

Prediction of the Stress–Strain Curve of Sn–40Pb–2.5Sb Alloy Using an LSTM Neural Network

Research Article

Seyyede Hosna Seyyedmousavi¹, Abolfazl Rezaee-Bazzaz²

1. Abstract

The mechanical reliability of electronic devices is heavily dependent on the integrity of solder joints, which undergo complex plastic deformation, creep, and fatigue under thermal and mechanical cycling. The Sn–40Pb–2.5Sb alloy is favored in these applications due to its metallurgical stability and improved mechanical strength, attributed to the addition of antimony.

2. Introduction and Problem Statement

Characterization of the stress–strain curve of such alloys is problematic. Classical phenomenological models, such as the Anand viscoplastic model, require a high volume of experimental testing for parameter calibration, and their accuracy often degrades under varying loading paths. The research identifies a critical research gap: the need for a robust, efficient, and predictive model capable of capturing nonlinear and rate-dependent behaviors without the computational and experimental burden associated with classical constitutive approaches. The motivation for this work lies in utilizing deep learning to create a surrogate model that can reduce the need for or complement traditional laboratory-based testing.

3. Experimental Procedure and Data Acquisition

The dataset used in this study consists of experimental stress–strain data obtained for the Sn–40Pb–2.5Sb alloy with a grain size of approximately 2.4 μm . The data were divided into training and testing sets as follows:

- **Training Set:** Data collected at a strain rate of 0.01 s^{-1} .
- **Evaluating Set:** Data collected at strain rates of 0.0066 s^{-1} , 0.0033 s^{-1} , and 0.0013 s^{-1} .

The raw experimental data underwent significant cleaning, where random noise and outliers were removed to ensure the stability of the neural network training process. The dataset was structured as time-series sequences, essential for the sequential learning capability of the LSTM model.

4. Methodology: LSTM Model Architecture

The core of this research is the development of an LSTM neural network implemented in MATLAB. LSTM networks were selected for their ability to mitigate the vanishing gradient problem, allowing the model to capture long-term dependencies in sequential data—a prerequisite for predicting path-dependent mechanical curves.

A LSTM neural network with the following architecture was designed to capture the nonlinear dynamics of stress-strain evolution:

- **Input Layer:** 7 neurons (representing a 7-step lag).
- **Hidden Layers:** A dual-layer LSTM structure featuring 32 neurons in the first layer and 64 neurons in the second layer. This depth allows the network to abstract hierarchical features from the input data.
- **Output Layer:** A fully connected layer with a single neuron, predicting the next stress value in the sequence.

¹ MSc Student in Metallurgical and Materials Engineering, Ferdowsi University of Mashhad.

² Assistant Professor, Department of Metallurgical and Materials Engineering, Ferdowsi University of Mashhad.

The developed model employed the Adam optimization algorithm, which is well-suited for non-stationary objective functions, using the Mean Squared Error (MSE) as the primary loss function.

5. Preprocessing Strategies and Comparative Analysis

A significant contribution of this research is the systematic evaluation of two distinct data preprocessing protocols, highlighting that the quality of input data is as critical as the neural network architecture itself.

Method 1: Standard Normalization

This baseline method utilized min–max normalization ($[0, 1]$) based strictly on statistical parameters derived from the training data. While this approach is standard practice to prevent data leakage, it proved limited in managing the inherent fluctuations in the raw mechanical data, leading to moderate predictive performance.

Method 2: Weighted Time-Dependent Scaling and Filtering

The second method introduced a more sophisticated approach consisting of the following steps:

1. **Noise Reduction:** Application of a weighted moving-average filter to smooth the input sequences. This successfully eliminated nonphysical, random fluctuations while preserving the essential mechanical characteristics (e.g., yielding and hardening).
2. **Cumulative Scaling:** Rather than using static normalization, statistics were updated cumulatively. The final scaling parameters were calculated by combining training data statistics with historical test data statistics (weights of 0.8 and 0.2, respectively). This approach takes into account the temporal nature of the stress–strain path and ensures that the model remains grounded in the evolving mechanical context without violating causality.

6. Results and Discussion

The performance of the models was rigorously assessed using MAE (Mean Absolute Error), MSE, RMSE (Root Mean Squared Error), and the square of correlation factor (R^2).

The results clearly demonstrate the superiority of Method 2. Using Method 1, the model achieved R^2 values ranging from **0.90 to 0.92**, showing reasonable but incomplete alignment with experimental data; while by implementing Method 2, the model demonstrated a substantial performance boost. The main achievements of implementing method 2 of preprocessing are as follows:

- R^2 Increased to the range of **0.993–0.996**.
- A significant decrease was observed in MAE, MSE, and RMSE.
- At a strain rate of 0.0033 s^{-1} , the model achieved an R^2 of 0.9959, demonstrating near-perfect correlation with experimental curves.

Considerable improvement in accuracy confirms that the success of an LSTM model is heavily dependent on feature engineering. The application of weighted moving-average filtering effectively removed noise, allowing the LSTM to better identify the subtle transitions in the stress–strain curve, particularly during the transition from elastic to plastic deformation. The cumulative scaling method successfully adapted the model to different strain rates without the need for retraining.

7. Conclusion

This study shows that LSTM neural networks provide a highly effective framework for modeling the mechanical response of Sn–40Pb–2.5Sb solder alloys. By exploiting the temporal memory inherent in the LSTM architecture, the model can accurately predict stress–strain curves, offering a viable, low-cost alternative to the exhaustive experimental and calibration procedures required by conventional constitutive models.

The success of the proposed preprocessing strategy—specifically the integration of filtering and time-dependent scaling—highlights the importance of data refinement in deep learning applications for materials science.

نسخه پیش انتشار

پیش بینی نمودار تنش - کرنش آلیاژ Sn-40Pb-2.5Sb با استفاده از شبکه عصبی LSTM

سیده حسنا سیدموسوی^۱، ابوالفضل رضائی بزازی*

^۱گروه مهندسی مواد و متالورژی، دانشکده مهندسی، دانشگاه فردوسی مشهد

*نویسنده مسئول: bazaz-r@um.ac.ir

چکیده

آلیاژهای لحیم سربدار همچنان در بسیاری از سامانه‌های الکترونیکی و اتصالات صنعتی، به‌ویژه در کاربردهایی که نیازمند پایداری حرارتی، رفتار مکانیکی قابل اعتماد و قابلیت فرآوری مناسبی هستند؛ مورد توجه قرار می‌گیرند. در این میان، آلیاژ Sn-40Pb-2.5Sb به دلیل حضور عنصر آنتیموان و تأثیر آن بر استحکام مکانیکی، ریزساختار و مقاومت به خزش از اهمیت ویژه‌ای برخوردار است. با این حال، رفتار تنش-کرنش این آلیاژ تحت شرایط بارگذاری مختلف ماهیتی غیرخطی، وابسته به مسیر و متأثر از ویژگی‌های متالورژیکی دارد و مدل‌سازی دقیق آن با استفاده از روابط تحلیلی کلاسیک همواره با محدودیت‌هایی همراه است. در پژوهش حاضر، برای پیش‌بینی نمودار تنش-کرنش آلیاژ Sn-40Pb-2.5Sb از یک رویکرد داده‌محور مبتنی بر شبکه عصبی بازگشتی حافظه کوتاه مدت طولانی (Long Short-Term Memory) استفاده می‌شود. بدین منظور، داده‌های تجربی حاصل از آزمون‌های مکانیکی پس از پیش‌پردازش، نرمال‌سازی و سازمان‌دهی به‌صورت دنباله‌های زمانی به شبکه LSTM اعمال می‌شوند. ساختار مدل به‌گونه‌ای طراحی می‌شود که بتواند وابستگی بین نقاط متوالی منحنی تنش-کرنش را فراگرفته و روند تکامل رفتار مکانیکی ماده را بازسازی کند. آموزش شبکه با استفاده از الگوریتم Adam انجام می‌شود و عملکرد مدل با بهره‌گیری از شاخص‌هایی نظیر میانگین خطای مطلق (Mean Absolute Error)، میانگین مجذور خطا (Mean Squared Error) و ضریب تعیین R^2 (Coefficient of Determination) ارزیابی می‌گردد. نتایج نشان می‌دهد که مدل پیشنهادی قادر است با دقتی مناسب، روند کلی و جزئیات منحنی تنش-کرنش را بازتولید کند و هم‌خوانی مطلوبی با داده‌های تجربی داشته باشد. یافته‌های این تحقیق نشان می‌دهد که استفاده از شبکه عصبی LSTM می‌تواند به‌عنوان ابزاری کارآمد برای مدل‌سازی رفتار مکانیکی آلیاژهای لحیم، در جهت کاهش وابستگی به آزمون‌های پرهزینه و توسعه روش‌های هوشمندسازی در مهندسی مواد مورد استفاده قرار گیرد.

واژگان کلیدی: آلیاژ Sn-40Pb-2.5Sb، نمودار تنش-کرنش، شبکه عصبی LSTM، یادگیری عمیق، مدل‌سازی داده‌محور، رفتار مکانیکی.

Abstract

Lead-based solder alloys are still widely used in many electronic systems and industrial joining applications, particularly where thermal stability, reliable mechanical performance, and suitable processability are required. Among these alloys, Sn-40Pb-2.5Sb is of particular interest due to the role of antimony in improving mechanical strength, microstructural stability, and creep resistance. However, the stress-strain behavior of this alloy under mechanical loading is inherently nonlinear, path-dependent, and strongly influenced by metallurgical features, making its accurate prediction by conventional constitutive models challenging. In the present study, a data-driven approach based on a Long Short-Term Memory (LSTM) neural network was employed to predict the stress-strain curve of the Sn-40Pb-2.5Sb alloy. Experimental mechanical test data were first preprocessed, normalized, and organized as sequential input data for the LSTM model. The network architecture was designed to capture the dependency between successive points along the stress-strain curve and to reconstruct the evolution of the alloy's

mechanical response. The model was trained using the Adam optimization algorithm. Its predictive performance was evaluated using statistical criteria such as R^2 , MAE, and MSE. The results demonstrated that the proposed model can accurately reproduce both the overall trend and the detailed features of the experimental stress-strain curves, showing good agreement with the measured data. The findings of this study indicate that LSTM-based modeling is a promising and efficient tool for predicting the mechanical behavior of solder alloys, reducing dependence on costly experimental testing, and advancing intelligent data-driven approaches in materials engineering.

Keywords: Sn-40Pb-2.5Sb alloy, stress-strain curve, LSTM neural network, deep learning, data-driven modeling, mechanical behavior.

