکربندی در تشخیص ناهنجاری نسبت به سایر گزینه‌های آزمایشی شده، دقت بالاتر و هشدارهای کاذب کمتری داشت. با توجه به نتایج به‌دست‌آمده این روش در تشخیص برخط ناهنجاری در فرایند حفاری مؤثر خواهد بود.

### اطلاعات مقاله

دریافت: ۱۴۰۳/۰۵/۲۲

پذیرش: ۱۴۰۳/۰۸/۲۹

دسترس آنلاین: ۱۴۰۳/۱۲/۲۵

ISSN: 2588-5316

Online ISSN: 2588-5324

### کلیدواژه‌ها

نمودارگیری از گل حفاری

شبکه عصبی حافظه کوتاه‌مدت طولانی

سیلان چاه

گیر رشته حفاری

الگوریتم نمونه‌گیری دسته اقلیت

## ۱ مقدمه

با توجه به افزایش تقاضای جهانی برای نفت و گاز، بهره‌وری هر چه بیشتر از مخازن هیدروکربوری امری ضروری است. آغاز عملیات حفاری در میدان نفتی گامی مؤثر در راستای دستیابی به این هدف محسوب می‌شود. شایان ذکر است که اقتصاد کشورهایی نظیر ایران به شدت به درآمدهای حاصل از فروش نفت و گاز وابسته است. لذا، افزایش تولید از مخازن هیدروکربوری، از اهمیت بالایی برخوردار است، با این تفاسیر باید توجه داشت که امروزه حفر چاه به‌منظور حفاری و تولید نفت از مخازن با مشکلات متعددی همراه است که وقوع آن‌ها می‌تواند ضرر و زیان‌های جبران‌ناپذیری بر پیکره اقتصاد کشور و همچنین محیط زیست به جا گذارد. همان‌طور که می‌دانید یکی از ملزومات حفاری، استفاده از ابزارآلات و تجهیزات با کیفیت است و هر گونه خرابی یا مشکل در چاه، خسارات سنگینی را به ارمغان می‌آورد. مطالعات متعدد در حوزه حفاری بیانگر این است که حفاری همچنان با مشکلاتی از قبیل سیلان چاه و هرزروی سیال حفاری (Lost Circulation) و گیر رشته لوله حفاری و نشت لوله‌های حفاری (Washout) روبه‌رو است.

طی سالیان اخیر محققان، با به‌کارگیری مفهومی نو به‌نام هوش مصنوعی به سراغ حل چنین مشکلاتی رفته‌اند. این علم به حفار این امکان را می‌دهد که قبل از وقوع هر یک از این حوادث، به‌صورت برخط از وقوع احتمالی آن باخبر شود و اقدامات مناسب برای جلوگیری از آن را عملی سازد. در سال ۲۰۱۷ محققان توانستند با استفاده از عوامل فشار پمپ (Pump Pressure)، وزن گل حفاری (Mud Weight)، نرخ جریان پمپ (Flow Pump) و عمق و به کمک شبکه‌های عصبی مصنوعی، آسیب هرزروی سیال را تشخیص دهند [۱]. در سال ۲۰۱۹ محققان با استفاده از عوامل وزن گل حفاری، لیتولوژی، نرخ جریان سیال، نرخ نفوذ مته حفاری (Rate of Penetration)، درصد جامدات، فشار سیال در حال چرخش (Mud Pressure Circulation) و ... به کمک شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان (Support Vector Machine) به تشخیص هرزروی سیال حفاری پرداختند که شبکه ماشین بردار پشتیبان، عملکرد قابل قبولی داشت [۲]. در سال ۲۰۲۰ نیز محققان به‌کمک خصوصیات زمین‌شناسی، عوامل عملیات حفاری و خصوصیات سیال حفاری و به کمک شبکه‌های عصبی مصنوعی به تشخیص هرزروی سیال حفاری پرداختند [۳].

از دیگر آسیب‌هایی که حین حفاری به وفور رخ می‌دهد سیلان چاه است. سیلان چاه زمانی رخ می‌دهد که فشار سیال حفاری بر فشار سازند غلبه نکند و سیال داخل سازند به داخل چاه جریان می‌یابد. در سال ۲۰۱۹ محققان توانستند با استفاده از عوامل نرخ نفوذ مته حفاری، وزن روی مته (Weight on Bit)، بار قلاب (Hook Load)، فشار لوله ایستاده (Standpipe Pressure)، نرخ ورودی سیال، نرخ خروجی سیال، دمای گاز و نرخ چرخش رشته حفاری (Rotary per Minutes) و به کمک شبکه‌های عصبی (Neural Networks)، آسیب سیلان چاه را حین حفاری تشخیص دهند [۴]. همچنین در سال ۲۰۲۰ نیز به کمک عوامل فشار ته چاهی، نرخ خروجی گل حفاری، چگالی سیال حفاری (Mud Density)، نرخ سیال حفاری پمپ

شده، وزن روی مته و سرعت چرخش مته حفاری و به کمک شبکه‌های عصبی مصنوعی انتشار رو به جلو آسیب سیلان چاه را حین حفاری تشخیص دهند [۵]. در سال ۲۰۲۱ نیز محققان با استفاده از عوامل فشار ته‌چاهی، چگالی گل حفاری، رسانایی گل حفاری (Mud Conductivity) و نرخ جریان سیال حفاری و با کمک دسته‌بندی‌کننده بیز ساده (Naïve Bayes)، سیلان چاه را حین حفاری تشخیص دهند [۶].

از دیگر آسیب‌هایی که به وفور در حین حفاری رخ می‌دهد، گیر رشته حفاری است. گیر رشته حفاری معمولاً زمانی رخ می‌دهد که رشته حفاری نمی‌تواند در امتداد محور چاه حرکت کند. در سال ۲۰۱۹ محققان با استفاده از عوامل سطحی، رئولوژیکی (Rheology)، میزان شدت انحراف (Dogleg Severity) و با استفاده از شبکه‌های ماشین بردار پشتیبان، جنگل تصادفی (Random Forest)، شبکه‌های عصبی مصنوعی و درخت تصمیم‌گیری (Decision Tree) به تشخیص زود هنگام گیر رشته حفاری پرداختند که مدل جنگل تصادفی عملکرد بهتری نسبت به دیگر شبکه‌ها داشت [۷]. همچنین در سال ۲۰۱۹ محققان توانستند با استفاده از عوامل نیروی قلاب، میانگین بار قلاب برای ۱۲ گام زمانی، نرخ تغییر بار قلاب، میزان نرخ نفوذ حفاری، فشار لوله ایستاده، میانگین فشار لوله ایستاده برای ۱۲ گام زمانی، نرخ تغییر فشار لوله ایستاده، گشتاور، میانگین گشتاور برای ۱۲ گام زمانی، نرخ تغییر گشتاور، وزن روی مته، نرخ جریان سیال حفاری ورودی و نرخ چرخش رشته حفاری و با کمک شبکه‌های عصبی مصنوعی، درخت تصمیم‌گیری و ماشین بردار پشتیبان به تشخیص گیر رشته حفاری بپردازند که شبکه ماشین بردار پشتیبان عملکرد بهتری داشت [۸].

