Journal of Aerospace Mechanics/ 2024/ Vol.20/ No.2/ 105-121

Journal of Aerospace Mechanics

DOR: 20.1001.1.26455323.1403.20.2.8.2

# Predicting the Plastic Response of Circular Metal Plates under Uniform Dynamic Load Using Deep Neural Network

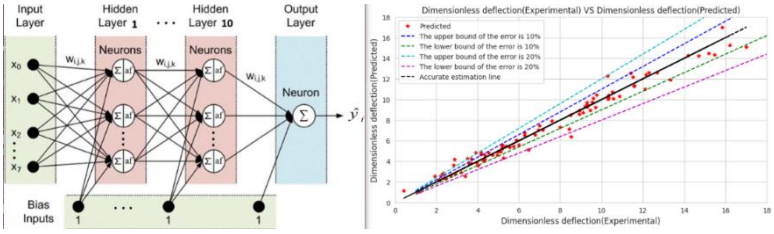
**Saeed Sarabadan [1*], Amirhossein Bagherian [2], Tohid Mirzababaie Mostofi [3]**

[1] Assistant Professor, Faculty of Science, Imam Hossein University, Tehran, Iran
[2] M.Sc Student, Faculty of Science, Imam Hossein University, Tehran, Iran
[3] Assistant Professor, Faculty of Mechanical Engineering, University of Eyvanekey, Eyvanekey, Iran

## HIGHLIGHTS

- Using deep learning to predict the permanent deflection of metal sheets.
- Using Adam's optimizer algorithm in deep neural network model architecture.
- Determining optimal metaparameters based on evaluator search engines.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Article Type: Research paper
Received: 8 February 2024
Received in revised form: 17 March 2024
Accepted: 20 April 2024
Available online: 18 June 2024
*Correspondence:
ssarabadan@ihu.ac.ir

*How to cite this article:*
S. Sarabadan, A. Bagherian, T.M. Mostofi. Predicting the plastic response of circular metal plates under uniform dynamic load using deep neural network. Journal of Aerospace Mechanics. 2024; 20(2):105-121.

*Keywords:*
Uniform dynamic loading
Circular metal sheets
Deep neural network
Artificial intelligence
Deflection
Metallic plates

## ABSTRACT

In the upcoming research, using deep neural networks, it is used to predict the maximum yield of circular metal sheets under uniform dynamic load. The neural network presented in this research was designed in the Python programming language and using the libraries available in it, including Tensorflow. The network is based on the regression problem and is of sequential type and includes 10 hidden layers which are the activation function in neurons of Leaky RELU type. The network optimizer algorithm was set to Adam and the objective function of the MSE problem and the number of network iterations was set to 700 times. The data set used in this article consists of 581 samples obtained from 16 series of experiments during the last forty years, which were standardized by the Scikit_learn library. The metal sheets are of 4 types: steel, aluminum, copper and titanium, and there is no separation between different metals. The number of training data in the model was determined to be 443 equals to 75% of the data set. Also, the number of experimental and evaluator data was selected as 88 numbers equivalent to 15% and 50 numbers equivalent to 10% of the entire data set. Each sample has 8 features as neural network inputs and one label as output. The presented intelligent model among the 88 test data that was completely randomly selected from the data set was able to classify 76% of the data, approximately equivalent to 67 numbers, within the error range of less than 10% and 88% of the data, or in other words, equivalent to approximately 78 numbers within the error range. Predict less than 20%. The amount of the root mean square error index decreased 102 times compared to the analytical and traditional predictive relationships available in the research records. Also, the coefficient of determination criterion, which is an important indicator for evaluating the performance of neural networks based on regression problems, includes the value of 0.96.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

**Publisher:** Imam Hossein University © Authors

مکانیک هوافضا/ سال ۱۴۰۳/ دوره ۲۰/ شماره ۲/ صفحه ۱۰۵-۱۲۱

نشریه علمی مکانیک هوافضا

DOR: 20.1001.1.26455323.1403.20.2.8.2

# پیش‌بینی پاسخ پلاستیک ورق‌های فلزی دایره‌ای تحت بار دینامیکی یکنواخت با استفاده از شبکه عصبی عمیق

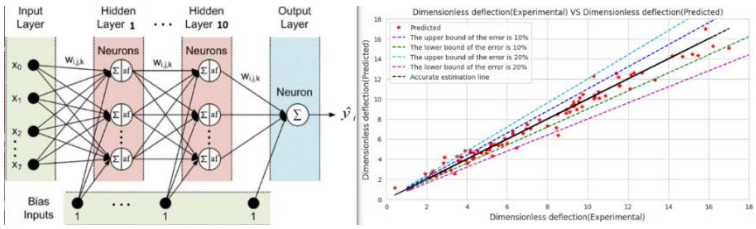
**سعید سرآبادان[1*]، امیرحسین باقریان[2]، توحید میرزابابای مستوفی[3]**

[1] استادیار، دانشکده علوم پایه، دانشگاه امام حسین(ع)، تهران، ایران

[2] کارشناسی ارشد، دانشکده علوم پایه، دانشگاه امام حسین(ع)، تهران، ایران

[3] استادیار، دانشکده مهندسی مکانیک، دانشگاه ایوان‌کی، ایوان‌کی، ایران

## چکیده گرافیکی




## برجسته‌ها

- استفاده از یادگیری عمیق جهت پیش‌بینی خیز دائمی ورق‌های فلزی
- استفاده از الگوریتم بهینه‌ساز آدام در معماری مدل شبکه عصبی عمیق
- تعیین فراپارامترهای بهینه بر اساس جستجو گرهای ارزیاب