۱. مقدمه

در دهه‌های اخیر، آلیاژهای لحیم پایه قلع نقش بسیار مهمی در صنایع الکترونیک و فناوری‌های بسته‌بندی الکترونیکی ایفا کرده‌اند. این آلیاژها به عنوان ماده اتصال‌دهنده میان اجزای مختلف مدارهای الکترونیکی مورد استفاده قرار می‌گیرند و علاوه بر ایجاد اتصال الکتریکی، وظیفه انتقال حرارت و تحمل تنش‌های مکانیکی و حرارتی را نیز بر عهده دارند. در میان انواع آلیاژهای لحیم، آلیاژهای قلع-سرب به دلیل ویژگی‌هایی نظیر نقطه ذوب مناسب، قابلیت ترشوندگی مطلوب و پایداری متالورژیکی، برای مدت زمان طولانی در صنعت الکترونیک مورد استفاده قرار گرفته‌اند [۱]. بررسی‌های انجام‌شده نشان می‌دهد که این آلیاژها از نظر خواص متالورژیکی و رفتار مکانیکی، عملکرد قابل قبولی در بسیاری از کاربردهای بسته‌بندی الکترونیکی دارند [۲]. علاوه بر این، تحقیقات گسترده‌ای درباره ساختار، ریزساختار و خواص مکانیکی این آلیاژها انجام شده‌است که آن‌ها را به یکی از سیستم‌های مرجع در مطالعه رفتار مکانیکی آلیاژهای لحیم تبدیل کرده‌است [۳].

رفتار مکانیکی آلیاژهای لحیم تأثیر مستقیمی بر قابلیت اطمینان و دوام اتصالات الکترونیکی دارد. در بسیاری از کاربردهای عملی، اتصالات لحیمی تحت بارگذاری‌های مکانیکی و سیکل‌های حرارتی قرار می‌گیرند که این شرایط می‌تواند منجر به تغییر شکل پلاستیک، خزش و در نهایت شکست اتصال شود. به همین دلیل شناخت دقیق منحنی تنش- کرنش این آلیاژها برای پیش‌بینی عملکرد و افزایش خواص مکانیکی اتصالات الکترونیکی از اهمیت بالایی برخوردار است. مطالعات تجربی نشان داده‌اند که آلیاژهای Sn-Pb رفتار خزشی قابل توجهی از خود نشان می‌دهند که به شدت به دما و شرایط بارگذاری وابسته است [۴]. همچنین بررسی‌های انجام‌شده درباره رشد ترک در اتصالات لحیمی نشان می‌دهد که رفتار مکانیکی ماده نقش تعیین‌کننده‌ای در عمر خستگی این اتصالات دارد [۵]. علاوه بر این، تحلیل خواص مکانیکی لحیم‌ها نشان می‌دهد که تغییرات ریزساختاری می‌تواند تأثیر قابل توجهی بر استحکام و مقاومت به تغییر شکل این مواد داشته باشد [۶].

به منظور بهبود خواص مکانیکی و افزایش مقاومت در برابر خزش، استفاده از عناصر آلیاژی مختلف در ترکیب آلیاژهای قلع-سرب مورد توجه پژوهشگران قرار گرفته‌است. یکی از عناصر آلیاژی مهم در این زمینه، آنتیموان است که می‌تواند با ایجاد تغییر در ریزساختار آلیاژ، موجب افزایش استحکام و بهبود پایداری مکانیکی آن شود. حضور آنتیموان در آلیاژهای قلع-سرب باعث تغییر در مورفولوژی فازها و تشکیل ترکیبات بین‌فلزی می‌شود که در نهایت رفتار مکانیکی آلیاژ را تحت تأثیر قرار می‌دهد [۷]. همچنین گزارش شده‌است که افزودن عناصر آلیاژی می‌تواند موجب بهبود مقاومت خزشی و افزایش استحکام اتصالات لحیمی شود. از این رو آلیاژ $\text{Sn} \square 40\text{Pb} \square 2.5\text{Sb}$ به عنوان یکی از ترکیبات اصلاح‌شده آلیاژهای لحیم مورد توجه قرار گرفته‌است [۸].

[۹].

به طور سنتی، مدل سازی رفتار مکانیکی آلیاژهای لحیم با استفاده از مدل های پدیدارشناختی و روابط ساختاری انجام می شود. یکی از مدل های شناخته شده در این زمینه، مدل ویسکوپلاستیک آند است که برای توصیف رفتار تغییر شکل مواد در دماهای بالا توسعه یافته است [۱۰]. با این حال تعیین پارامترهای این مدل ها معمولاً نیازمند انجام آزمایش های متعدد است و در شرایط بارگذاری پیچیده ممکن است دقت پیش بینی آن ها کاهش یابد. به همین دلیل پژوهشگران در سال های اخیر به دنبال استفاده از روش های نوین مدل سازی مبتنی بر داده برای توصیف رفتار مکانیکی مواد بوده اند [۱۱].

در سال های اخیر استفاده از روش های یادگیری ماشین و یادگیری عمیق به عنوان رویکردی نوین در مدل سازی رفتار مواد مورد توجه قرار گرفته است. این روش ها قادرند روابط پیچیده و غیرخطی میان متغیرهای مختلف را از داده های تجربی استخراج کنند و در نتیجه دقت پیش بینی رفتار مواد را افزایش دهند [۱۲]. در میان مدل های مختلف یادگیری عمیق، شبکه های عصبی بازگشتی و به ویژه شبکه حافظه کوتاه مدت طولانی به دلیل توانایی در یادگیری وابستگی های زمانی و داده های ترتیبی، برای مدل سازی فرآیندهای وابسته به مسیر مانند منحنی تنش- کرنش بسیار مناسب هستند [۱۳]. مطالعات اخیر نشان داده اند که روش های یادگیری عمیق می توانند در پیش بینی رفتار پلاستیسیته وابسته به مسیر عملکرد موفقیت آمیزی داشته باشند [۱۴]. مبنای نظری بسیاری از این روش ها نیز در چارچوب یادگیری عمیق توسعه یافته است [۱۵].

یکی از پرکاربردترین حوزه ها در زمینه استفاده از شبکه عصبی LSTM، پیش بینی منحنی های تنش- کرنش است. در این راستا، پژوهش های متعددی با استفاده از این معماری و با در نظر گرفتن متغیرهایی نظیر شرایط بارگذاری، کرنش سختی، تغییر شکل الاستیک، دما و ترکیب شیمیایی، موفق به پیش بینی دقیق رفتار کششی ورقه های آلومینیوم شده اند. به طور مشخص، در بررسی آلیاژ آلومینیوم ۶۰۱۶ در دماها و نرخ کرنش های متفاوت، با بهره گیری از پارامترهایی نظیر ترکیب شیمیایی، اندازه دانه و مقدار کرنش، توانایی مدل در بازنمایی منحنی تنش- کرنش با دقت بسیار بالای $R^2 = 0.9990$ به اثبات رسیده است [۱۶]. همچنین در حوزه سوپرآلیاژهای پایه نیکل، مدل سازی رفتار کششی آلیاژ FGH98 در دماها و نرخ کرنش های مختلف، منجر به دستیابی به دقت $R^2 = 0.951$ گردیده است [۱۷]. علاوه بر این، پیش بینی منحنی تنش- کرنش سوپرآلیاژ IN718 در دماهای بالا و نرخ کرنش های بسیار پایین ($0.005 s^{-1}$)، دقت بسیار بالای این شبکه را با مقدار $R^2 = 0.9925$ نشان داده است [۱۸].

علاوه بر رفتار کششی، قابلیت تعمیم این شبکه ها در پیش بینی پدیده های مخرب دیگر نیز به اثبات رسیده است. برای مثال، در بررسی رفتار خستگی سوپرآلیاژ پایه نیکل GH4169، مدل LSTM توانسته با لحاظ کردن پیچیدگی های هندسی و انواع بارگذاری، عمر خستگی را با دقت $R^2 = 0.991$ پیش بینی نماید [۱۹]. در حوزه رفتار خزشی نیز، مدل سازی منحنی های تنش- زمان برای فولادها با استفاده از متغیرهای ترمومکانیکی، منجر به دستیابی به خطای بسیار ناچیز ($RMSE = 0.000729$) شده است [۲۰]. در نهایت، پتانسیل این معماری در حوزه فرآیندهای ساخت نیز مشهود است؛ به طوری که در پیش بینی دمای حوضچه مذاب در فرآیند شکل دهی لیزری بر پایه پارامترهای فرآیندی و ویژگی های ماده، دقت $R^2 = 0.925$ حاصل شده است [۲۱]. مجموع این یافته ها، نشان دهنده تسلط شبکه های LSTM بر طیف گسترده ای از تحلیل های مهندسی، از مرحله تولید و شکل دهی تا ارزیابی عملکرد مکانیکی در شرایط عملیاتی است.

با وجود مطالعات متعدد در زمینه رفتار مکانیکی آلیاژهای لحیم، استفاده از روش های یادگیری عمیق به ویژه شبکه های LSTM برای پیش بینی نمودار تنش- کرنش آلیاژ Sn-40Pb-2.5Sb هنوز به طور گسترده مورد بررسی قرار نگرفته است. از این رو با توجه به اهمیت شناخت رفتار مکانیکی آلیاژهای لحیم و قابلیت روش های یادگیری عمیق در مدل سازی روابط پیچیده،

در این پژوهش تلاش شده است تا با استفاده از شبکه عصبی LSTM نمودار تنش- کرنش آلیاژ Sn-40Pb-2.5Sb پیش‌بینی شود. در این راستا داده‌های تجربی تنش و کرنش به عنوان ورودی شبکه مورد استفاده قرار گرفته و عملکرد مدل در بازسازی و پیش‌بینی رفتار مکانیکی آلیاژ مورد ارزیابی قرار گرفته است. نتایج این تحقیق می‌تواند گامی در جهت توسعه روش‌های داده‌محور برای تحلیل رفتار مکانیکی آلیاژهای لحیم و بهبود پیش‌بینی عملکرد اتصالات الکترونیکی باشد.

۲. روش تحقیق

در این پژوهش به منظور پیش‌بینی رفتار تنش- کرنش آلیاژ لحیم Sn-40Pb-2.5Sb از یک رویکرد داده‌محور مبتنی بر شبکه عصبی حافظه کوتاه مدت طولانی استفاده شده است. شبکه‌های LSTM یکی از انواع شبکه‌های عصبی بازگشتی هستند که توانایی یادگیری وابستگی‌های بلندمدت در داده‌های ترتیبی را دارند و به همین دلیل برای مدل‌سازی فرآیندهای وابسته به م سیر، مانند رفتار تنش- کرنش مواد مناسب هستند [۱۳]، [۲۲]. در این تحقیق؛ فرآیند کلی مدل سازی شامل مراحل جمع‌آوری داده‌ها، پیش‌پردازش داده‌ها، طراحی ساختار شبکه عصبی، آموزش مدل و در نهایت ارزیابی عملکرد مدل پیش‌بینی شده با استفاده از داده‌های آزمون می‌باشد.