در سال ۲۰۲۰ محققان به کمک دو عامل استخراج روند ویژگی در مقیاس چند زمانه و شبکه عصبی احتمالی، به‌صورت برخط از وقوع چند حادثه در ته چاه حین حفاری مطلع و از آن جلوگیری کردند [۹]. رویکرد استخراج روند ویژگی دوقسمتی خواهد بود که عبارتند از: روند تغییر سریع، روند تغییر آهسته. عوامل دارای روند تغییر سریع نسبت به زمان را می‌بایست در مقیاس زمانی کوتاه مورد بررسی قرار داد که عبارتند از: وزن روی مته، گشتاور و نرخ چرخش رشته حفاری. همچنین عواملی که دارای روند تغییر آهسته هستند را می‌بایست در مقیاس زمانی بلند مورد بررسی قرار داد که عبارتند از: نرخ حجم محفظه گل، فشار لوله ایستاده، تعداد ضربه بر دقیقه، نرخ خروجی گل. در این مطالعه محققان برای تشخیص آسیب‌های حین حفاری از شبکه عصبی احتمالی استفاده کردند [۹]. در این شبکه علاوه بر عوامل اصلی و کلیدی حفاری، از عوامل استخراج روند ویژگی در مقیاس زمانی بلند و کوتاه نیز استفاده شده که در قدم اول و قبل از کاربرد آن به‌عنوان ورودی شبکه، داده‌ها به‌نچارج شده‌اند. مدل پیشنهادی دقت و کارایی قابل‌توجهی در تشخیص و شناسایی حوادث رایج حفاری از قبیل گیر رشته حفاری، هرزروی سیال حفاری، شستشوی لوله حفاری و بیش کشش (Overpulling) از خود نشان می‌دهد. این بهبود عملکرد در مقایسه با مواردی که تنها از عوامل اصلی در لایه ورودی شبکه استفاده می‌شود، حاصل شده است [۹]. در سال ۲۰۲۱ محققان با مطالعه بر روی همبستگی

ایستاده، میانگین وزن روی مته، میانگین نرخ سیال ورودی، میانگین نرخ سیال خروجی، گازهای اتان تا هپتان، میانگین گشتاور، میانگین نرخ چرخش رشته حفاری و میانگین بار قلاب. داده‌های مذکور، مربوط به ۸۶ نقطه برحسب عمق آسیب رخ داده حین حفاری است که در ۲۹ نقطه وقوع سیلان چاه و در ۵۶ نقطه وقوع گیر رشته حفاری گزارش شده و مابقی داده‌ها در شرایط بهنجار هستند.

## ۲-۲ پیش‌پردازش

شروع هر نوع کار و عملیاتی در مرحله اول، دارای مجموعه‌ای از مقدمات و پیش‌نیازها است. داده‌کاو نیز از این قانون مستثنی نبوده و نیازمند آماده‌سازی و پردازش‌های مقدماتی است. پیش‌پردازش نقش اساسی را در روند پردازش داده‌ها و نتایج حاصل از آن ایفا می‌کند. روش بهنجارسازی داده‌ها از جمله مهم‌ترین مراحل پیش‌پردازش (Preprocessing) در علم داده‌کاو است. متغیرها دارای بازه وسیعی از مقادیر است که همین امر موجب کاهش دقت و عملکرد شبکه خواهد بود. با استفاده از بهنجارسازی کمینه-بیشینه (Min-Max Normalization)، بازه مقادیر متغیرها به بازه [۱۰] نگاشت شده است. رابطه ریاضی این روش بهنجارسازی به صورت زیر است:

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

که در این رابطه  $x$  مقدار اصلی،  $\hat{x}$  مقدار بهنجار شده،  $x_{min}$  کمینه مقدار ویژگی مورد نظر در مجموعه داده و  $x_{max}$  بیشینه مقدار ویژگی مورد نظر در مجموعه داده است.

## ۲-۳ بیش‌برازش

بررسی داده‌ها نشان می‌دهد که تعداد داده‌های مربوط به آسیب‌های ذکر شده در مقایسه با داده‌های بهنجار شده به شدت کم است. به همین علت میزان دقت و عملکرد شبکه برای تشخیص آسیب‌ها بر حسب عمق، کاهش یافته و قابلیت تعمیم ندارد و در این حالت شبکه دچار بیش‌برازش (Overfitting) می‌شود. برای جلوگیری از وقوع بیش‌برازش، روش‌های مختلفی به کار گرفته می‌شود. در این پژوهش از روش حذف موقت (Dropout) استفاده شده است. در این روش احتمال حذف هر گره به صورت تصادفی تعیین شده است. در این روش شبکه عصبی تمایل به اختصاص وزن بالا به ویژگی اصلی ندارد زیرا که امکان حذف آن‌ها نیز وجود دارد. به همین دلیل مقادیر وزن کمتر بین همه ویژگی‌ها تقسیم شده است که این امر تاثیر به‌سزایی در نظم‌دهی و جلوگیری از بیش‌برازش مدل دارد.

## ۲-۴ الگوریتم نمونه‌گیری دسته اقلیت

در بسیاری از کاربردهای هوش مصنوعی و علم داده، محققان و دانشمندان با داده‌های نامتوازن روبه‌رو هستند. یکی از فنون قابل قبول و پر کاربرد برای به‌کارگیری و مقابله با داده‌های نامتوازن، روش‌های بازنمونه‌گیری

زمانی بین عوامل حفاری، روش تشخیص آسیب مبتنی بر خوشه‌بندی ویژگی برحسب توالی زمانی ارائه داده‌اند. عوامل مورد استفاده در این مطالعه عبارتند از: ارتفاع بلاک (Block Position)، وزن روی مته، بار قلاب، نرخ چرخش رشته حفاری، گشتاور، فشار لوله ایستاده، نرخ ورودی سیال، نرخ خروجی سیال، حجم‌گل‌محفظه (Mud Pit)، چگالی گل، رسانایی گل، دمای ورودی گل و دمای خروجی گل.