## چکیده

در تحقیق پیش رو با استفاده از شبکه‌های عصبی عمیق به پیش‌بینی میزان بیشترین خیز ورق‌های فلزی دایره‌ای شکل تحت بار شدید دینامیکی یکنواخت پرداخته می‌شود. شبکه عصبی ارائه‌شده در این تحقیق در محیط زبان برنامه‌نویسی پایتون و با استفاده از کتابخانه‌های موجود در آن ازجمله تنسورفلو طراحی گردید. مدل طراحی شبکه مبتنی بر مسئله رگرسیون و از نوع سکوئنشال و شامل ۱۰ لایه می‌باشد که تابع فعال‌ساز موجود در نورون‌ها از نوع لیکی رلو (Leaky RELU) هستند. الگوریتم بهینه‌ساز مدل روی آدام و تابع هدف مسئله میانگین مربعات خطا و تعداد تکرار شبکه روی ۷۰۰ مرتبه تنظیم شد. مجموعه داده مورداستفاده در این مقاله متشکل از ۵۸۱ نمونه حاصل از ۱۶ سری آزمایش در طی چهل سال گذشته می‌باشد که به‌وسیله کتابخانه سایکیت-لرن استانداردسازی شدند. ورق‌های فلزی از چهار جنس فولادی، آلومینیوم، مس و تیتانیوم می‌باشند و هیچ‌گونه تفکیکی میان فلزات مختلف صورت نگرفته است. تعداد داده‌های آموزشی در مدل ۴۴۳ عدد معادل ۷۵٪ از مجموعه داده تعیین شد. همچنین تعداد داده‌های آزمایشی و ارزیاب به ترتیب ۸۸ عدد معادل ۱۵٪ و ۵۰ عدد معادل ۱۰٪ از کل مجموعه داده انتخاب شد. هر نمونه دارای ۸ ویژگی به‌عنوان ورودی‌های شبکه عصبی و یک برچسب به‌عنوان خروجی می‌باشد. مدل هوشمند ارائه‌شده از میان ۸۸ داده آزمایشی که به‌صورت کاملاً تصادفی از مجموعه داده انتخاب‌شده بود، توانست ۷۶٪ از داده‌ها تقریباً معادل ۶۷ عدد را در محدوده خطای کمتر از ۱۰٪ و ۸۸٪ از داده‌ها یا به‌عبارت‌دیگر معادل تقریباً ۷۸ عدد را در محدوده خطای کمتر از ۲۰٪ پیش‌بینی کند. میزان شاخص ریشه میانگین مربعات خطا $10^{-2}$ برابر نسبت به روابط پیش‌بینی کننده تحلیلی و سنتی موجود در سوابق تحقیق کاهش پیدا کرد. همچنین معیار ضریب تعیین که شاخصی مهم جهت ارزیابی عملکرد شبکه‌های عصبی مبتنی بر مسائل رگرسیون می‌باشد مقدار ۰/۹۶ را در بر گرفته است.

## مشخصات مقاله

تاریخچه مقاله:

نوع مقاله: علمی پژوهشی

دریافت: ۱۴۰۲/۱۱/۱۹

بازنگری: ۱۴۰۲/۱۲/۲۷

پذیرش: ۱۴۰۳/۰۲/۰۱

ارائه برخط: ۱۴۰۳/۰۳/۲۹

*نویسنده مسئول:

ssarabadan@ihu.ac.ir

کلیدواژه‌ها:

بارگذاری دینامیکی یکنواخت

ورق‌های فلزی دایره‌ای شکل

شبکه عصبی عمیق

هوش مصنوعی

خیز

ورق‌های فلزی

* این مقاله یک مقاله با دسترسی آزاد است که تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY) توزیع‌شده است.

© نویسندگان

**ناشر:** دانشگاه جامع امام حسین(ع)

## ۱- مقدمه

با پیشرفت تکنولوژی و لزوم کسب دانش دقیق جهت ارزیابی پاسخ سازه‌ها تحت بارگذاری دینامیکی شدید، روش‌های تجربی جای خود را به روش‌های مبتنی بر یادگیری ماشین و یادگیری عمیق داده‌اند که نقش به سزایی در پویایی ارزیابی‌ها، کاهش زمان، هزینه‌ها و آسیب‌های احتمالی خواهند داشت. در جدیدترین تحقیقات در این حوزه ساکاریدس [۱] با استفاده از نتایج آزمایش‌های منتشرشده در سالیان گذشته در زمینه پاسخ کمانش سازه‌های لوله‌ای و استفاده از مفاهیم یادگیری ماشین نسبت به فروپاشی سازه‌ها پیش‌بینی‌های دقیق‌تری انجام داده است. شایان توجه است که در زمینه ارائه مدل‌های تحلیلی و تجربی جهت پیش‌بینی پاسخ دینامیکی ورق‌های دایره‌ای و مستطیلی تحت بارگذاری دینامیکی تحقیقات گسترده‌ای صورت گرفته که به برخی از آن‌ها در ادامه اشاره می‌شود. [۲-۱۲].