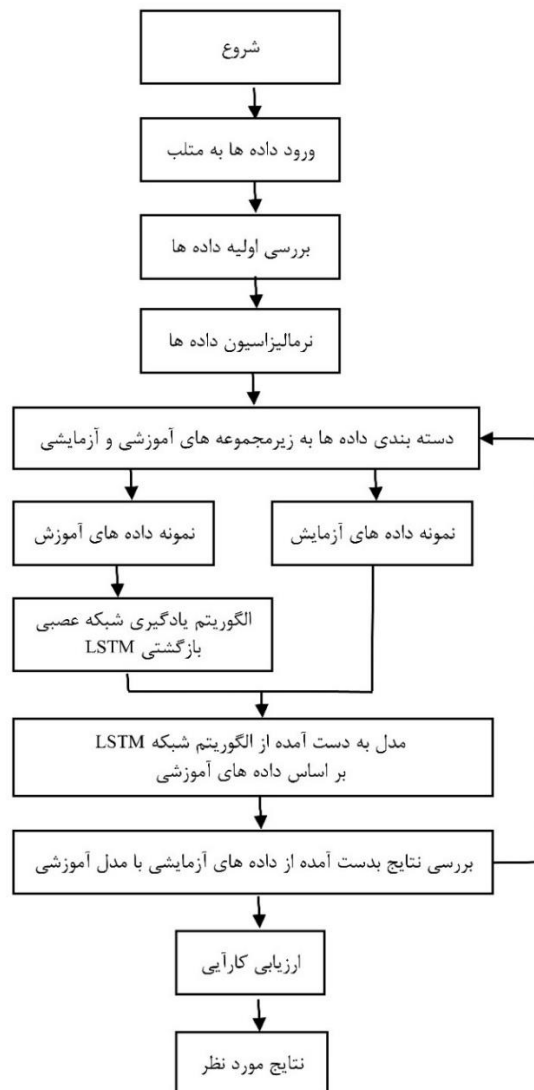
۱.۲. تهیه داده‌های تجربی

برای تحلیل رفتار مکانیکی آلیاژ Sn-40Pb-2.5Sb از داده‌های حاصل از آزمون کشش تک‌محوره که تحت سرعت فک ثابت در محدوده $0.5 - 1.5 \text{ mm/min}$ انجام شده است؛ استفاده می‌شود [۹]. آزمون کشش یکی از مهم‌ترین و متداول‌ترین آزمایش‌ها برای تعیین خواص مکانیکی مواد است و اطلاعات مهمی از جمله مدول الاستیسیته، تنش تسلیم، استحکام کششی نهایی و رفتار تغییر شکل پلاستیک را ارائه می‌دهد [۴]، [۲۳]. در این آزمایش نمونه‌های استاندارد تحت بارگذاری کششی یکنواخت در دمای 370°C درجه کلین قرار می‌گیرند و مقادیر تنش و کرنش تا زمان شکست ثبت می‌شود [۹].

رفتار مکانیکی آلیاژهای لحیم به دلیل دمای ذوب پایین و ساختار میکروسکوپی خاص، وابستگی زیادی به شرایط دما و نرخ کرنش دارد [۲۴]. به همین دلیل در مطالعات مختلف برای توصیف دقیق رفتار این مواد از مجموعه داده‌های دقیق تنش- کرنش استفاده می‌شود. در این پژوهش داده‌های مربوط به تنش و کرنش به صورت نقاط پیوسته از منحنی تنش- کرنش استخراج شده و به عنوان مجموعه داده آموزشی برای شبکه عصبی مورد استفاده قرار می‌گیرند [۹]، [۱۳].

۲.۲. پیش‌پردازش داده‌ها

پیش‌پردازش داده‌ها یکی از مراحل مهم در استفاده از الگوریتم‌های یادگیری ماشین محسوب می‌شود. در این مرحله داده‌ها به گونه‌ای آماده‌سازی می‌شوند که فرآیند آموزش شبکه با دقت و پایداری بیشتری انجام‌شود [۲۵]. شکل ۱، پیاده‌سازی روش‌های یادگیری ماشین را نشان می‌دهد.



شکل ۱. فلوچارت پیاده سازی روش های یادگیری ماشین.

در این پژوهش ابتدا داده های تنش و کرنش مرتب سازی شده و داده های غیرعادی یا تصادف حذف شدند. سپس برای جلوگیری از اثر مقیاس داده ها بر فرآیند یادگیری، داده ها با استفاده از روش نرمال سازی مینیمم-ماکسیمم در بازه صفر تا یک تبدیل شدند. رابطه نرمال سازی به صورت زیر تعریف می شود:

$$z = \frac{x - \min(x)}{[\max(x) - \min(x)]} \quad (1)$$

که در آن x مقدار اولیه داده، $\min(x)$ و $\max(x)$ به ترتیب حداقل و حداکثر مقدار داده در مجموعه داده ها هستند. نرمال سازی داده ها موجب افزایش سرعت همگرایی شبکه و جلوگیری از ناپایداری در فرآیند آموزش می شود [۲۶].

۳-۲. طراحی شبکه عصبی LSTM

شبکه‌های LSTM نوعی از شبکه‌های عصبی بازگشتی هستند که برای حل مشکل ناپدید شدن گرادینان در شبکه‌های عصبی بازگشتی کلاسیک، توسعه داده شده‌اند. این شبکه‌ها با استفاده از ساختار حافظه و دریچه‌های کنترلی قادرند، اطلاعات مهم را در طول توالی داده‌ها حفظ کنند [۱۵]، [۲۷].

شکل ۲، معماری شبکه عصبی LSTM را نشان می‌دهد. معماری شبکه عصبی LSTM، با هدف بهبود توانایی یادگیری وابستگی‌های زمانی در داده‌های سری زمانی طراحی شده‌است. برخلاف ساختارهای ساده، LSTM از یک ساختار سلولی استفاده می‌کند که اجازه می‌دهد اطلاعات با دقت بیشتری در طول زمان حفظ یا حذف شوند. این ساختار باعث می‌شود شبکه LSTM، بتواند وابستگی‌های بلندمدت موجود در داده‌های ترتیبی را به خوبی یاد بگیرد.

فرآیند مدیریت اطلاعات، توسط سه دریچه کنترلی هدایت می‌شود. این سه دریچه عبارتند از دریچه فراموشی، دریچه ورودی و دریچه خروجی. در ابتدا، دریچه فراموشی با بررسی ورودی‌های موجود، تعیین می‌کند که کدام بخش از حافظه سلول باید پاک شود تا از انباشت اطلاعات غیرضروری جلوگیری گردد. سپس، دریچه ورودی با شناسایی اطلاعات مهم در گام زمانی فعلی، فرآیند به‌روزرسانی حافظه سلول حالت را انجام می‌دهد. در نهایت، دریچه خروجی با فیلتر کردن اطلاعات موجود در سلول حالت، خروجی نهایی شبکه را برای گام بعدی تولید می‌کند. این مکانیزم کنترلی باعث می‌شود که شبکه بتواند اطلاعات مهم را برای مدت طولانی حفظ کرده و از نوسانات تصادفی جلوگیری کند. جزئیات ریاضیاتی و نحوه تعامل این دریچه‌ها برای مدیریت جریان اطلاعات، در قالب روابط ۲، ۳، ۴ و ۵ تعریف می‌گردند [۲۸].

دریچه ورودی i_t میزان اطلاعات جدیدی را که باید وارد حافظه شود را تعیین می‌کند و نشان می‌دهد که هر بخش از ورودی جدید، چه وزنی برای ورود به حافظه دارد و خروجی آن عددی بین ۰ و ۱ می‌باشد (۱ یعنی کاملاً پذیرفته شود و ۰ یعنی نادیده گرفته شود) و همچنین از فرمول زیر پیروی می‌کند:

$$i_t = \sigma(w_{hi} h_{t-1} + w_{ii} x_t + b_{hi} + b_{ii}) \quad (2)$$

σ تابعی کران‌دار و مشتق‌پذیر است که به ازای کلیه مقادیر حقیقی قابل تعریف می‌باشد و به آن تابع سیگموئید گفته می‌شود. h_{t-1} وضعیت پنهان (خروجی) سلول در گام زمانی قبلی است و در واقع حافظه کوتاه‌مدت گام قبلی است که به گام فعلی منتقل شده‌است.

x_t داده‌های ورودی جدید در گام زمانی فعلی است (به عنوان مثال، مقدار تنش یا کرنش در لحظه t).

w_{hi} ، ماتریس وزن‌های مرتبط با وضعیت پنهان قبلی (h_{t-1}) برای دریچه ورودی است. این ماتریس تعیین می‌کند که چه مقدار از حافظه قبلی برای تصمیم‌گیری فعلی اهمیت دارد.

w_{ii} ، ماتریس وزن‌های مرتبط با ورودی جدید (x_t) برای دریچه ورودی است. این ماتریس تعیین می‌کند که ورودی جدید چه مقدار در تصمیم‌گیری برای فراموش کردن اطلاعات قبلی تأثیر دارد.

b_{hi} یا b_{ii} بردار بایاس دریچه ورودی است. این پارامتر به مدل اجازه می‌دهد تا با یک آستانه اولیه، حساسیت دریچه به فراموشی را تنظیم کند.

سپس مقدار حافظه سلول به صورت زیر به‌روزرسانی می‌شود:

$$C_t = f_t C_{t-1} + i_t C'_t \quad (3)$$

C_t ، همان حافظه بلندمدتی است که تمام اطلاعات انتخاب شده از گذشته تا لحظه فعلی در آن جمع شده و قرار است از زمان $t-1$ به زمان t منتقل شود.

C_{t-1} ، حافظه‌ی گام قبلی است.

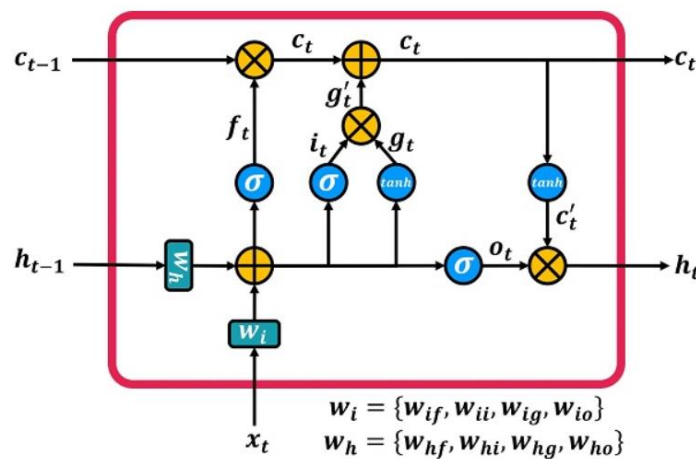
C'_t ، نیز حاوی اطلاعات جدید کاندید برای ورود به حافظه است (به عنوان مثال مقدار تنش در گام جدید).