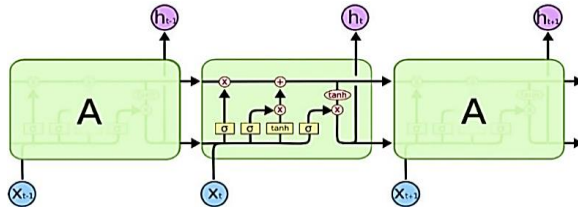
فرایند تشخیص آسیب درون چاه حین حفاری دارای ۴ مرحله است. مرحله اول استفاده از روش همبستگی فاصله‌ای (Distance Correlation)، مرحله دوم استفاده از روش منکندل (Mann Kendell Trend Test)، مرحله سوم استفاده از روش خوشه‌بندی مبتنی بر چگالی (Feature Clustering)، و مرحله چهارم استفاده از دسته‌بندی‌کننده بیز ساده. در این مدل، هریک از خوشه‌ها به‌عنوان داده ورودی به شبکه مذکور داده شده، در نهایت نوع آسیب در آن دسته‌بندی می‌شود. دقت این مدل بالای ۹۵ درصد است [۱۰]. با توجه به مطالب ارائه شده، می‌توان نتیجه گرفت که استفاده از هوش مصنوعی در تشخیص و پیش‌بینی حوادث حفاری به‌عنوان رویکردی نوین، مورد توجه بسیاری از محققان قرار گرفته است. مطالعات پیشین عمدتاً بر تشخیص تک‌حادثه‌ای همچون هرزروی سیال، سیلان چاه یا گیر رشته حفاری با استفاده از روش‌های مبتنی بر داده متمرکز بوده‌اند. در آن مطالعات از روش‌های مختلفی مانند ماشین بردار پشتیبان، شبکه‌های عصبی پرسپترون چندلایه و ... برای تشخیص حوادث حفاری استفاده شده است که این روش‌ها توالی زمانی موجود در داده‌های حفاری را در نظر نگرفته‌اند. به این دلیل، با عنایت به ماهیت توالی زمانی داده‌های حفاری و وجود وابستگی‌های طولانی‌مدت بین آن‌ها، استفاده از مدل‌های یادگیری مبتنی بر شبکه‌های عصبی بازگشتی (Recurrent Neural Network) و به‌طور خاص، حافظه کوتاه مدت طولانی (LSTM) می‌تواند نتایج بهتری را به همراه داشته باشد. LSTM به دلیل توانایی در یادگیری الگوهای پیچیده و حفظ اطلاعات در طول زمان، انتخاب مناسبی برای تحلیل داده‌های توالی زمانی در حوزه حفاری است. در این پژوهش، با بهره‌گیری از قابلیت‌های LSTM، به مدلی توسعه می‌یابد که می‌تواند حوادث حفاری شامل گیر لوله و سیلان چاه را با دقت بیشتری تشخیص دهد و به این ترتیب، به بهبود ایمنی و بهره‌وری عملیات حفاری کمک کند.

## ۲ بخش نظری

### ۲-۱ جمع‌آوری داده

در این پژوهش، با هدف شناسایی و پیش‌بینی چالش‌های حفاری، داده‌های حاصل از عملیات حفاری در یکی از میادین نفتی ایران مورد تحلیل قرار گرفت. تعداد و نوع حوادث رخ داده در این چاه‌ها، بر اساس داده‌های موجود، به‌عنوان مبنایی برای این مطالعه در نظر گرفته شد. بررسی نمودارهای گل حفاری نشان داد که دو حادثه اصلی سیلان چاه و گیر رشته حفاری به‌عنوان چالش‌های عمده حفاری در این میدان نفتی شناسایی شده‌اند. عوامل مورد استفاده عبارتند از: میانگین نرخ نفوذ مته حفاری، میانگین فشار لوله

لاپه‌های مخفی، چندین درگاه طوری سازمان‌دهی شده‌اند که بتوانند داده‌های ورودی با طول زیاد را به‌خاطر بسپارد. شکل ۱ شبکه عصبی حافظه کوتاه‌مدت طولانی را نشان می‌دهد [۱۳].



شکل ۱ شبکه عصبی حافظه کوتاه‌مدت طولانی

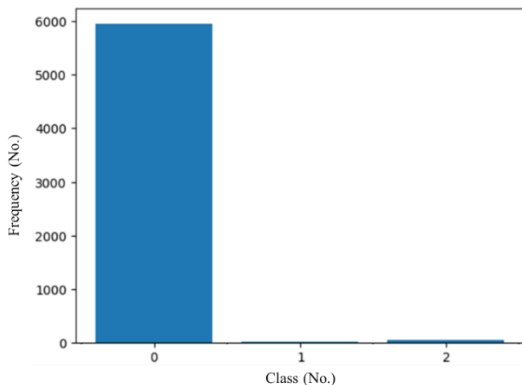
Figure 1. Long-Short Term Memory Neural Network

یکی از اساسی‌ترین ابرمتغیرها در این نوع شبکه‌ها، مقدار نگاه به عقب (Look back) است. این مفهوم تعیین می‌کند که چند مرحله زمانی قبلی برای پیش‌بینی گام‌های زمانی بعدی استفاده می‌شود؛ زیرا که در قسمت ورودی این شبکه، علاوه بر متغیر ورودی در لحظه فعلی، اطلاعات لحظه قبلی برای افزایش کارایی و دقت مدل به‌عنوان ورودی در شبکه استفاده خواهد شد. برای تشخیص نوع آسیب در لایه آخر از شبکه عصبی مصنوعی با تابع فعال‌سازی بیشینه هموار (Softmax Activation Function) استفاده شده که در آن خروجی شبکه عصبی بازگشتی کوتاه‌مدت طولانی به‌عنوان ورودی شبکه عصبی مصنوعی قابل استفاده است و نوع آسیب را متناسب با هر عمق نشان می‌دهد.

### ۳ نتایج و بحث

#### ۳-۱ نمونه‌گیری دسته اقلیت

با بررسی و به‌کارگیری داده‌های نمودارگیری گل حفاری چاه‌های موجود در یکی از میادین نفتی ایران، مشاهده شد که آسیب سیلان چاه در اعماق ۱۳۸۰ و ۱۶۸۲ متر و گیر رشته حفاری نیز در اعماق ۲۳۷۵، ۲۶۱۸ و ۳۹۵۴ متر رخ داده است. شکل ۲ توزیع اولیه داده‌ها را قبل از پیش‌پردازش و به‌کارگیری آن در شبکه نشان می‌دهد. در این پژوهش حفاری عادی، سیلان چاه و گیر رشته حفاری به‌ترتیب با شماره دسته‌های ۰، ۱ و ۲ نمایش داده شده‌اند.