در سال ۱۹۸۹ نوریک و مارتین [۱۳] اصلاحاتی را روی عدد بی‌بعد جانسون که در سال ۱۹۵۱ ارائه داده بود انجام دادند و عدد بی‌بعد ایمپالس $\phi_c$ را برای پیش‌بینی رفتار دینامیکی ورق‌های دایره‌ای تحت بارگذاری دینامیکی یکنواخت به‌صورت رابطه (۱) پیشنهاد کردند:

$$\phi_c = \frac{I}{\pi R H^2 \sqrt{\rho \sigma_0}} \qquad (۱)$$

سپس در سال ۲۰۰۷ ژاکوب و همکاران [۱۴] عدد بی‌بعد ایمپالس $\phi_{cs}$ را با اعمال فاصله استخراج خرج از سطح ورق که با نماد S نمایش داده می‌شود، برای بارگذاری دینامیکی به‌صورت رابطه (۲) ارائه دادند:

$$\phi_{cs} = \frac{I(1 + ln(R/R_0))}{(1 + ln(S/R_0))\pi R H^2 \sqrt{\rho \sigma_0}} \qquad (۲)$$

همچنین نوریک [۸] در سال ۲۰۱۷ رابطه تجربی خود را برحسب عدد بی‌بعد ایمپالس $\phi_{cs}$ برای پیش‌بینی بیش‌ترین خیز دائمی ورق‌های دایره‌ای تحت بار دینامیکی یکنواخت ($W_0$) نسبت به ضخامت به‌صورت رابطه (۳) ارائه کرده است.

$$\frac{W_0}{H} = 425\phi_{cs} + 0.277 \qquad (۳)$$

در مقاله ارائه‌شده توسط مستوفی و همکاران [۳] با استفاده از بی‌بعدسازی روابط حاکم بر ورق، این اثر به تحلیل اضافه گردید. همچنین یکی دیگر از مشکلات موجود در مراجع قبلی [۸، ۱۴ و ۱۵] عدم بیان دلیل برای معرفی اعداد بی‌بعد بود که برای حل این مشکل، اعداد بی‌بعد پیشنهادی، از بی‌بعدسازی معادلات تعادل دینامیکی حاکم بر ورق استخراج شدند [۳]. همچنین، کمیت‌های مهمی نظیر هندسه ورق، توانایی مقاومت دینامیکی ماده در برابر بار وارده، حساسیت ماده به نرخ کرنش و همچنین اینرسی بار دفعی واردشده به موارد فوق اضافه گردید. درنهایت معادلات ساختاری تجربی بر اساس نتایج تجربی موجود و همچنین روش تجزیه مقادیر منفرد بهینه‌شده، به دست آمد.

در مقاله حاضر، علاوه بر افزایش دقت، به افزایش شاخص سرعت در محاسبات نیز پرداخته شد تا مدل بتواند به‌صورت هوشمند، داده‌های ورودی را در سریع‌ترین زمان ممکن با حداقل خطا پردازش و خروجی نهایی را در اختیار کاربر قرار دهد. همچنین این مدل با داشتن تقریباً ۶۰ هزار پارامتر، در صورت استفاده از منابع محاسباتی کارآمد قادر خواهد بود حجم عظیمی از داده‌های دریافتی را پردازش و مورد پیش‌بینی قرار دهد. در روابط تحلیلی پیشین، هدف تعیین بهترین منحنی گذرنده از داده‌ها با حل دستگاه معادلات و تعیین مجهولات موجود در رابطه بود که نیاز به حل روابط و معادلات پیچیده ریاضیاتی داشت و به‌صورت آزمون‌وخطا جهت رسیدن به کمترین میزان شاخص ارزیابی خطا که عموماً میانگین مربعات خطا بود صورت می‌گرفت. روش ارائه‌شده در این تحقیق مبتنی بر روش‌های تکراری و استفاده از روابط ساده ریاضیاتی که پیچیدگی محاسباتی بسیار کم‌تری نسبت به روابط تحلیلی دارند صورت می‌پذیرد. این پیچیدگی محاسباتی کم‌تر این امکان را می‌دهد تا بتوان روش تکراری را تا مراتب بسیار بالاتر اجرا کنیم و مدل روابط موجود میان داده‌ها را با دقت بیشتری جست‌وجو و برآورد می‌کند. همچنین، ضعف مدل‌های کلاسیک تأخیر در پردازش داده‌های با حجم بالا می‌باشد. درواقع با افزایش میزان داده‌های ورودی مدل مجبور به حل معادلات سنگین خواهد شد که نیاز به منابع محاسباتی گران و زمان طولانی خواهد داشت. همچنین در مدل‌های مبتنی بر یادگیری ماشین ارائه‌شده، شبکه هیچ‌گاه داده‌های آموزشی را در ارزیابی و محاسبه خطا لحاظ نکرده و عملکرد شبکه بر اساس داده‌های

آزمایشی که مدل بر اساس آن‌ها تعلیم ندیده و کاملاً برای شبکه جدید هستند ارزیابی می‌گردد که منجر به ارزیابی واقعی و دقیق‌تر مدل خواهد شد [۱۶-۱۹]. در این مقاله، تلاش گردیده تا با استفاده از یادگیری عمیق و ایجاد یک مدل شبکه عصبی در راستای بهبود میزان دقت و شاخص‌های ارزیابی خطا، پاسخ ورق‌های تک لایه فلزی دایره‌ای تحت بارگذاری شدید دینامیکی یکنواخت مورد پیش‌بینی قرارگرفته و با روش‌های سنتی و تحلیلی موجود در سوابق تحقیق مقایسه گردد [۲۰-۴۵]. داده‌های موجود در ادبیات تحقیق از ۴ نوع فلز فولاد، آلومینیوم، تیتانیوم و مس که سهم هرکدام به ترتیب ۴۸۶، ۷۰، ۱۲ و ۱۳ عدد از ۵۸۱ داده موجود در مجموعه داده می‌باشد، تشکیل‌شده است. هر داده دارای ۸ ویژگی شامل میزان نیروی واردشده بر ورق، ضخامت ورق، شعاع ورق، مقاومت استاتیکی ورق، چگالی ورق، کمیت‌های مربوط به حساسیت ماده به نرخ کرنش و همچنین یک برچسب تحت عنوان میزان خیز دائمی ورق می‌باشد.