و در نهایت دریچه خروجی، مقدار خروجی شبکه را مشخص می‌کند:

$$h_t = o_t \tanh(c_t) \quad (4)$$

h_t ، خروجی نهایی سلول در زمان t می‌باشد و همان حافظه کوتاه مدت است که مدل در نهایت برای پیش‌بینی مقدار تنش یا کرنش استفاده می‌کند.

o_t ، دریچه‌ای است که تعیین می‌کند چه بخشی از حافظه بلندمدت یعنی C_t قرار است در خروجی نهایی ظاهر شود و مانند یک فیلتر عمل می‌کند. مقدار عددی آن نیز بین ۰ و ۱ می‌باشد.



شکل ۲. معماری شبکه عصبی LSTM [۲۲].

دریچه فراموشی، حیاتی‌ترین بخش برای مدیریت حافظه به‌شمار می‌آید و تصمیم می‌گیرد چه بخشی از اطلاعات گذشته برای گام‌های زمانی آینده ضروری و چه بخشی قابل حذف است. شکل ۳، تصویر شماتیکی از دریچه فراموشی را نشان می‌دهد. رابطه این دریچه به صورت زیر بیان می‌شود:

$$f_t = \sigma(w_{hf} h_{t-1} + w_{if} x_t + b_{hf} + b_{if}) \quad (5)$$

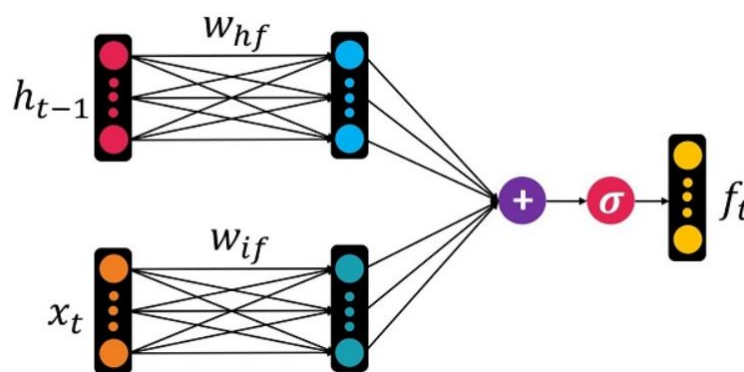
f_t ، خروجی این دریچه است که یک بردار می‌باشد نه یک عدد اسکالر و مقداری بین ۰ و ۱ دارد. بدین معنا که اگر خروجی نزدیک به ۰ باشد، یعنی اطلاعات قبلی باید کاملاً فراموش شود و اگر خروجی نزدیک به ۱ باشد، یعنی اطلاعات قبلی باید حفظ شود.

w_{hf} ، ماتریس وزن‌های مرتبط با وضعیت پنهان قبلی (h_{t-1}) برای دریچه فراموشی است. این ماتریس تعیین می‌کند که چه مقدار از حافظه قبلی برای تصمیم‌گیری فعلی اهمیت دارد.

w_{if} ، ماتریس وزن‌های مرتبط با ورودی جدید (x_t) برای دریچه فراموشی است. این ماتریس تعیین می‌کند که ورودی جدید چه مقدار در تصمیم‌گیری برای فراموش کردن اطلاعات قبلی تأثیر دارد.

b_{hf} یا b_{if} بردار بایاس دریچه فراموشی است. این پارامتر به مدل اجازه می‌دهد تا با یک آستانه اولیه، حساسیت دریچه به فراموشی را تنظیم کند.

تعاریف سایر پارامترهای دریچه فراموشی نیز، دقیقاً مشابه پارامترهای دریچه ورودی می‌باشند.



شکل ۳. تصویر شماتیکی از دریچه فراموشی f_t [۲۲].

۲-۴. آموزش شبکه عصبی LSTM

پس از طراحی ساختار شبکه LSTM، فرآیند آموزش شبکه با استفاده از داده‌های تجربی انجام شد. در این مرحله مجموعه داده‌ها به سه بخش شامل داده‌های آموزش، داده‌های اعتبارسنجی و داده‌های آزمون تقسیم شدند. معمولاً در بسیاری از مطالعات یادگیری ماشین حدود ۷۰ درصد داده‌ها برای آموزش، ۱۵ درصد برای اعتبارسنجی و ۱۵ درصد برای آزمون استفاده می‌شود [۲۹]. این تقسیم‌بندی موجب می‌شود تا عملکرد مدل به طور دقیق ارزیابی شده و از بیش‌برازش جلوگیری شود.

در فرآیند آموزش شبکه، پارامترهای مدل از طریق الگوریتم بهینه‌سازی به‌روزرسانی می‌شوند تا مقدار خطای پیش‌بینی به حداقل برسد. در این تحقیق از الگوریتم بهینه‌سازی آدام استفاده شد که یکی از پرکاربردترین روش‌های آموزش شبکه‌های عصبی عمیق محسوب می‌شود و سرعت همگرایی بالایی دارد [۳۰]. برای درک بهتر شبکه عصبی LSTM، مانند دیگر مدل‌های یادگیری در متلب راهنمای مفیدی ارائه گردیده است که با توجه به آن دستوراتی برای کار با شبکه عصبی LSTM استفاده می‌شود. ابتدا داده‌ها را به فرمت مناسب تبدیل کرده تا برای شبکه عصبی LSTM قابل استفاده باشد. برای این کار، می‌توان از توابع `seq2con` و `con2seq` استفاده کرد. در واقع این توابع، واسطه‌های پیش‌پردازش و پس‌پردازش داده‌ها هستند که برای تبدیل داده‌های خام آزمایشگاهی به فرمت قابل فهم برای LSTM، استفاده می‌شوند. در تحقیق حاضر، از تابع `seq2con` استفاده می‌شود. این تابع خروجی‌های گسسته را می‌گیرد، آن‌ها را مرتب کرده و هم‌پوشانی‌های احتمالی را اصلاح می‌کند و به صورت یک سری زمانی یا یک نمودار پیوسته به ما تحویل می‌دهد.

پس از اجرای توابع مورد نظر، مدل آموزش می‌بیند. نحوه‌ی کار به این صورت است که برای پیش‌بینی داده‌ی بعدی، به منظور بهبود عملکرد شبکه، انحراف معیار و میانگین داده‌های قبل مورد استفاده قرار می‌گیرند (منظور از انحراف معیار و میانگین مورد استفاده، همان نرمال‌سازی داده‌ها می‌باشد که بعد از دادن داده‌های تنش به ورودی شبکه، محاسبه شده و جهت آموزش به شبکه داده می‌شود). در این پژوهش مقدار لگ که برابر با ۷ تعریف می‌شود، در نتیجه برای پیش‌بینی داده‌ی هشتم از انحراف معیار و میانگین داده‌های اول تا هفتم استفاده می‌شود و همینطور به ترتیب برای داده‌های بعدی نیز این روند ادامه می‌یابد. به این صورت که برای پیش‌بینی داده‌ی نهم، از انحراف معیار و میانگین داده‌های دوم تا هشتم استفاده می‌شود. پس از آموزش شبکه، نتایج پیش‌بینی شده استخراج می‌شود و برای تحلیل و ترسیم نمودارهای تنش - کرنش استفاده می‌شود.

۲-۵. ارزیابی عملکرد مدل

پس از آموزش شبکه، عملکرد مدل با استفاده از داده‌های آزمون مورد ارزیابی قرار می‌گیرد. در واقع بخشی از داده‌های ورودی را که به عنوان داده‌های آزمایش در نظر گرفته می‌شود، مورد ارزیابی قرار داده تا دقت الگوریتم شبکه عصبی مورد ارزیابی قرار گیرد. در نتیجه، داده‌های آزمایش با خروجی مدل آموزشی مورد مقایسه قرار می‌گیرد.

در این پژوهش، به منظور ارزیابی جامع عملکرد شبکه عصبی در مواجهه با سناریوهای مختلف توزیع داده، دو روش متمایز برای پیش‌پردازش و مقیاس‌سازی داده‌ها پیاده‌سازی می‌شود. در روش اول، با رعایت اصول استاندارد یادگیری ماشین، فرآیند نرمال‌سازی صرفاً بر پایه آماره‌های توصیفی مجموعه داده‌های آموزشی (شامل میانگین و انحراف معیار) انجام می‌گیرد تا از هرگونه نشت اطلاعات از مجموعه آزمون به مدل جلوگیری شود و در روش دوم، یک پروتکل مقیاس‌سازی ترکیبی جهت بررسی حساسیت مدل به تغییرات توزیع آماری اتخاذ می‌گردد.

نکته حائز اهمیت در روش دوم این است که به منظور پیشگیری از هرگونه نشت داده و حفظ ماهیت زمانی مسئله، محاسبات آماری در هر گام زمانی، تنها بر پایه داده‌های مشاهده شده در گذشته تا لحظه فعلی انجام می‌شود. به عبارت دیگر، مقیاس‌گذاری در این روش به صورت تجمعی و مبتنی بر تاریخچه داده‌های موجود صورت می‌گیرد، به طوری که میانگین و انحراف معیار نهایی، از ترکیب وزنی آماره‌های داده‌های آموزش (با وزن ۰/۸) و آماره‌های داده‌های آزمون پیشین (با وزن ۰/۲) به دست می‌آید. این رویکرد، توانایی مدل را در مواجهه با داده‌های عملیاتی که دارای روند یا تغییر توزیع در طول زمان هستند، به دقت ارزیابی می‌کند [۳۱].

برای سنجش دقت پیش‌بینی از چند معیار آماری استفاده می‌شود که یکی از مهم‌ترین شاخص‌ها، ضریب تعیین R^2 است که میزان تطابق بین داده‌های واقعی و مقادیر پیش‌بینی شده را نشان می‌دهد. مقدار این شاخص بین صفر و یک قرار دارد و هرچه مقدار آن به یک نزدیک‌تر باشد، دقت مدل بیشتر است. این شاخص از رابطه زیر پیروی می‌کند [۳۲].

$$R^2 = 1 - \frac{\left(\text{واریانس غیرقابل توضیح} \right)}{\left(\text{واریانس کل} \right)} \quad (6)$$

تابع خطای مورد استفاده در شبکه عصبی، میانگین مربعات خطا است که به صورت زیر تعریف می‌شود:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

که در آن n تعداد کل داده‌ها، y_i مقدار واقعی تنش و $\hat{y}_i$ مقدار پیش‌بینی شده توسط شبکه عصبی است. کاهش مقدار این تابع نشان‌دهنده بهبود عملکرد مدل در پیش‌بینی داده‌ها است [۳۳].