شکل ۲ توزیع اولیه داده‌ها

Figure 2 Initial Data Distribution

(Resampling) است [۱۱]. از پرکاربردترین روش‌ها برای مقابله با عدم تعادل این دسته در داده‌های مربوط، استفاده از الگوریتم نمونه‌گیری دسته اقلیت (Synthetic Minority Oversampling Technique) است. در این روش نمونه‌های بیشتری به دسته اقلیت (Minority Class) اضافه می‌شوند و به این ترتیب بین نمونه‌های مربوط به هر دسته، تعادل برقرار می‌شود.

#### ۲-۵ جستجوی شبکه با اعتبارسنجی متقاطع

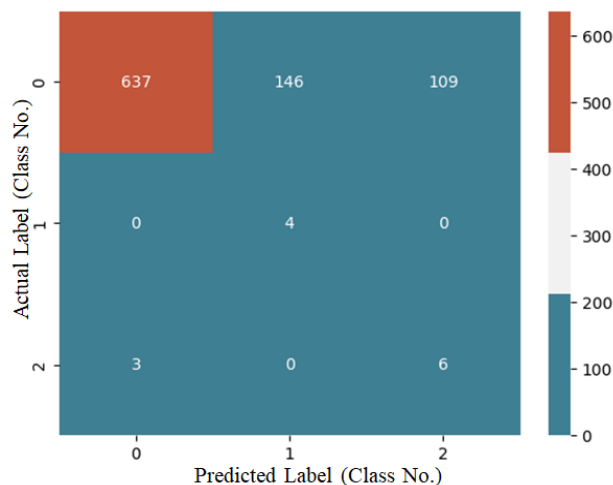
متغیرهای مدل، دارای مقادیری هستند که برای پیکربندی مدل‌ها یا الگوریتم یادگیری استفاده می‌شوند. متغیرها را هم می‌توان از داده‌ها آموخت و هم به‌صورت دستی تنظیم کرد. مقادیر متغیرها، روش کار الگوریتم یا مدل را تعیین می‌کنند. بنابراین متغیرها می‌توانند تأثیر قابل توجهی بر عملکرد مدل‌ها داشته باشند. درحالی‌که ابرمتغیرها (Hyperparameters) توسط متخصص علم داده و حین آموزش مشخص می‌شوند. تعیین بهترین ابرمتغیرها معمولاً کاری غیرممکن است؛ ولی با آزمون و خطا می‌توان به مقادیر مناسبی دست یافت. از پرکاربردترین روش‌ها در تعیین مقادیر بهینه ابرمتغیرها، روش جستجوی شبکه‌ای با اعتبارسنجی متقاطع (Grid Search CV) است. در این روش به‌جای نمونه برداری تصادفی، بر روی جستجوی خود متمرکز شده و همه ترکیبات مورد ارزیابی قرار می‌گیرد [۱۲].

#### ۲-۶ الگوریتم بهینه‌سازی آدام

یکی از اساسی‌ترین مؤلفه‌ها در طراحی مدل باکیفیت، متغیرهای موجود در شبکه است. عدم توجه به تنظیم و بهینه‌سازی آن‌ها عملکرد شبکه را کاهش می‌دهد و شبکه دچار خطا می‌شود. بهینه‌سازها به الگوریتم‌ها یا روش‌هایی اطلاق می‌شود که برای تغییر متغیرهای شبکه عصبی از قبیل وزن و نرخ یادگیری به‌کار گرفته می‌شوند. این کار، باعث کاهش تابع هزینه (Loss Function) می‌شود. در این مطالعه از روش الگوریتم بهینه‌سازی آدام (Adam Optimization Algorithm) استفاده شده است. الگوریتم آدام یک الگوریتم بهینه‌سازی است که می‌توان از آن برای به‌روزرسانی وزن‌های شبکه بر اساس تکرار در داده‌های آموزشی استفاده کرد.

#### ۲-۷ الگوی ترکیبی

آسیب‌های حین حفاری از قبیل کیک چاه و گیر رشته حفاری می‌تواند خسارات جبران‌ناپذیری بر پیکره اقتصاد کشور و همچنین محیط زیست به‌جای بگذارد. برای شناسایی و پیش‌بینی آسیب‌های مذکور قبل از وقوع، از الگوی ترکیبی شبکه عصبی بازگشتی کوتاه‌مدت طولانی و شبکه عصبی مصنوعی استفاده شده است. در شبکه‌های عصبی بازگشتی حافظه دارای محدودیت است و به‌همین علت با افزایش طول داده‌های ورودی، اثر اطلاعات حفظ شده لحظات ابتدایی کمتر شده و مشکل محوشدگی گرادبان (Vanishing Gradient) رخ داده است. برای بهبود این مشکل و تقویت حافظه از شبکه عصبی حافظه کوتاه‌مدت طولانی به‌جای آن استفاده شده است. تفاوت این شبکه با شبکه عصبی بازگشتی معمولی این است که در



شکل ۳ ماتریس درهم‌ریختگی داده‌های آزمایشی (روش کاهش نمونه‌های دسته اکثریت)

Figure 3 Confusion Matrix Test Data (Undersampling Method)

روش‌های دیگر نیز که ترکیبی از روش‌های کم کردن نمونه‌های اکثریت و افزایش نمونه‌های اقلیت هستند، به‌زمان پردازش بیشتری نیاز دارند و از طرفی با توجه به کم کردن نمونه‌های اکثریت ممکن است اطلاعات از دست برود. از این رو استفاده از روش SMOTE باعث می‌شود که مدلی با دقت مناسب با زمان پردازشی کم ایجاد شود.

### ۳-۲ بررسی اثر مقدار نگاه به عقب

برای افزایش کارایی شبکه در راستای ارزیابی داده‌های اعتبارسنجی، ابرمتغیرهای موجود در شبکه باید در گام بعدی بهینه شوند. ابرمتغیرهای موجود در شبکه عبارتند از: نگاه به عقب، نرخ یادگیری، تعداد لایه‌های پنهان (Hidden Layer)، تعداد بخش‌های پنهان (Hidden Unit)، اندازه دسته و نرخ حذف موقت. براین اساس مقادیر {۷، ۵، ۳} برای نگاه به عقب، مقادیر {۰/۱، ۰/۰۰۱، ۰/۰۰۰۱} برای نرخ یادگیری، مقادیر {۰/۲، ۰/۴}، {۰/۱} برای نرخ حذف موقت، مقادیر {۱۲۸، ۶۴، ۳۲، ۱۶} و برای اندازه دسته و برای بررسی تعداد لایه‌های پنهان و بخش‌های آن، مقادیر {۵۱۲، ۲۵۶، ۱۲۸، ۶۴، ۳۲}، {۵۱۲، ۲۵۶، ۱۲۸، ۶۴}، {۵۱۲، ۲۵۶، ۱۲۸}، {۵۱۲، ۲۵۶، ۱۲۸، ۶۴، ۳۲} در راستای بررسی و ارزیابی مقادیر ابرمتغیرهای مذکور، از الگوریتم جستجوی شبکه اعتبارسنجی متقاطع استفاده شده است.