## ۲- هوش مصنوعی و استفاده از داده‌های تجربی

کار با داده‌ها پس از گسترش مدل‌های هوشمند در میان محققان و تحلیلگران دچار تغییرات زیادی شد. دیگر منابع ذخیره‌سازی اطلاعات، گنجایش داده‌های تولیدی روزافزون را نداشته و مدل‌های مبتنی بر ریاضیات کلاسیک توانایی پردازش و کار با حجم عظیم داده‌ها را از دست دادند. در اینجا نقش‌آفرینی پدیده نوظهور حوزه تکنولوژی و کامپیوتر یعنی هوش مصنوعی در کنار پردازنده‌های گرافیکی در پردازش داده‌ها آشکار می‌شود. جایی که روابط تحلیلی سنتی در تخمین و پردازش داده‌هایی که نیاز به هزاران پارامتر در معادلات خود را دارند قادر به پاسخگویی نخواهند بود. حتی در خوش‌بینانه‌ترین حالت اگر رابطه‌ای تحلیلی برای تخمین و پردازش داده‌های خود بیابیم، زمان عملکرد الگوریتم وجه تمایز اصلی میان روش‌های هوشمند امروزی با مدل‌های کلاسیک گذشته خواهد بود. اگرچه امروزه با افزایش داده‌های ورودی در طراحی مدل‌های مبتنی بر هوش مصنوعی با مشکلاتی از قبیل افزایش بیش‌ازحد پارامترها، اشباع شبکه و افزایش پیچیدگی محاسباتی مواجه خواهیم شد اما می‌توان با استفاده از شبکه‌های عصبی کانولوشنی تا حد بسیار زیادی از میزان پارامترهای اضافی کاست و مدلی سبک‌تر به لحاظ پردازش محاسباتی داشت.

### ۲-۱- یادگیری عمیق

یادگیری عمیق یکی از الگوریتم‌های حوزه یادگیری ماشین و نوین‌ترین دستاورد حوزه هوش مصنوعی است که به جهت کاربردهای متفاوت آن در مدل بندی مسائل علم و فناوری، موردتوجه قرارگرفته و گرایش‌های متفاوتی از علوم را به خود وابسته نموده است. در یادگیری عمیق طراح شبکه عصبی با عمیق کردن شبکه یا به عبارت بهتر، اضافه کردن لایه‌های پنهان به شبکه آن را قادر به یادگیری الگوریتم‌های پیچیده‌تر و دریافت حجم عظیمی از داده‌ها می‌کند.

هر مدل یادگیری عمیق به‌صورت کلی متشکل از لایه ورودی، لایه پنهان و لایه خروجی می‌باشد. در هر لایه، بسته به نوع مسئله تعدادی نورون قرار می‌گیرد که در هسته هر نورون توابع پله‌ای و یا تابع فعال‌ساز قرار داده می‌شود [۱۶]. در مدل یادگیری عمیق با رفتن از هر لایه به لایه دیگر جمع وزن‌دار مجموعه نورون‌های لایه قبل محاسبه و با اعمال یک تابع فعال‌ساز غیرخطی به لایه دیگر منتقل می‌شوند؛ اما اینکه چگونه هر وزن باید تغییر کند کمی دشوار است، این بدان معنی است که خطای به‌دست‌آمده در خروجی لایه نهایی می‌تواند ناشی از نورون‌های لایه آخر و همچنین نورون‌های لایه‌های قبلی باشد که توسط الگوریتم پس انتشار سهم خطای هر نورون در هر لایه مشخص و فرآیند به‌روزرسانی وزن‌ها تا رسیدن به بهترین ضرایب ادامه می‌یابد [۱۷]. از مفاهیم اصلی در فرآیند یادگیری در شبکه‌های عمیق، بهینه‌سازی الگوریتم‌های تابع خطا می‌باشد که توسط فرآیندی به نام گرادیان کاهشی انجام می‌پذیرد [۱۸]. پیش‌شرطی‌سازی مؤثر با گرادیان کاهشی به معنای انتخاب نرخ یادگیری متفاوت برای هر متغیر است، این امر سبب می‌گردد تا بهینه‌سازها با روش‌های بهتری با توجه به نوع مسئله در نظر گرفته شوند که می‌توان به بهینه‌سازهای تصادفی، پویا و تطبیقی اشاره کرد [۱۹]. شایان توجه است که جهت مدل‌سازی از داده‌های موجود در ادبیات تحقیق که در جدول **۱** ارائه‌شده، استفاده گردیده است.