مقدار خطای جذر میانگین مربعات نیز که به عنوان پارامتر دیگری از خطای شبکه به حساب می‌آید، به صورت زیر تعریف می‌شود [۳۳]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

همچنین از معیار میانگین قدر مطلق خطا برای ارزیابی عملکرد مدل استفاده می‌شود که به صورت زیر تعریف می‌شود [۳۳]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

علاوه بر این، نمودارهای تنش- کرنش پیش‌بینی‌شده توسط مدل با داده‌های تجربی مقایسه می‌شوند تا میزان توانایی شبکه در بازسازی رفتار مکانیکی آلیاژ بررسی شود. مقایسه بصری منحنی‌های پیش‌بینی شده و داده‌های واقعی یکی از روش‌های متداول در ارزیابی مدل‌های داده‌محور در مهندسی مواد است [۳۳]، [۳۴].

۲-۶. ابزار و محیط پیاده‌سازی

مدل پیشنهادی در محیط برنامه‌نویسی متلب پیاده‌سازی شد. این چارچوب‌ها به دلیل قابلیت بالا در توسعه مدل‌های یادگیری عمیق و امکان پیاده‌سازی شبکه‌های عصبی پیچیده، در بسیاری از تحقیقات علمی مورد استفاده قرار می‌گیرند [۳۵].

۲-۷. تشریح فرآیند پیاده‌سازی مدل پیش‌بینی توسط شبکه عصبی LSTM

در این روش، فرآیند پیاده‌سازی مدل مبتنی بر شبکه عصبی LSTM در دو فاز اصلی آماده‌سازی و اجرا انجام می‌شود. در مرحله نخست، داده‌های خام پس از استخراج تحت فرآیند نرمال‌سازی قرار گرفته تا با حذف مقیاس‌های نامتوازن، پایداری و سرعت همگرایی مدل در طول آموزش افزایش یابد. سپس با اعمال مفهوم پنجره زمانی، داده‌ها به صورت توالی‌های زمانی بازسازی شده و به دو مجموعه آموزشی و آزمونی تقسیم می‌گردند تا قابلیت تعمیم‌پذیری مدل بر روی داده‌های مشاهده‌نشده، تضمین گردد.

در مرحله دوم، معماری شبکه شامل لایه‌های ورودی، لایه حافظه کوتاه مدت طولانی (LSTM) برای استخراج ویژگی‌های زمانی و لایه‌های متصل و رگرسیون برای تولید خروجی عددی، طراحی و پیاده‌سازی می‌شود. فرآیند آموزش مدل با بهره‌گیری از یک بهینه‌ساز مدرن و با هدف کمینه کردن تابع هزینه بر روی مجموعه آموزشی انجام می‌گیرد. در نهایت، پس از اتمام آموزش، مقادیر پیش‌بینی شده بر روی داده‌های آزمون اعمال شده و با بازگرداندن آن‌ها به مقیاس اصلی، دقت مدل از طریق محاسبه معیارهای آماری خطا مانند RMSE مورد ارزیابی قرار می‌گیرد. شبه کد ۱، مراحل کلی از روش پیشنهادی پژوهش را نشان می‌دهد.

شبه کد ۱: مراحل کلی از روش پیشنهادی پژوهش

الگوریتم: آموزش و پیش‌بینی با شبکه عصبی LSTM

ورودی: داده‌های سری زمانی خام (X_raw)

خروجی: مقادیر پیش‌بینی شده (Y_pred)

۱. مرحله پیش پردازش داده‌ها:

بارگذاری داده‌ها از منبع مشخص؛

نرمال سازی داده‌ها برای مقیاس بندی (Normalization)؛

تعیین طول پنجره زمانی (Lag) جهت ایجاد توالی‌ها؛

تبدیل داده‌ها به مجموعه‌های متوالی (Sequences)؛

ساخت مجموعه آموزش: $[X_train, Y_train]$ ← از داده‌های آموزشی؛

ساخت مجموعه آزمون: $[X_test, Y_test]$ ← از داده‌های آزمون؛

۲. مرحله تعریف معماری شبکه:

تعریف لایه ورودی (Sequence Input Layer)؛

تعریف لایه حافظه طولانی کوتاه (LSTM Layer)؛

تعریف لایه کاملاً متصل (Fully Connected Layer)؛

تعریف لایه خروجی برای رگرسیون (Regression Layer)؛

۳. مرحله آموزش مدل:

تنظیم پارامترهای بهینه‌ساز (Optimizer) و تنظیمات آموزش؛

اجرای فرآیند آموزش (Training Process) با استفاده از $[X_train, Y_train]$ ؛

۴. مرحله پیش بینی و ارزیابی:

استفاده از مدل آموزش دیده برای پیش بینی مقادیر آزمون (Y_hat)؛

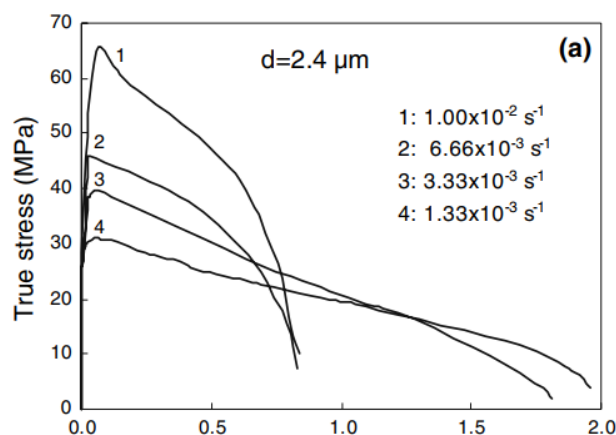
بازگرداندن مقادیر به مقیاس اصلی (Denormalization)؛

محاسبه معیار خطای مدل (Error Metric) بین Y_hat و Y_test ؛

۳. نتایج و بحث

۳-۱. نتایج آزمایش تجربی

در پژوهشی، رفتار سوپرپلاستیک آلیاژ Sn-40Pb-2.5Sb در نرخ کرنش‌های مختلف و با اندازه دانه ۲/۴ میکرومتر بررسی می‌شود [۹]. نتایج حاصل که به صورت منحنی تنش-کرنش درآمده در شکل ۴ مشاهده می‌شود.



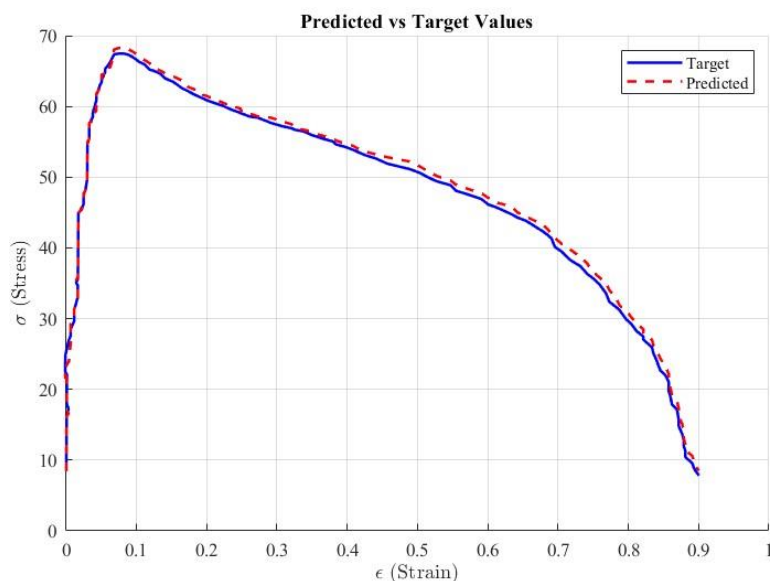
شکل ۴. منحنی تنش - کرنش آلیاژ Sn-40Pb-2.5Sb در نرخ کرنش‌های مختلف و با اندازه دانه ۲/۴ میکرومتر [۹].

۲-۳. نتایج مدل‌سازی شده

همان‌طور که گفته شد، مدل شبکه عصبی LSTM با بهره‌گیری از داده‌های موجود ابتدا آموزش یافته و سپس مورد ارزیابی قرار می‌گیرد. پس از آن داده‌های خروجی مدل تحلیل می‌شود و برای مقایسه بهتر میان نتایج تجربی و خروجی مدل‌سازی، این داده‌ها به صورت نمودار ارائه می‌گردد.

در این پژوهش، شبکه پیشنهادی شامل یک لایه ورودی با تعداد ۷ نورون، دو لایه LSTM (لایه اول، شامل ۳۲ نورون و لایه دوم شامل ۶۴ نورون می‌باشد) و یک لایه خروجی با یک نورون کاملاً متصل است. ورودی شبکه شامل مقادیر تنش و کرنش دوره‌های قبلی و خروجی شبکه مقدار تنش دوره بعدی می‌باشد. هدف شبکه نیز، یادگیری رابطه غیرخطی بین این دو متغیر و بازسازی منحنی تنش - کرنش آلیاژ مورد مطالعه است.

به همین منظور، داده‌های مربوط به نمونه آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.01 s^{-1} به عنوان داده‌های آموزش، به شبکه عصبی LSTM داده می‌شود. در واقع شبکه خودش از روی تفاوت کرنش‌ها در آزمایش‌های مختلف، متوجه نرخ کرنش و اثر آن بر تنش می‌شود. نتایج حاصل در قالب نمودار تنش - کرنش در شکل ۵ قابل مشاهده است [۹]. همان‌طور که مشاهده می‌شود، منحنی پیش‌بینی‌شده توسط شبکه عصبی LSTM در قسمت آموزش شبکه؛ انطباق قابل توجهی در تمامی مقادیر تنش و کرنش از خود نشان می‌دهد.



شکل ۵. مقایسه منحنی تنش- کرنش آلیاژ Sn-40Pb-2.5Sb مدل سازی شده توسط شبکه عصبی LSTM با اندازه گیری تجربی در نرخ کرنش 0.01 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده های پیش بینی شده توسط شبکه عصبی LSTM می باشد و منحنی ممتد آبی، داده های آزمایش را نشان می دهد).

۳-۳. بررسی میزان تعمیم پذیری مدل

جهت بررسی دقیق تر نتایج، از محاسبه ضریب تعیین کمک گرفته می شود که نتایج آن در جدول ۱ آورده شده است. اعداد ذکر شده نشان دهنده خطای بسیار کم و دقت بالای شبکه عصبی در قسمت آموزش می باشد.