در این مرحله، به کمک مقدار نگاه به عقب ۳ داده‌های آموزش بازنمایی و در راستای شناسایی بهترین ابرمتغیرها مورداستفاده قرار می‌گیرد. در همین راستا بعد از آموزش شبکه و ارزیابی آن به کمک داده‌های اعتبارسنجی و بررسی میزان خطای موجود در شبکه، در نهایت با به‌کارگیری الگوریتم بهینه‌سازی جستجوی شبکه اعتبارسنجی متقاطع، مقادیر بهینه ابرمتغیرها و به تبع آن عملکرد قابل قبول شبکه شناسایی می‌شود. جدول ۱ خلاصه‌ای از شبکه‌ی مذکور به ازای مقدار نگاه به عقب ۳ را نشان می‌دهد.

با توجه به شکل ۲ مشاهده می‌کنید که در اکثریت داده‌ها شرایط حفاری به‌صورت نرمال بوده است و ۰/۴ درصد از داده‌ها دارای آسیب سیلان چاه و ۰/۹ درصد نیز دارای آسیب رخداد گیر رشته حفاری است. همان‌طور که ملاحظه می‌شود عدم تعادل در دسته خروجی بسیار شدید بوده که همین امر موجب کاهش دقت و عملکرد شبکه خواهد بود، زیرا که شبکه به‌درستی آموزش ندیده و قابلیت تعمیم و ارزیابی ندارد. در گام اول برای بهبود این مشکل و قبل از به‌کارگیری داده در شبکه مذکور، ابتدا با استفاده از روش بهنجارسازی کمینه-بیشینه بازه مقادیر متغیرها به بازه جدید [۰ و ۱] تغییر خواهد کرد. در گام بعدی حدود ۷۰ درصد داده‌ها برای آموزش شبکه، ۱۵ درصد داده‌ها برای اعتبارسنجی شبکه و مابقی برای آموزش مورد استفاده قرار می‌گیرد، سپس برای برقراری توازن بین داده‌ها در طبقه‌بندی از الگوریتم نمونه‌گیری دسته اقلیت استفاده می‌شود. این الگوریتم بر روی داده‌های آزمایشی اعمال نمی‌شود و داده‌های آزمایشی در حالت واقعی خود باقی می‌مانند. این کار به این دلیل انجام شده که مدل آموزش‌دیده باید بر روی داده‌های آزمایشی‌ای که نمایانگر وضعیت واقعی داده‌ها هستند، ارزیابی شود. اگر نمونه‌گیری بر روی داده‌های آزمایشی نیز انجام شود، در واقع اطلاعاتی از داده‌های آزمایشی به مدل آموزش داده می‌شود که منجر به نشت اطلاعات و ارزیابی نادرست عملکرد مدل می‌شود. به بیان دیگر، اعمال الگوریتم نمونه‌گیری بر روی داده‌های آزمایشی می‌تواند منجر به انحراف در ارزیابی عملکرد مدل شود و دقت واقعی مدل در دنیای واقعی را به‌درستی منعکس نکند. پس از اعمال این روش، سهم هر کلاس از اعمال این روش ۳۳/۳۳ درصد است و توازن بین داده‌ها برقرار می‌شود.

لازم به ذکر است که الگوریتم SMOTE به این دلیل انتخاب شده که یکی از کارآمدترین روش‌ها برای متعادل‌سازی داده‌های نامتوازن است. در این روش، با ایجاد داده‌های مصنوعی برای دسته اقلیت به جای تکرار داده‌ها یا حذف داده‌های اکثریت، تعادل داده‌ها حفظ می‌شود و تنوع بیشتری در داده‌های اقلیت ایجاد می‌شود. در مقابل، روش‌های کاهش نمونه‌های دسته اکثریت (Undersampling) هرچند ساده‌تر هستند، اما با کاهش تعداد نمونه‌ها از دسته اکثریت، امکان از دست رفتن اطلاعات مهم وجود دارد و معمولاً دقت مدل پایین می‌آید. با اعمال روش کاهش نمونه‌های دسته اکثریت بر روی مجموعه داده‌ی موجود با استفاده از بهترین ابرمتغیرهای مدل با نگاه رو به عقب ۳، دقت مدل برای مجموعه داده‌ی آزمایشی به ۸۰/۹۵ درصد رسیده است که برای روش SMOTE دقت مجموعه‌ی مجموعه‌ی آزمایشی ۹۲/۴۵ درصد است. نمودار ماتریس درهم‌ریختگی (Confusion Matrix) برای این حالت نیز در شکل ۳ نشان داده شده است. همان‌طور که مشاهده می‌شود در مقایسه با نتایج SMOTE (شکل ۶) هشدارهای کاذب بسیار زیادی را ایجاد کرده است که این امر باعث می‌شود قابلیت اطمینان مدل و دقت آن از دست برود.

جدول ۱ خلاصه‌ای از شبکه به ازای مقدار نگاه به عقب ۳

Table 1 Summary of Model Structure for Value of Look Back 3

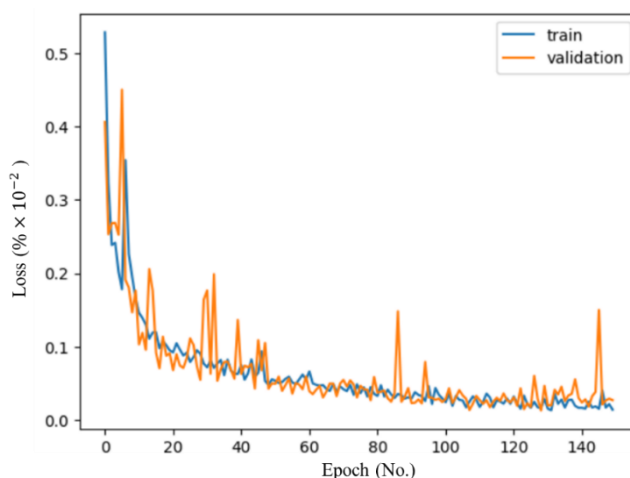
Layer (type)	Output Shape (size)	Param (No.)
lstm_1032 (LSTM)	(None, 3, 512)	1073152
dropout_1032 (Dropout)	(None, 3, 512)	0
lstm_1033 (LSTM)	(None, 3, 256)	787456
dropout_1033 (Dropout)	(None, 3, 256)	0
lstm_1034 (LSTM)	(None, 256)	525312
dropout_1034 (Dropout)	(None, 256)	0
dense_274 (Dense)	(None, 3)	771

جدول ۲ نتایج ترکیب شبکه‌های مذکور به ازای مقدار نگاه به عقب ۳

Table 2 Results of the Combination of the Networks with Look Back 3

Dataset	Long-Short Term Memory + Artificial Neural Network			
	Dropout Rate (proportion)	Learning Rate (proportion)	Hidden Unit (No.)	Accuracy (%)
Training	0.4	0.0001	{(512,256,256)}	<b>86.44</b>
Testing	0.4	0.0001	{(512,256,256)}	<b>92.45</b>

شکل ۴ نمودار تابع هزینه (زیان) را برحسب تکرار دوره (Epoch) نشان می‌دهد.