**جدول (۱):** جمع‌آوری نتایج تجربی موجود در ادبیات تحقیق.

| مرجع | سال تحقیق | تعداد آزمایش | قطر ورق (mm) | ضخامت ورق (mm) | تنش تسلیم (MPa) | محدوده بیشترین خیز (mm) | محدوده ایمپالس (N·s) |
|---|---|---|---|---|---|---|---|
| [۲۰] | ۱۹۶۶ | ۴۲ | ۲۰۰ | ۶/۱ و ۶/۴ | ۵۴۵ و ۲۹۰ | ۴۲/۷-۲/۴ | ۱۱۳/۰-۲۴/۲ |
| [۲۱] | ۱۹۷۰ | ۳۰ | ۲۰۳ | ۶/۲ و ۶/۴ | ۲۸۳ و ۲۹۰ | ۲۸/۳-۳/۸ | ۱۳۷/۶-۳۲/۲ |
| [۲۲] | ۱۹۷۹ | ۲۰ | ۶۴ | ۱/۹ و ۲/۳ | ۲۲۳ و ۲۵۱ | ۱۲/۷-۰/۸ | ۷/۲-۰/۹۱ |
| [۱۵] | ۱۹۸۹ | ۲۱ | ۱۰۰ | ۱/۶ | ۲۸۲ | ۱۹/۸-۶/۱ | ۱۵/۶-۵/۶ |
| [۲۳] | ۱۹۹۱ | ۹ | ۱۰۰ | ۱/۶ | ۲۶۴ | ۳۰/۱-۳/۷ | ۲۹/۲-۴/۸ |
| [۲۴] | ۱۹۹۴ | ۱۴۲ | ۱۰۰ | ۱/۶ | ۲۷۰ | ۲۷/۹-۴/۶ | ۲۲/۰-۴/۶ |
| [۲۵] | ۱۹۹۵ | ۲۰ | ۱۰۰ | ۲/۰-۱/۵ | ۲۶۲ | ۲۰/۹-۵/۶ | ۱۶/۳-۴/۶ |
| [۲۶] | ۱۹۹۶ | ۱۱۴ | ۶۰، ۸۰، ۱۰۰ و ۱۲۰ | ۱/۶ | ۲۹۰ | ۳۴/۳-۲/۷ | ۳۰/۹-۱/۴ |
| [۲۷] | ۱۹۹۶ | ۷ | ۱۰۰ | ۱/۶ | ۲۵۵ | ۲۰/۴-۷/۳ | ۱۶/۵-۵/۲ |
| [۱۴] | ۲۰۰۷ | ۴۸ | ۱۰۶ | ۱/۹ | ۲۴۰ | ۱۹/۰-۵/۴ | ۳۳/۵-۹/۶ |
| [۴۵] | ۲۰۰۷ | ۲۵ | ۱۰۶ | ۲/۰ | ۲۶۱ | ۱۸/۶-۰/۵ | ۲۹/۰-۵/۴ |
| [۲۸] | ۲۰۱۰ | ۴۲ | ۱۰۰ | ۱/۶، ۲/۰ و ۳/۰ | ۱۰۲، ۲۰۱، ۲۷۷ و ۳۱۸ | ۲۶/۳-۳/۲ | ۳۴/۶-۵/۲ |
| [۲۹] | ۲۰۱۱ | ۷ | ۱۷۰ | ۳/۰ | ۲۱۵ | ۳۳/۸-۱۲/۷ | ۲۴/۳-۱۷/۴ |
| [۳۰] | ۲۰۱۴ | ۱۳ | ۱۰۶ | ۲/۰ و ۳/۰ | ۷۵۰ | ۱۷/۷-۱/۳ | ۵۹/۸-۱۱/۹ |
| [۳۱] | ۲۰۱۵ | ۱۵ | ۱۲۰ | ۱/۰، ۲/۰ و ۳/۰ | ۲۸۹ | ۳۵/۸-۶/۱ | ۵۱/۶-۴/۱ |
| [۳۲] | ۲۰۲۰ | ۲۶ | ۱۰۰ | ۲/۰، ۳/۰ و ۶/۰ | ۱۱۰ و ۳۵۰ | ۱۶/۱-۶/۶ | ۲۱/۷-۶/۵ |

## ۲-۲- رگرسیون در یادگیری عمیق

رگرسیون یک فرآیند آماری است که رابطه‌ی معناداری بین متغیرهای وابسته و مستقل پیدا می‌کند [۳۳]. الگوریتم‌های رگرسیون برای پیش‌بینی مقادیر پیوسته استفاده می‌شوند و هدف در این مدل بندی جستجوی ضرایب بهینه جهت رسیدن به بهترین پیش‌بینی کننده می‌باشد که نماینده خوبی برای داده‌ها است. در مسائل مربوط به شبکه‌های عصبی عمیق، ابعاد بسیار بالاتر برای متغیرهای مستقل و نیاز به محاسبات پیچیده جهت تعیین ضرایب بهینه موردتوجه قرار می‌گیرد. از شاخص‌های ارزیابی خطا در مسائل مبتنی بر رگرسیون که بیانگر کیفیت عملکرد برآوردگر است (شکل ۱)، میانگین مربعات خطا و ضریب تعیین می‌باشد که نحوه محاسبه آن به ترتیب به‌صورت روابط (۴) و (۵) است [۳۴].