جدول ۱. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.01 s^{-1}

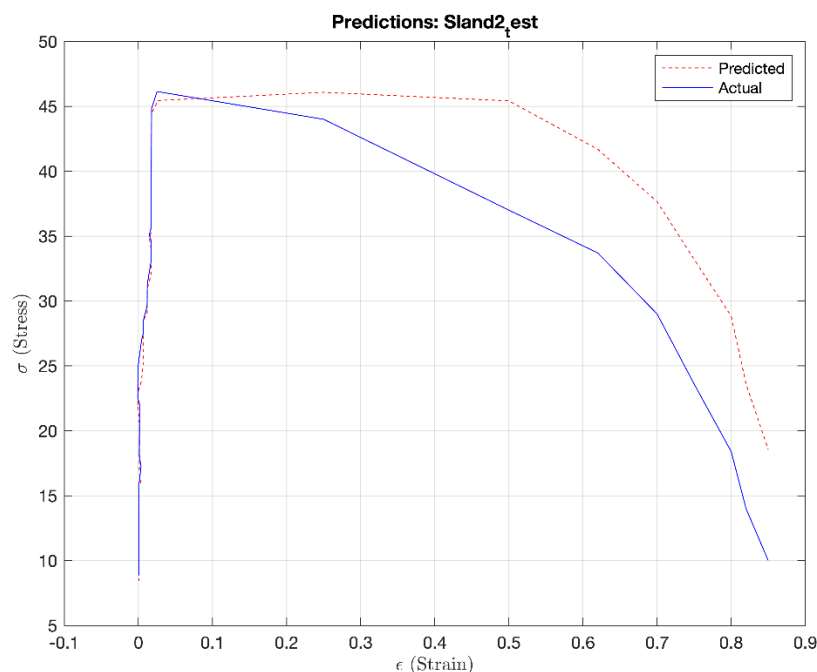
Regression Statistics	
Mean absolute Error	۱,۳۶۹,۰۵۱
Mean Square Error	۴,۵۴۰,۱۱۴
Root Mean Square Error	۲,۱۳۰,۷۵۴
R Square	۰,۹۸۵,۷۵۲
Adjusted R Square	۰,۹۸۵,۶۹۹

۳-۴. نتایج ارزیابی شبکه عصبی LSTM

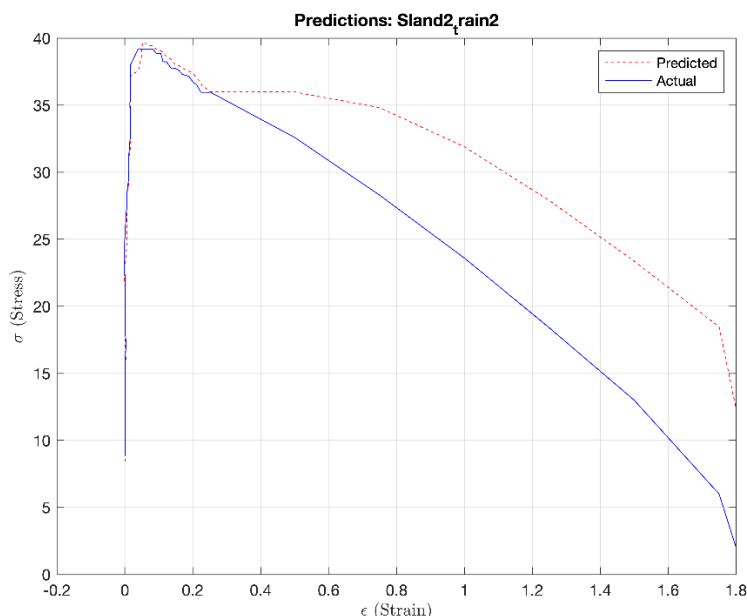
در این بخش، عملکرد شبکه عصبی LSTM در پیش بینی رفتار غیرخطی نمودار تنش- کرنش مورد بررسی قرار می گیرد. همچنین برای ارزیابی شبکه عصبی، داده های مربوط به نمونه آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش های 0.013 s^{-1} و 0.033 s^{-1} ، به عنوان داده های آزمون شبکه عصبی LSTM، مورد استفاده قرار می گیرند. این نتایج، حاصل عملکرد شبکه به روش اول می باشد که در شکل های ۶، ۷ و ۸ قابل مشاهده است. در شکل ۶ که منحنی تنش- کرنش در نرخ کرنش 0.066 s^{-1} را نشان می دهد، منحنی پیش بینی شده توسط شبکه عصبی تا مقادیر کرنش 0.025 s^{-1} تطابق بالایی با مقادیر اصلی منحنی دارد ولی از مقدار کرنش 0.025 s^{-1} به بعد خطای بالایی در پیش بینی منحنی به روش اول از خود نشان می دهد. شکل ۷ نیز منحنی تنش- کرنش در نرخ کرنش 0.033 s^{-1} را نشان می دهد که تطابق بالایی منحنی پیش بینی شده تا مقدار کرنش 0.025 s^{-1} به

وضوح دیده می‌شود. از مقدار کرنش 0.25 s^{-1} به بعد نیز از دقت شبکه عصبی کاسته می‌شود و تطابق با منحنی اصلی کاهش می‌یابد. شکل ۸ منحنی تنش-کرنش در نرخ کرنش 0.0013 s^{-1} می‌باشد که تا مقدار کرنش 0.3 s^{-1} ، منحنی پیش‌بینی شده بسیار نزدیک به منحنی اصلی می‌باشد و از مقدار کرنش 0.3 s^{-1} به بعد از دقت پیش‌بینی شبکه عصبی کاسته می‌شود و خطای شبکه به روش اول زیاد می‌گردد.

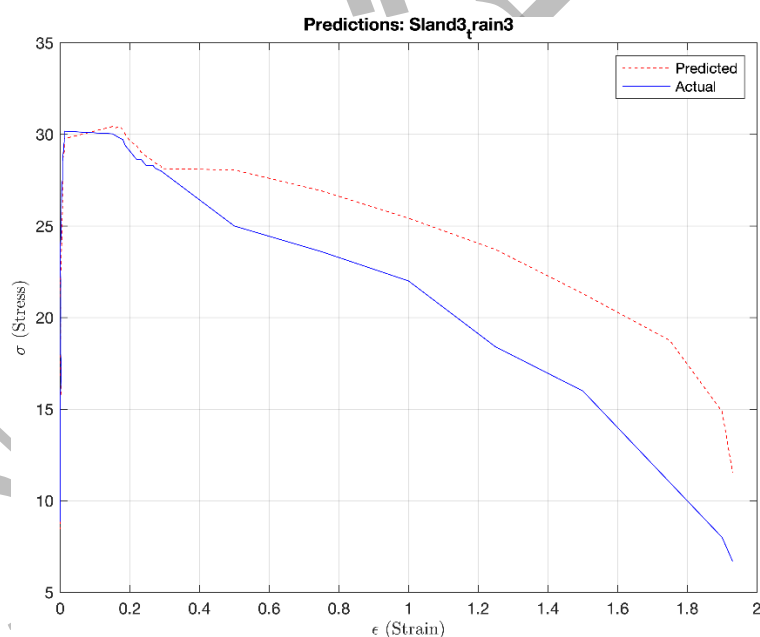
در حقیقت در این بخش، نتایج حاصل از روش اول مورد بررسی قرار می‌گیرد. داده‌های حاصل از این مدل نشان می‌دهد که اگرچه این روش قادر است روند کلی تغییرات تنش را دنبال کند، اما در انطباق با داده‌های تجربی به‌ویژه در نواحی تغییر شکل‌های غیرخطی و نقاط بحرانی، دچار خطای مشهودی است. این پدیده، عمدتاً ناشی از حساسیت بالای مدل به نوسانات تصادفی موجود در داده‌های خام تجربی است که باعث می‌شود شبکه عصبی نتواند روند هموار تغییرات را به‌طور کامل در این نقاط بازسازی کند.



شکل ۶. مقایسه منحنی تنش-کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش اول با اندازه‌گیری تجربی در نرخ کرنش 0.0066 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).



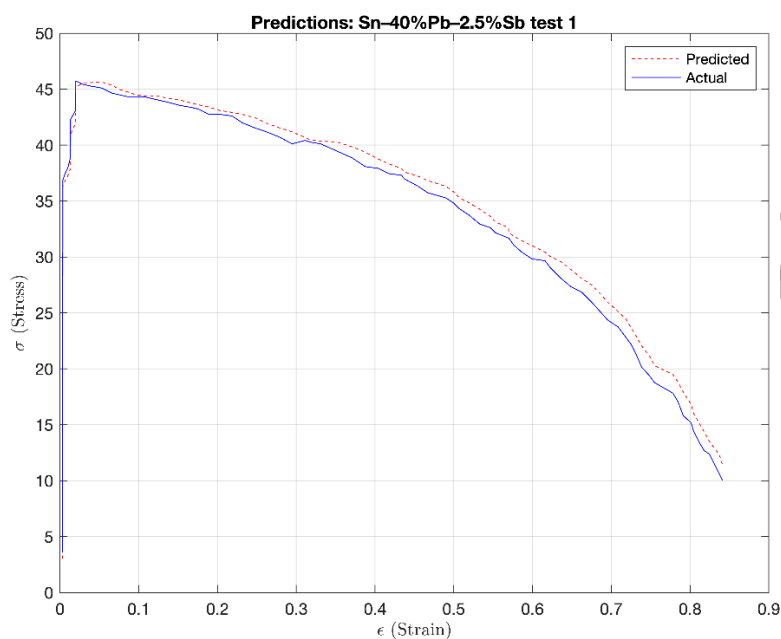
شکل ۷. مقایسه منحنی تنش - کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش اول با اندازه‌گیری تجربی در نرخ کرنش 0.0023 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).