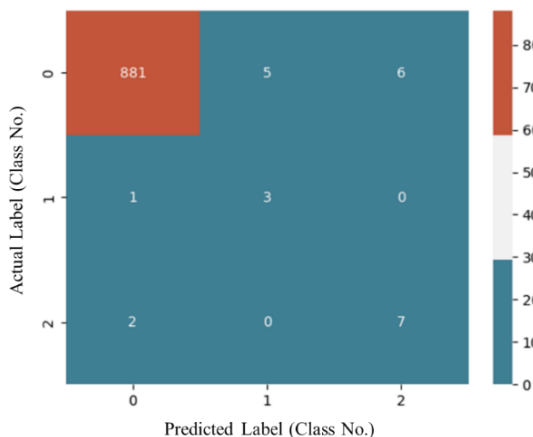


شکل ۴ نمودار تابع هزینه بر حسب تکرار دوره به ازای مقدار نگاه به عقب ۳

Figure 4 Loss Function Diagram for Look Back 3

جدول ۲، نتایج شبکه را به ازای مقادیر بهینه ابرمتغیرها بر روی داده‌های آموزش و آزمایش نشان می‌دهد.

شکل ۵ در مرحله بعدی نمودار ماتریس درهم‌ریختگی برای داده‌های آزمایش با مقادیر بهینه ابرمتغیر را با اندازه دسته ۳۲ نشان می‌دهد.



شکل ۵ ماتریس درهم‌ریختگی داده‌های آزمایش به ازای مقدار نگاه به عقب ۳

Figure 5 Confusion Matrix Test Data for Look Back 3



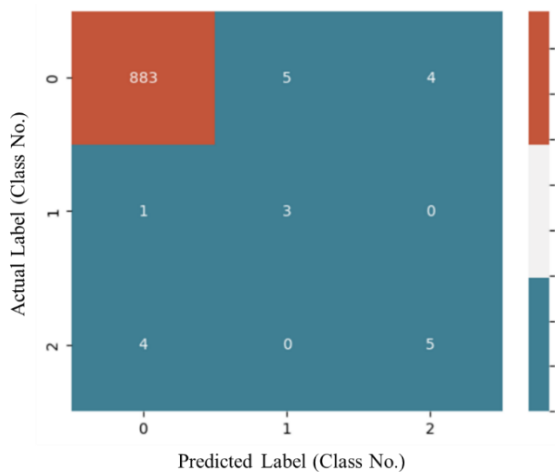
جدول ۴، نتایج شبکه را به ازای مقادیر بهینه ابرمتغیرها بر روی داده‌های آموزش و آزمایش نشان می‌دهد.

جدول ۴ نتایج ترکیب شبکه‌های مذکور به ازای مقدار نگاه به عقب ۵

Table 4 Results of the Combination of the Networks with Look Back 5

Dataset	Long-Short Term Memory + Artificial Neural Network			
	Dropout Rate (proportion)	Learning Rate (proportion)	Hidden Unit (No.)	Accuracy (%)
Training	0.2	0.01	{(512,256,256)}	85.19
Testing	0.2	0.01	{(512,256,256)}	91.45

شکل ۷ نمودار ماتریس درهم‌ریختگی برای داده‌های آزمایش با مقادیر بهینه فرامغیر را با اندازه دسته ۳۲ نشان می‌دهد.



شکل ۷ ماتریس درهم‌ریختگی داده‌های آزمایش با مقدار نگاه به عقب ۵

Figure 7 Confusion Matrix Test Data for Value of Look Back 5

با مشاهده شکل ۷ مقادیر زیادی از داده‌های آزمایش به‌درستی و به حالت بهنجار طبقه بندی شده است، اما ۹ مورد از آن به اشتباه توسط شبکه طبقه بندی شده است. ۴ مورد از داده‌های مربوط به حالت سیلان چاه طبقه بندی شده است که از آن، ۳ مورد به درستی توسط شبکه دسته بندی شده است و همچنین ۹ مورد از داده‌ها به حالت گیر رشته حفاری طبقه بندی شده که ۵ مورد از آن به درستی دسته بندی شده است. در چنین حالتی ملاحظه می‌کنید که عملکرد مدل برای تشخیص آسیب گیر رشته حفاری مطلوب نبوده و در اغلب موارد دچار خطا شده و به تبع آن میزان دقت کاهش یافته و برابر ۹۱ درصد می‌شود.