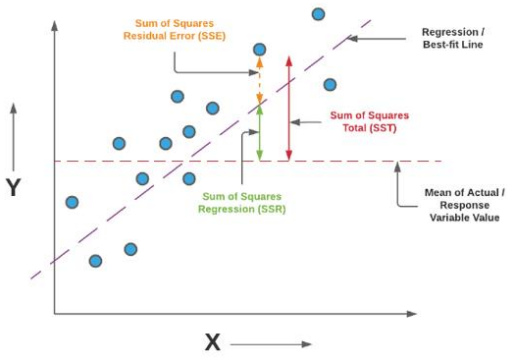
$$MSE(y,\hat{y}) = \frac{1}{n}\left(\sum_{i=1}^{n}(y_i - \hat{y}_i)\right)^2 \quad (۴)$$

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (۵)$$

در رابطه (۵) همان‌طور که در ادامه رابطه بسط داده‌شده است، منظور از SSR مجموع مربعات خطا به ترتیب بین مقدار

پیش‌بینی‌شده هر داده توسط مدل و مقدار میانگین مقادیر واقعی داده‌ها می‌باشد. همچنین منظور از SST مجموع مربعات خطا به ترتیب بین مقدار واقعی هر داده و میانگین داده‌های واقعی می‌باشد.

یکی از مباحث مهم و مرسوم در تعلیم مدل‌های مبتنی بر شبکه‌ی عصبی، استفاده از شبکه‌های بزرگ آموزش‌دیده و ارائه‌شده توسط شرکت‌های بزرگ فعال در حوزه‌ی صنعت دیجیتال جهت شخصی‌سازی آن‌ها با داده‌های موردنظر کاربر است. یکی از چالش‌های موجود در مدل‌سازی شبکه‌ی عصبی ارائه‌شده در این تحقیق علی‌رغم تلاش‌های صورت گرفته، عدم دسترسی به شبکه‌ای مناسب از نوع رگرسیون و سازگار با داده‌های موجود در مجموعه‌ی داده جهت افزایش کارایی مدل در برخورد با طیف وسیعی از فلزات دیگر بود. ازاین‌رو ایجاد یک معماری جدید مبتنی بر مسئله رگرسیون و سازگار با داده‌های محدود موجود از پایه انجام شد.




**شکل (۱):** شاخص‌های ارزیابی خط رگرسیون.

## ۲-۳- ضریب همبستگی

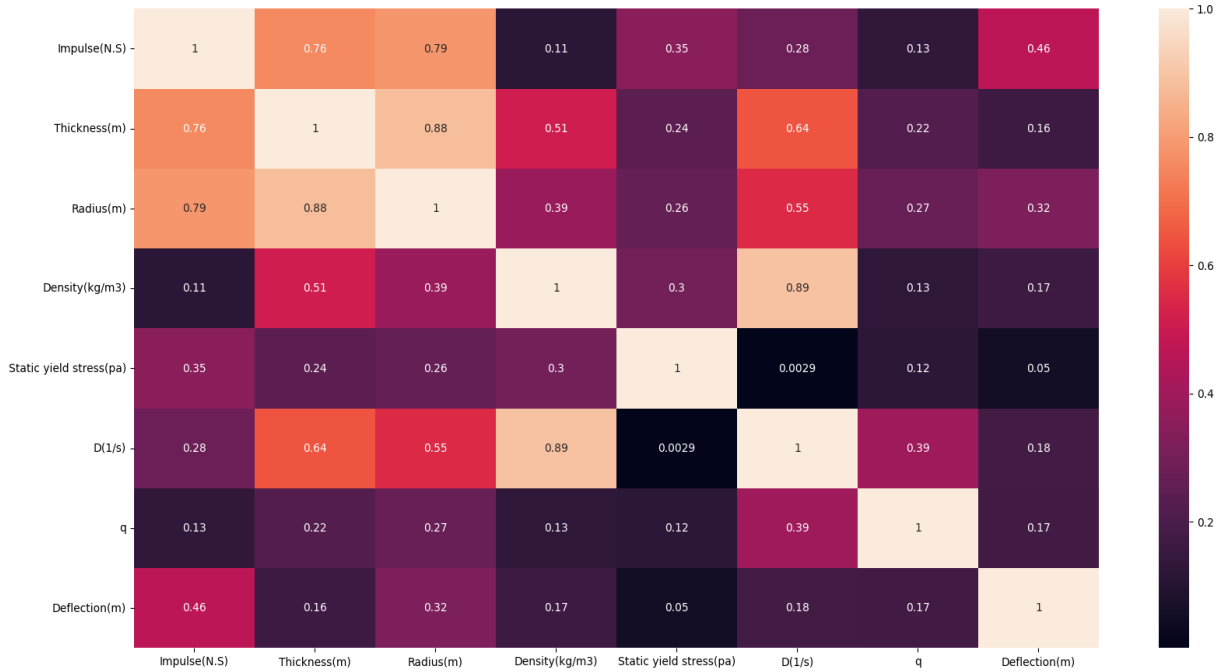
اغلب بیان همبستگی میان دو متغیر تصادفی بر اساس شاخص کوواریانس مشهور می‌باشد اما مشکل این شاخص وابستگی به واحد اندازه‌گیری متغیرها است، همین دلیل مبنایی بر استفاده از شاخص ارزیابی دیگری به نام ضریب همبستگی شد که واحدی نداشته و بی‌بعد می‌باشد. ضریب همبستگی میان دو متغیر تصادفی $X$ و $Y$ به شکل زیر محاسبه می‌شود:

$$\rho_{X,Y} = \frac{\operatorname{cov}\ (X,Y)}{\sqrt{\operatorname{var}(X)}\sqrt{\operatorname{var}(Y)}} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \quad (۶)$$

که در رابطه بالا منظور از $E$ امید ریاضی، $\mu$ میانگین داده‌ها به ازای هر متغیر تصادفی و $\sigma$ انحراف معیار به ازای هر متغیر تصادفی است.