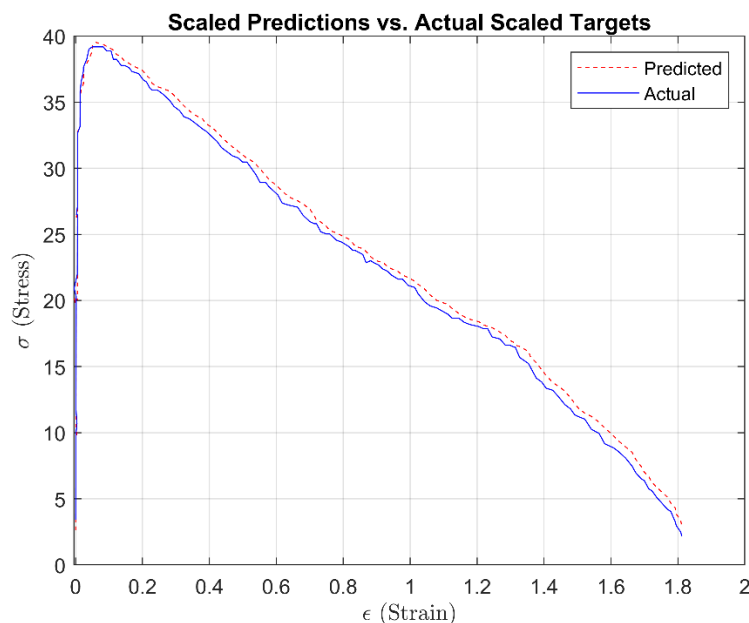
شکل ۸. مقایسه منحنی تنش - کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش اول با اندازه‌گیری تجربی در نرخ کرنش 0.0013 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).

شکل‌های ۹، ۱۰ و ۱۱، نتایج حاصل عملکرد شبکه به روش دوم می‌باشد. همان‌طور که مشاهده می‌گردد، نقاط پیش‌بینی شده در تمامی مقادیر کرنش بسیار نزدیک به مقادیر تجربی می‌باشند که این نتایج گویای این است که ارزیابی شبکه عصبی در این روش، با دقت بالایی صورت گرفته و مقادیر پیش‌بینی شده تطابق بالایی با مقادیر اصلی منحنی دارند. در واقع در این روش

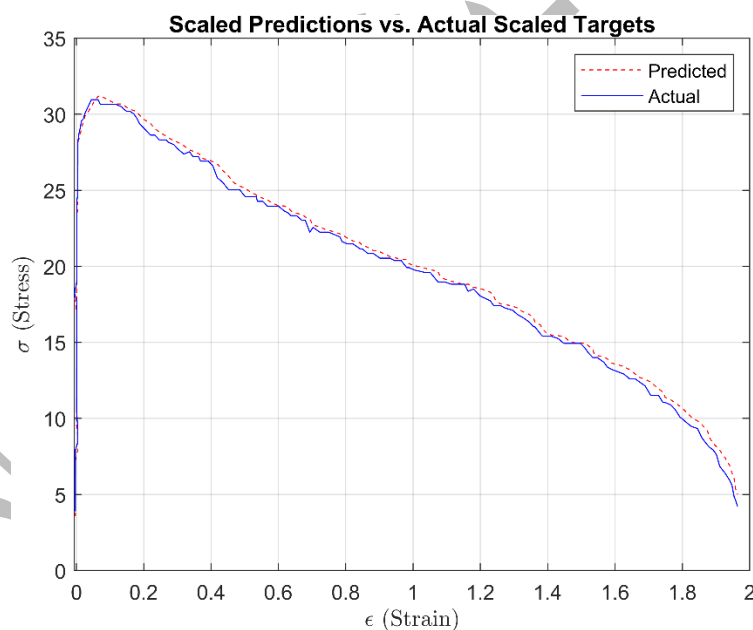
(روش دوم)، با هدف ارتقای دقت و بهبود پایداری پیش‌بینی، از فیلتر میانگین وزن‌دار متحرک بر روی داده‌های ورودی شبکه عصبی استفاده می‌گردد. نتایج نشان می‌دهد که این اصلاح، تفاوت چشمگیری در کیفیت بازسازی منحنی ایجاد می‌کند. اعمال میانگین وزن‌دار متحرک، باعث حذف نوسانات تصادفی اضافی و هموارسازی هوشمندانه روند داده‌ها شده و این امکان را برای لایه حافظه LSTM فراهم می‌آورد تا الگوهای اصلی حاکم بر منحنی تنش- کرنش را، با دقت بسیار بالاتری استخراج کند. در نمودارهای حاصل از این روش، انطباق منحنی پیش‌بینی‌شده با داده‌های تجربی در مقایسه با روش اول، بهبود قابل توجهی داشته و نرخ خطا در تمامی بازه‌های کرنش به شکل معناداری کاهش می‌یابد.



شکل ۹. مقایسه منحنی تنش- کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش دوم با اندازه‌گیری تجربی در نرخ کرنش 0.0066 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی‌شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).



شکل ۱۰. مقایسه منحنی تنش- کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش دوم با اندازه‌گیری تجربی در نرخ کرنش 0.0033 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).



شکل ۱۱. مقایسه منحنی تنش- کرنش آلیاژ Sn-40Pb-2.5Sb ارزیابی شده توسط شبکه عصبی LSTM به روش دوم با اندازه‌گیری تجربی در نرخ کرنش 0.0013 s^{-1} با اندازه دانه $2/4$ میکرومتر. (منحنی خط چین قرمز، داده‌های پیش‌بینی شده توسط شبکه عصبی LSTM می‌باشد و منحنی ممتد آبی، داده‌های واقعی را نشان می‌دهد).

در جداول ۲، ۳ و ۴، نتایج مربوط به مدل LSTM به روش اول ارائه می‌شود. یافته‌ها نشان می‌دهد که این مدل به‌طور قابل‌قبولی قادر به بازسازی روند کلی رفتار مکانیکی ماده و دنبال کردن منحنی تنش- کرنش است که نشان‌دهنده کارایی معماری شبکه عصبی LSTM در یادگیری الگوهای زمانی می‌باشد.

جدول ۲. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0066 s^{-1} به روش اول

Regression Statistics	
Mean absolute Error	۱,۷۶۴۳۰۹
Mean Square Error	۱۰,۱۱۳۴۴
Root Mean Square Error	۳,۱۸۰۱۶۴
R Square	۰,۹۲۰۸۳۸
Adjusted R Square	۰,۹۱۹۵۴۰

جدول ۳. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0033 s^{-1} به روش اول

Regression Statistics	
Mean absolute Error	۱,۵۳۹۵۳۰
Mean Square Error	۹,۰۲۸۱۴۶
Root Mean Square Error	۳,۰۰۴۶۸۷
R Square	۰,۹۱۸۴۲۵
Adjusted R Square	۰,۹۱۷۲۴۲

جدول ۴. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0013 s^{-1} به روش اول

Regression Statistics	
Mean absolute Error	۱,۴۷۳۶۵۰
Mean Square Error	۵,۱۵۴۹۵۶
Root Mean Square Error	۰,۵۰۱۳۵۵
R Square	۰,۹۰۲۱۰۹۵
Adjusted R Square	۰,۹۰۰۱۱۱۷

مقایسه شاخص‌های خطا برای روش دوم که مبتنی بر فیلتر میانگین وزن دار متحرک می‌باشد، در جداول ۵، ۶ و ۷ ارائه می‌شوند که نشان‌دهنده بهبود چشمگیر عملکرد مدل است. کاهش هم‌زمان مقادیر RMSE و MAE نسبت به مدل پایه، گویای دقت بالای این روش در بازسازی دقیق منحنی‌های مکانیکی است. حذف نوسانات ناخواسته از طریق پیش‌پردازش، منجر به تثبیت ورودی‌های شبکه می‌شود و در نتیجه، مدل می‌تواند با خطای بسیار کمتر، الگوهای حاکم بر رفتار ماده را در تمامی بازه‌های کرنش بازسازی کند.

جدول ۵. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0066 s^{-1} به روش دوم

Regression Statistics	
Mean absolute Error	۰,۸۸۵۷۲۸
Mean Square Error	۱,۰۴۳۳۶۴
Root Mean Square Error	۱,۰۲۱۴۵۲
R Square	۰,۹۹۲۸۴۶
Adjusted R Square	۰,۹۹۲۷۹۹

جدول ۶. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0033 s^{-1} به روش دوم

Regression Statistics	
Mean absolute Error	۰,۶۱۵۲۳۰
Mean Square Error	۰,۴۶۹۱۷۵
Root Mean Square Error	۰,۶۸۴۹۶۳
R Square	۰,۹۹۵۹۳۷
Adjusted R Square	۰,۹۹۵۹۱۵

جدول ۷. آمار رگرسیون برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0013 s^{-1} به روش دوم

Regression Statistics	
Mean absolute Error	۰,۴۲۴۹۴۰
Mean Square Error	۰,۲۵۱۳۵۶
Root Mean Square Error	۰,۵۰۱۳۵۵
R Square	۰,۹۹۵۲۹۴
Adjusted R Square	۰,۹۹۵۲۷۰

ارزیابی میزان دقت مدل‌ها، از طریق معیار خطای میانگین و ریشه میانگین مربعات خطا در جداول بالا ارائه شده است. مشاهده می‌شود که رویکرد مبتنی بر LSTM با فیلتر میانگین وزن دار متحرک نسبت به مدل استاندارد، کاهش چشمگیری در هر دو شاخص دارد. کاهش مقدار RMSE نیز، نشان‌دهنده کاهش شدت خطاهای بزرگ و نوسانات غیرمنطقی در منحنی‌های پیش‌بینی شده می‌باشد. این امر بیانگر آن است که حذف نوسانات تصادفی از داده‌های ورودی، می‌تواند پایداری مدل را در بازسازی نواحی حساس، نظیر محدوده تسلیم و تغییر شکل‌های غیرخطی به طور معناداری ارتقا دهد.

در نهایت با مقایسه این دو روش، نقش کلیدی پیش‌پردازش داده‌ها در عملکرد شبکه‌های عصبی LSTM به‌وضوح دیده می‌شود. اگرچه هر دو مدل از ساختار حافظه کوتاه مدت طولانی بهره می‌برند، اما روش دوم با بهره‌گیری از فیلتر میانگین وزن دار متحرک، توانسته بر ضعف روش اول در مواجهه با داده‌های دارای نوسانات تصادفی غلبه کند. تحلیل‌های آماری تأیید می‌کنند که روش دوم نه تنها نوسانات غیرفیزیکی مدل اول را حذف نموده، بلکه با تثبیت ویژگی‌های ورودی، پایداری و دقت پیش‌بینی را در شرایط مختلف بارگذاری به شکل مؤثری افزایش می‌دهد. همچنین، این مقایسه نشان می‌دهد که ادغام یک روش هوشمندانه در پیش‌پردازش داده‌ها با ساختار قدرتمند LSTM، راهکاری نوین و بسیار کارآمد برای دستیابی به مدل‌های پیش‌بینی با دقت بالا در تحلیل‌های مکانیکی است.