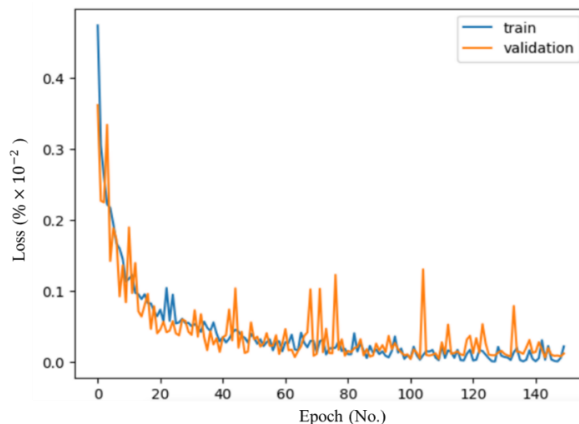
در شکل ۵ برچسب صفر وقوع شرایط بهنجار، برچسب یک رخداد سیلان چاه و برچسب دو رخداد گیر رشته حفاری را نشان می‌دهد. محور افقی نمودار نتایج پیش‌بینی شده توسط شبکه و محور عمودی آن نتایج واقعی را نشان داده است. با مشاهده شکل ۵ می‌توان دریافت که مقادیر زیادی از داده‌های آزمایش به درستی و به حالت بهنجار طبقه بندی شده است، اما ۱۱ مورد از آن به اشتباه توسط شبکه طبقه بندی شده است. ۴ مورد از داده‌های مربوط به حالت سیلان چاه طبقه بندی شده است که از آن ۳ مورد به درستی توسط شبکه دسته بندی شده است و همچنین ۹ مورد از داده‌ها به حالت گیر رشته حفاری طبقه بندی شده که ۷ مورد از آن به درستی دسته بندی شده است و در چنین حالتی عملکرد مدل قابل قبول بوده و میزان دقت آن بر روی داده‌های آزمایش با توجه به جدول ۲ مقدار ۹۲ درصد به دست آمده است. در مرحله بعدی، مجدداً شبکه را به ازای مقدار دور نگاه به عقب ۵ مورد بررسی قرار گرفته است. فرایند در این مرحله، مشابه فرایند در مرحله قبل به ازای دور نگاه به عقب ۳ خواهد بود. بعد از به کارگیری الگوریتم بیش نمونه گیری اقلیت مصنوعی و دور نگاه به عقب با مقدار ۵ در مسئله، با کمک الگوریتم جستجوی شبکه اعتبارسنجی متقاطع مقادیر بهترین ابرمتغیرهای موجود در شبکه شناسایی و انتخاب شده است. جدول ۳ خلاصه‌ای از ساختار شبکه به ازای مقدار نگاه به عقب ۵ را نشان می‌دهد.

جدول ۳ خلاصه‌ای از شبکه به ازای مقدار نگاه به عقب ۵

Table 3 Summary of Model Structure for Value of Look Back 5

Layer (type)	Output Shape (size)	Param (No.)
lstm_1020 (LSTM)	(None, 5, 512)	1073152
dropout_1020 (Dropout)	(None, 5, 512)	0
lstm_1021 (LSTM)	(None, 5, 256)	787456
dropout_1021 (Dropout)	(None, 5, 256)	0
lstm_1022 (LSTM)	(None, 256)	525312
dropout_1022 (Dropout)	(None, 256)	0
dense_274 (Dense)	(None, 3)	771

شکل ۶ نیز نمودار تابع هزینه (زیان) را برحسب تکرار دوره نشان می‌دهد:



شکل ۶ نمودار تابع هزینه بر حسب تکرار دوره به ازای مقدار نگاه به عقب ۵

Figure 6 Loss Function Diagram for Value of Look Back 5

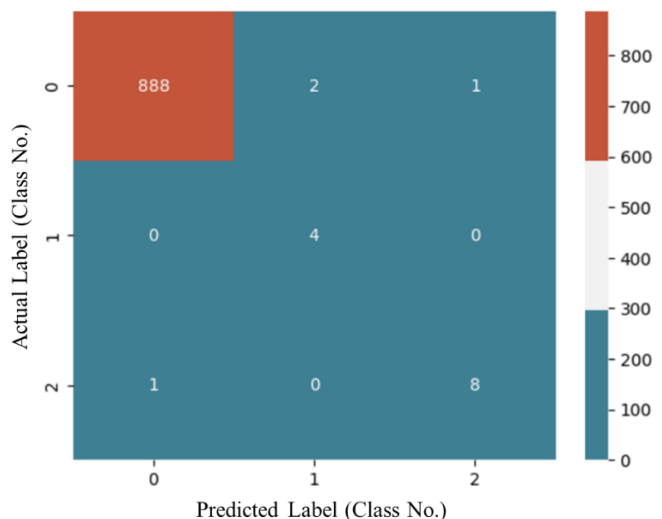


جدول ۶ نتایج ترکیب شبکه‌های مذکور به ازای مقدار نگاه به عقب ۷

**Table 6** Results of the Combination of the Networks with Look Back 7

Dataset	Long-Short Term Memory + Artificial Neural Network			
	Dropout Rate (proportion)	Learning Rate (proportion)	Hidden Unit (No.)	Accuracy (%)
Training	0.2	0.001	{{(512,256,256)}}	95.32
Testing	0.2	0.001	{{(512,256,256)}}	94.45

شکل ۹ نمودار ماتریس درهم‌ریختگی برای داده‌های آزمایش را با مقادیر بهینه ابرمتغیر را با اندازه دسته ۳۲ نشان می‌دهد.



شکل ۹ ماتریس درهم‌ریختگی داده‌های آزمایش به ازای مقدار نگاه به عقب ۷  
**Figure 9** Confusion Matrix Testing Data for Value of Look Back 7

با مشاهده شکل ۹ مقادیر زیادی از داده‌های آزمایش به‌درستی و به‌حالت بهنجار طبقه‌بندی شده است و تنها ۳ مورد از آن به اشتباه توسط شبکه طبقه‌بندی شده است. ۴ مورد از داده‌های مربوط به‌حالت سیلان چاه طبقه‌بندی شده است که در این حالت هر ۴ مورد به‌درستی و به کلاس سیلان چاه طبقه‌بندی شده و همچنین ۹ مورد از داده‌ها به حالت گیر رشته حفاری طبقه‌بندی شده که ۸ مورد از به‌درستی دسته‌بندی شده است و تنها در ۱ مورد تشخیص درست نبوده است. با بررسی نمودارها و نتایج به‌دست‌آمده مشاهده می‌کنید که افزایش داده‌های زمان‌های قبلی تأثیر به‌سزایی در دقت شبکه در زمان فعلی دارد و میزان دقت برابر ۹۴ درصد است.

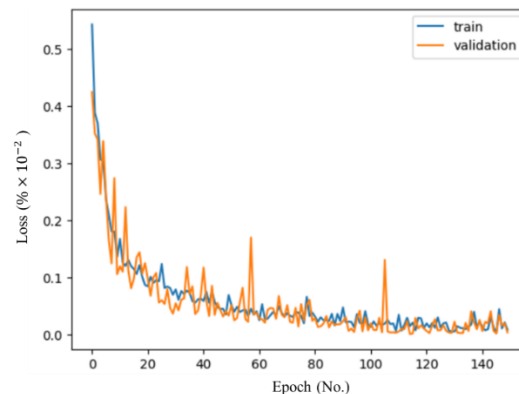
در این مرحله، مسئله مجدداً به‌ازای مقدار نگاه به عقب ۷ مورد بررسی قرار گرفته است. در ابتدا داده‌های آموزش به کمک الگوریتم بیش‌نمونه‌گیری با اقلیت مصنوعی بازنمایی شده و مقدار دور نگاه به عقب ۷ را اعمال خواهید کرد. در همین راستا در مسئله به کمک الگوریتم جستجوی شبکه با اعتبارسنجی متقاطع، مقادیر بهترین ابرمتغیرها را شناسایی و انتخاب شده است. جدول ۵ خلاصه‌ای از ساختار شبکه به ازای مقدار نگاه به عقب ۷ را نشان می‌دهد:

جدول ۵ خلاصه‌ای از شبکه به ازای مقدار نگاه به عقب ۷

**Table 5** Summary of Model Structure for Value of Look Back 7

Layer (type)	Output Shape (size)	Param (No.)
lstm-1029 (LSTM)	(None, 7, 512)	1073152
dropout_1029 (Dropout)	(None, 7, 512)	0
Lstm-1030 (LSTM)	(None, 7, 256)	787456
dropout_1030 (Dropout)	(None, 7, 256)	0
Lstm-1031 (LSTM)	(None, 256)	525312
dropout_1031 (Dropout)	(None, 256)	0
Dense-274 (Dense)	(None, 3)	771

شکل ۸ نمودار تابع هزینه (زیان) برحسب تکرار دوره را نشان می‌دهد.



شکل ۸ نمودار تابع هزینه بر حسب تکرار دوره به ازای مقدار نگاه به عقب ۷  
**Figure 8** Loss Function Diagram for Value of Look Back 7

جدول ۶، نتایج شبکه را به ازای مقادیر بهینه ابر متغیرها بر روی داده‌های آموزش و آزمایش نشان می‌دهد.

#### ۴ نتیجه‌گیری

ابرمغیرهای مهم در شبکه مذکور تعیین مقدار بهینه نگاه به عقب است که حالت‌های مختلف آن بررسی شده است، اما در داده‌های نمودارگیری از گل حفاری چاه‌های موجود در یکی از میداین نفتی ایران آسیب‌های سیلان چاه و گیر رشته حفاری به‌ندرت رخ داده است که همین امر موجب عدم توازن در طبقه‌بندی و کاهش میزان دقت و عملکرد شبکه خواهد شد. برای رفع آن از الگوریتم نمونه‌گیری دسته اقلیت استفاده شده است. همچنین برای تعیین مقدار بهینه نگاه به عقب نیز از الگوریتم جستجوی شبکه‌ای با اعتبارسنجی متقاطع استفاده شده که با مشاهده و بررسی نتایج فوق، به‌ازای مقدار نگاه به عقب ۷، شبکه مذکور هشدار کاذب کم و عملکرد قابل‌قبولی را نسبت به مابقی حالت‌ها دارد و میزان دقت به‌دست آمده از آن برابر ۹۴ درصد است.

عملیات حفاری در صنعت نفت بسیار با اهمیت بوده و همواره دارای ریسک‌های عملیاتی است و در صورت عدم کنترل دقیق، می‌تواند با آسیب‌ها و خطرهای جدی روبه‌رو شود. تاکنون مشکلات حین حفاری به‌صورت کامل حل نشده است. به‌همین دلیل محققان طی سالیان اخیر با به‌کارگیری مدل‌های هوش مصنوعی دنبال پیش‌بینی سریع نوع حادثه، قبل از وقوع آن بوده‌اند. شبکه‌های عصبی حافظه کوتاه مدت طولانی با گذشت زمان، اطلاعات مفید از زمان‌های قبلی را حفظ کرده و در پیش‌بینی خروجی زمان فعلی از آن استفاده می‌کنند و در نهایت با لحاظ شبکه عصبی تمام‌متصل در لایه آخر آسیب، داده مربوط شناسایی و در آن طبقه‌بندی خواهد شد. یکی از

#### مراجع

- Hosseini P., Estimation of Lost Circulation Amount Occurs During Under Balanced Drilling Using Drilling Data and Neural Network, *Egyptian Journal of Petroleum*, 26, 627–634, 2017.
- Abbas A.K., Al-haideri N.A. and Bashikh A.A., Implementing Artificial Neural Networks and Support Vector Machines to Predict Lost Circulation, *Egyptian Journal of Petroleum*, 28, 339–347, 2019.
- Hou X., Yang J., Yin Q., Liu H., Chen H., Zheng J., Wang J., Cao B., Zhao X. and Hao M., Lost Circulation Prediction in South China Sea Using Machine Learning and Big Data Technology, *Presented at the Offshore Technology Conference*, 2020.
- Hou X., Yang J., Yin Q., Chen L., Cao B., Xu J., Meng L., Zhang Y., Xu D. and Zhao X., Automatic Gas Influxes Detection in Offshore Drilling Based on Machine Learning Technology, *Presented at the SPE Gas & Oil Technology Showcase and Conference*, 2019.
- Muojeke S., Venkatesan R. and Khan F., Supervised Data-Driven Approach to Early Kick Detection During Drilling Operation, *Journal of Petroleum Science and Engineering*, 192, 107324, 2020.
- Nhat D.M., Venkatesan R. and Khan F., Data-Driven Bayesian Network Model for Early Kick Detection in Industrial Drilling Process, *Process Safety and Environmental Protection*, 138, 130–138, 2020.
- Magana-Mora A., Gharbi S., Alshaikh A. and Al-Yami A., AccuPipePred: A framework for the Accurate and Early Detection of Stuck Pipe for Real-Time Drilling Operations, *Presented at the SPE Middle East Oil and Gas Show and Conference*, 2019.
- Alshaikh A., Magana-Mora A., Gharbi S.A. and Al-Yami A., Machine Learning for Detecting Stuck Pipe Incidents: Data Analytics and Models Evaluation, *Presented at the International petroleum technology conference*, 2019.
- Li Y., Cao W., Hu W. and Wu M., Diagnosis of Downhole Incidents for Geological Drilling Processes Using Multi-Time Scale Feature Extraction and Probabilistic Neural Networks, *Process Safety and Environmental Protection*, 137, 106–115, 2020.
- Zhang Z., Lai X., Wu M., Chen L., Lu C. and Du S., Fault Diagnosis Based on Feature Clustering of Time Series Data for Loss and Kick of Drilling Process, *Journal of Process Control*, 102, 24–33, 2021.
- Hairani H., Anggrawan A. and Priyanto D., Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link, *JOIV: International Journal on Informatics Visualization*, 7, 258–264, 2023.
- Passos D., Mishra P., A Tutorial on Automatic Hyperparameter Tuning of Deep Spectral Modelling for Regression and Classification Tasks, *Chemometrics and Intelligent Laboratory Systems*, 223, 104520, 2022.
- Wang J., Ozbayoglu E.M., Application of Recurrent Neural Network Long Short-Term Memory Model on Early Kick Detection, *Presented at the International Conference on Offshore Mechanics and Arctic Engineering*, 2022.