ضریب همبستگی، شاخصی در بازه ۱- تا ۱ می‌باشد. هرگاه میزان شاخص به ۱ یا ۱- نزدیک‌تر شود همبستگی میان دو متغیر بیشتر است. علامت مثبت و منفی بیانگر جهت رابطه و همبستگی میان دو متغیر است به‌گونه‌ای که اگر علامت ضریب همبستگی منفی باشد بیانگر آن است که رابطه بین دو متغیر معکوس یکدیگر است و هرگاه مثبت باشد آن دو متغیر باهم همبستگی مستقیم دارند یعنی با افزایش مقدار یک متغیر، میزان متغیر دیگر نیز با همان نسبت افزایش خواهد یافت. این شاخص به‌صورت درصد هم بیان می‌شود.

ضریب همبستگی در یادگیری ماشین شاخص و معیاری ارزشمند جهت تشخیص میزان همبستگی میان ویژگی‌های موجود در مجموعه داده می‌باشد. بر اساس این شاخص طراح شبکه می‌تواند رفتار داده‌های خود را بررسی و بر اساس میزان تأثیر هرکدام از ویژگی‌ها بر متغیر خروجی شبکه یا برچسب آن‌ها را تغییر دهد و یا گاهی حذف کند و همچنین چگونگی عملکرد شبکه را بر اساس میزان این شاخص تفسیر کند. شایان‌ذکر است هر چه میزان همبستگی میان ویژگی‌ها به‌صورت دوبه‌دو و به‌خصوص میان هر ویژگی با برچسب بیشتر باشد عملکرد شبکه بهتر و تأثیر ویژگی‌ها در پیش‌بینی مقدار دقیق‌تر برچسب خروجی از شبکه بیشتر است. در تحقیق پیش رو میزان ضریب همبستگی میان ویژگی‌های موجود در مجموعه داده به‌صورت دوبه‌دو محاسبه و به‌صورت شکل **۲** گزارش می‌شود. همان‌گونه که پیداست قطر اصلی این ماتریس ۱ می‌باشد زیرا هر ویژگی با خودش دارای ضریب همبستگی حداکثری با مقدار ۱ است.




**شکل (۲):** میزان همبستگی میان ویژگی‌های ورق

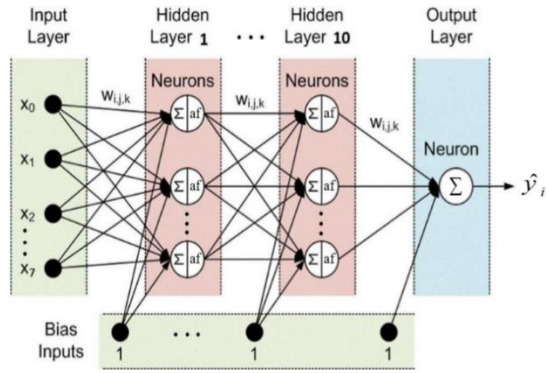
## ۲-۴- گرادیان کاهشی

نحوه اصلاح وزن‌ها در یک شبکه عصبی مصنوعی تا رسیدن به بهترین برآورد کننده داده‌ها بر اساس فرآیند بهینه‌سازی صورت می‌پذیرد. درواقع یک شبکه عصبی مصنوعی با تعیین تابع خطا به‌عنوان تابع هدف آن را تا یافتن حداقل مطلق کمینه می‌کند. این کمینه‌سازی به‌وسیله الگوریتم گرادیان کاهشی که در برابر روش‌های مشابه دیگر همچون معادله نرمال هزینه محاسباتی کمتر و در برابر داده‌های با حجم بالا سرعت و دقت بالاتری دارد انتخاب می‌شود. در یادگیری ماشین الگوریتم گرادیان کاهشی با داده‌ها با سه رویکرد برخورد می‌کند. رویکرد متداول گرادیان کاهشی تصادفی ریز دسته‌ای می‌باشد، در این شیوه داده‌ها به دسته‌بندی‌های متعدد تقسیم‌بندی می‌شود و با دریافت هر دسته فرآیند به‌روزرسانی انجام می‌پذیرد که تا حد امکان مشکلات اشاره‌شده در شیوه‌های قبلی را برطرف می‌کند. این رویکرد در کتابخانه‌های زبان پایتون به‌عنوان پیش‌فرض تنظیم و عمل می‌کند که در این شبکه نیز از آن استفاده‌شده است. در الگوریتم گرادیان کاهشی همان‌گونه که به‌صورت گذرا به آن اشاره شد، مشکل و خطری که مدل ما را تهدید می‌کند، گیر افتادن در کمینه‌های محلی یا موضعی است. این تهدید سبب پیدایش و ارائه الگوریتم‌های بهینه‌ساز پیشرفته و جدید همچون آدادلتا، آداگراد، آر ام اس پراپ و آدام شد. درواقع آنچه وجه تمایز میان این الگوریتم‌ها می‌باشد، تفاوت در تعیین نرخ‌های یادگیری و فراپارامترهای درون روابط بوده است. درواقع دانشمندان الگوریتم‌ها را به سویی هدایت کرده‌اند که با هر گام و هر مرتبه اعمال الگوریتم گرادیان کاهشی مدل اندازه گام خود را به‌صورت یکسان انتخاب نکند و در هر مرحله با توجه به سوابق گرادیان در گام‌های قبل و تابع هدف مسئله گام خود را تغییر دهد تا در حداقل‌های موضعی به دام نیفتد (شکل **۳**). عبارت ساده (۷) تابع خطا را به‌صورت یک معادله نشان می‌دهد.