نتیجه‌گیری

در این پژوهش، پتانسیل مدل‌سازی داده‌محور با استفاده از شبکه عصبی حافظه کوتاه مدت طولانی LSTM برای پیش‌بینی نمودار تنش-کرنش آلیاژ لحیم Sn-40Pb-2.5Sb به دو روش مورد ارزیابی قرار می‌گیرد. با توجه به ماهیت غیرخطی بودن و نیز وابستگی به نرخ کرنش در رفتار مکانیکی لحیم‌ها، مدل پیشنهادی دوم با آموزش بر روی داده‌های تجربی، می‌تواند الگوهای

پیچیده تغییر شکل پلاستیک این آلیاژ را با دقت بالایی بازسازی کند. نتایج حاصل از این مطالعه منجر به استنتاج‌های کلیدی زیر می‌گردد:

۱. **کارآمدی مدل LSTM در داده‌های سری زمانی:** برخلاف مدل‌های کلاسیک تنش-کرنش که نیازمند فرضیات ساده‌ساز در فرمول‌بندی معادلات ساختاری هستند، شبکه عصبی LSTM با بهره‌گیری از دریچه‌های حافظه (دریچه ورودی، دریچه فراموشی و دریچه خروجی)، می‌تواند وابستگی زمانی داده‌های تنش-کرنش را به خوبی درک کرده و با کمترین خطای میانگین مربعات نمودارهای تجربی را پیش‌بینی کند.
 ۲. **تأثیر بهینه‌سازی پارامترها:** به کارگیری الگوریتم بهینه‌سازی آدام در کنار نرمال‌سازی داده‌ها با روش مینیمم-ماکسیمم، سرعت همگرایی شبکه را به طور چشمگیری افزایش می‌دهد و از بروز بیش‌برازش در داده‌های آموزشی جلوگیری می‌کند. مقادیر بالای ضریب تعیین R^2 در داده‌های آزمون به روش دوم که برای آلیاژ Sn-40Pb-2.5Sb با نرخ کرنش 0.0033 s^{-1} عدد 0.995937 گزارش شده، حاکی از تعمیم‌پذیری بالای مدل پیشنهادی می‌باشد.
 ۳. **پتانسیل جایگزینی روش‌های آزمایشگاهی:** مدل توسعه‌یافته نشان می‌دهد که می‌توان با کاهش تعداد آزمایش‌های مکانیکی پرهزینه و زمان‌بر، با تکیه بر یک پایگاه‌داده غنی و مدل هوشمند LSTM، رفتار مکانیکی آلیاژهای لحیم حاوی آنتیموان را برای شرایط بارگذاری مختلف پیش‌بینی کرد. این امر در طراحی قطعات الکترونیکی و همچنین تخمین عمر خستگی اتصالات لحیم‌شده، صرفه‌جویی قابل‌توجهی در هزینه‌های تحقیق و توسعه به همراه خواهد داشت.
 ۴. **توسعه‌پذیری تحقیق:** این پژوهش به عنوان گامی در جهت استفاده از یادگیری عمیق در مهندسی مواد محسوب می‌شود. در مطالعات آتی، پیشنهاد می‌شود اثر تغییرات دمایی و نرخ‌های کرنش بالاتر نیز به عنوان ورودی‌های شبکه عصبی لحاظ گردد تا مدل قادر به پیش‌بینی رفتار آلیاژ در شرایط عملیاتی سخت‌تر باشد.
- در نهایت، نتایج این مطالعه تأیید می‌کند که تلفیق دانش متالورژی لحیم‌های Sn-Pb-Sb با الگوریتم‌های پیشرفته یادگیری عمیق، رویکردی قدرتمند و دقیق برای تحلیل رفتارهای ساختاری مواد مهندسی دارد و می‌تواند به عنوان جایگزین یا مکملی برای مدل‌های کلاسیک پلاستیسیته مورد استفاده قرار گیرد.

واژه‌نامه

Sn-Pb Alloys	آلیاژهای قلع-سرب
Sb	آنتیموان
Anand	آند
Machine Learning	یادگیری ماشین
Deep Learning	یادگیری عمیق
Long Short-Term Memory	حافظه کوتاه مدت طولانی
Min- Max Normalization	نرمال‌سازی مینیمم-ماکسیمم
Recurrent Neural Network (RNN)	شبکه‌های عصبی بازگشتی
Forget Gate	دریچه فراموشی
Input Gate	دریچه ورودی
Output Gate	دریچه خروجی
Cell State	سلول حالت

Sigmoid	سیگموئید
Training Dataset	داده‌های آموزش
Validation Dataset	داده‌های اعتبارسنجی
Test dataset	داده‌های آزمون
Overfitting	بیش‌برازش
Adam	آدام
Lag	لگ
Data Scaling	مقیاس‌سازی داده‌ها
Hybrid Scaling	مقیاس‌سازی ترکیبی
Statistical Distribution Shift	تغییرات توزیع آماری
Data Leakage	نشت داده
Mean Squared Error	میانگین مربعات خطا
Root Mean Squared Error	خطای جذر میانگین مربعات
Mean Absolute Error	میانگین قدر مطلق خطا
MATLAB	متلب

منابع

- [۱] Adetunji, O.R., et al., *Tensile, hardness and microstructural properties of Sn-Pb solder alloys*. Materials today: proceedings, 2021. **44**: p. 321-325.
- [۲] Morris, J.W., J.L. Freer, and Z. Mei, *Microstructure Influences on the Mechanical Properties of Solder*. 1993.
- [۳] Tu, K.-N. and K. Zeng, *Tin-lead (SnPb) solder reaction in flip chip technology*. Materials science and engineering: R: reports, 2001. **34**(1): p. 1-58.
- [۴] Shi, X., et al., *Creep behavior and deformation mechanism map of Sn-Pb eutectic solder alloy*. J. Eng. Mater. Technol., 2003. **125**(1): p. 81-88.
- [۵] Darveaux, R., *Effect of simulation methodology on solder joint crack growth correlation and fatigue life prediction*. J. Electron. Packag., 2002. **124**(3): p. 147-154.
- [۶] Siviour, C., et al., *Mechanical properties of SnPb and lead-free solders at high rates of strain*. Journal of Physics D: Applied Physics, 2005. **38**(22): p. 4131-4139.
- [۷] Kang, S.K., et al., *Microstructure and mechanical properties of lead-free solders and solder joints used in microelectronic applications*. IBM Journal of Research and Development, 2005. **49**(4.5): p. 607-620.
- [۸] Shangguan, D., *Lead-free solder interconnect reliability*. 2005: ASM international.
- [۹] Rezaee-Bazzaz, A., *Superplastic power-law creep of Sn-40% Pb-2.5% Sb peritectic alloy*. Materials Science and Engineering A, 2007.
- [۱۰] Anand, L., *Constitutive equations for hot-working of metals*. International Journal of plasticity, 1985. **1**(3): p. 213-231.

- [١١] Bessa, M.A., et al., *A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality*. Computer Methods in Applied Mechanics and Engineering, 2017. **320**: p. 633-667.
- [١٢] Wang, H., et al., *Computational methods and artificial intelligence-based modeling of magnesium alloys: a systematic review of machine learning, deep learning, and data-driven design and optimization approaches*. Frontiers in Materials, 2025. **12**: p. 1645227.
- [١٣] Mozaffar, M., et al., *Deep learning predicts path-dependent plasticity*. Proceedings of the National Academy of Sciences, 2019. **116**(52): p. 26414-26420.
- [١٤] Lou, Q., X. Meng, and G.E. Karniadakis, *Physics-informed neural networks for solving forward and inverse flow problems via the Boltzmann-BGK formulation*. Journal of Computational Physics, 2021. **4** : 47p. 110676.
- [١٥] Goodfellow, I., *Deep learning*. 2019, MIT press.
- [١٦] Zhao, H., et al., *A comparative study of hot tensile deformation behavior of 6016 aluminum alloy under LSTM neural network and Arrhenius model*. Materials Research Express, 2024. **11**(1) :0p. 106517.
- [١٧] Xu, S., J. Xiong, and T.-Y. Zhang, *Translating strain to stress: A single-layer Bi-LSTM approach to predicting stress-strain curves in alloys during hot deformation*. Materials Research Express, 2024. **11**(7): p. 076526.
- [١٨] Mei, H., et al., *Study on constitutive relation of nickel-base superalloy Inconel 718 based on long short term memory recurrent neural network*. Metals, 2020. **10**(12): p. 1588.
- [١٩] Guo, Y., et al., *Machine learning method for fatigue strength prediction of nickel-based superalloy with various influencing factors*. Materials, 2022. **16**(1): p. 46.
- [٢٠] Deng, B., G. Zeng, and R. Ge, *A Creep Model of Steel Slag–Asphalt Mixture Based on Neural Networks*. Applied Sciences, 2024. **14**(13): p. 5820.
- [٢١] Wang, J., et al., *Prediction of molten pool temperature in laser solid forming based on PSO-LSTM*. Frontiers in Materials, 2023. **10**: p. 1218222.
- [٢٢] Hochreiter, S. and J. Schmidhuber, *Long short-term memory*. Neural computation, 1997. **9**(8): p. 1735-1780.
- [٢٣] Callister Jr, W.D. and D.G. Rethwisch, *Materials science and engineering: an introduction*. 2020: John wiley & sons.
- [٢٤] Courtney, T.H., *Mechanical behavior of materials*. 2005: Waveland Press.
- [٢٥] Géron, A., *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow " :٢٠٢٢ .O'Reilly Media, Inc."*.
- [٢٦] Bishop, C.M. and N.M. Nasrabadi, *Pattern recognition and machine learning*. Vol. 4. 2006: Springer.
- [٢٧] Zhang, A., et al., *Dive into deep learning*. 2023: Cambridge University Press.
- [٢٨] Graves, A., *Supervised sequence labelling*, in *Supervised sequence labelling with recurrent neural networks*. 2012, Springer. p. 5-13.
- [٢٩] Hastie, T., R. Tibshirani, and J. Friedman, *The elements of statistical learning*. 2009, Springer series in statistics New-York.
- [٣٠] Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980, 2014.

- [۳۱] Yeganeh, A., A. Shadman, and A. Amiri, *A novel run rules based MEWMA scheme for monitoring general linear profiles*. Computers & Industrial Engineering, 2021. **15** :2p. 107031.
- [۳۲] Montgomery, D.C., E.A. Peck, and G.G. Vining, *Introduction to linear regression analysis*. 2021: John Wiley & Sons.
- [۳۳] Das, S., et al., *Recurrent neural networks (RNNs): architectures, training tricks, and introduction to influential research*. Machine learning for Brain disorders. 2023.
- [۳۴] Bessa, M. and S. Pellegrino, *Design of ultra-thin shell structures in the stochastic post-buckling range using Bayesian machine learning and optimization*. International Journal of Solids and Structures, 2018. **139**: p. 174-188.
- [۳۵] James, P.E., et al., *Recurrent neural network-based speech recognition using MATLAB*. International Journal of Intelligent Enterprise, 2020. 7(1-3): p. 56-66.