$$Loss = y - \hat{y} \qquad (۷)$$

که در آن به ترتیب $y$ و $\hat{y}$ مقادیر واقعی و پیش‌بینی‌شده توسط شبکه هستند. رابطه (۸) عملکرد الگوریتم گرادیان کاهشی روی تابع خطا، در به‌روزرسانی ضرایب شبکه را نمایش می‌دهد. در این رابطه، منظور از $m$ تعداد داده‌ها، $h_w(x^{(i)})$ تابع مدل برآوردکننده به ازای ورودی $i$، $y^{(i)}$ مقدار واقعی یا تجربی داده و $\eta$ نرخ یادگیری الگوریتم می‌باشد.

$$J(W) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_w\left(x^{(i)}\right) - y^{(i)}\right)^2$$
$$w_{new} = w - \eta\frac{\partial J(W)}{\partial w} = w - \eta \cdot [2x(wx + b - y)] \quad (۸)$$




**شکل (۳):** طرح‌واره شبکه عصبی عمیق.

## ۲-۵- الگوریتم بهینه‌ساز مدل (آدام)

بر اساس آنچه تابه‌حال بررسی کردیم دریافتیم که گرادیان کاهشی تصادفی در حل مسائل بهینه‌سازی از گرادیان کاهشی مؤثرتر است، به‌عنوان‌مثال، به دلیل انعطاف‌پذیری ذاتی آن در برابر داده‌های اضافی. همچنین گرادیان کاهشی ریز دسته‌ای با استفاده از مجموعه‌های بزرگ‌تری از مشاهدات در یک ریز دسته، کارایی اضافی قابل‌توجهی را نسبت به گرادیان کاهشی تصادفی ارائه می‌دهد. رویکرد پیشرفت عملکرد الگوریتم‌های بهینه‌ساز، وابسته به گرادیان‌های طی کرده در گام‌های گذشته و میزان اهمیت آن‌ها برای برداشتن گام بعدی در جهت همگرایی سریع و دقیق‌تر است. آدام ارائه‌شده در سال ۲۰۱۴ توسط کینگما [۳۵] تمام ایده‌های مطرح‌شده در سیر تکامل الگوریتم‌های بهینه‌ساز برای تجمیع تاریخچه گرادیان‌های گذشته و اعمال مربع گرادیان برای تسریع همگرایی و مقیاس بندی هر مختصات جهت ایجاد یک پیش‌شرط محاسباتی کارآمد را در نظر گرفت. اگرچه الگوریتم آدام عاری از مشکل نیست و بهبودهایی بر آن ارائه شد اما تا به امروز محبوبیت خود را در میان طراحان شبکه‌های عصبی از دست نداده و همچنان کلید حل مسائل دشوار یادگیری ماشین است که در مدل ارائه‌شده در این تحقیق نیز به‌عنوان الگوریتم بهینه‌ساز شبکه تعریف گردید.

## ۲-۶- تابع فعال‌ساز لیکی رلو

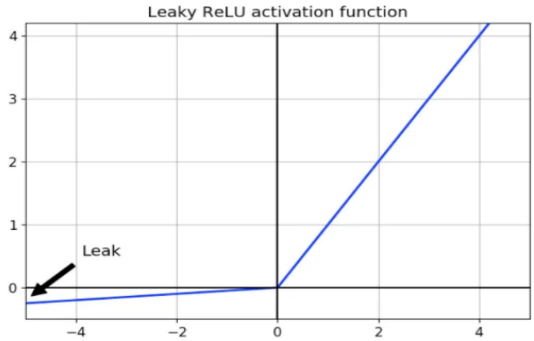
تابع فعال‌ساز در هر نورون مصنوعی اگر سیگنال‌های دریافتی اصطلاحاً به حد آستانه رسیده باشند، سیگنال‌های خروجی را برای نورون بعدی ارسال می‌کند. درواقع تابع فعال‌ساز تصمیم می‌گیرد که یک نورون باید فعال شود یا خیر. یک شبکه عصبی مصنوعی در یادگیری عمیق بدون وجود تابع فعال‌ساز یک مدل رگرسیون خطی ساده است به این دلیل که این توابع محاسبات غیرخطی را در ورودی یک شبکه عصبی مصنوعی انجام می‌دهند و آن را قادر به یادگیری و انجام وظایف پیچیده‌تر می‌کنند؛ بنابراین انتخاب تابع فعال‌ساز مناسب که بتواند غیرخطی بودن و دقت را یک مدل شبکه عصبی ارائه دهد حائز اهمیت می‌باشد. استفاده از توابع فعال‌ساز خانواده رلو مشکل محو گرادیان را تا حد زیادی بهبود می‌دهد اما دونقطه ضعف این توابع، عدم مشتق‌پذیری در نقطه صفر و همچنین تولید مقدار صفر یا ایجاد نورون مرده به ازای مقادیر منفی است. برای حل این مشکل در مقاله‌ای در سال ۲۰۱۰ توابع جدیدتری ازجمله لیکی رلو ارائه شدند که مشکلات تابع رلو را برطرف کرد.

توابع فعال‌ساز سیگموید و تانژانت هایپربولیک را به دلیل مشکل محوگرادیان نمی‌توان در شبکه‌هایی با لایه‌های زیاد به کار برد. تابع فعال‌ساز لیکی رلو که محبوب‌ترین انتخاب در یادگیری عمیق (در لایه پنهان) است بر مشکل محو گرادیان غلبه می‌کند و به شبکه اجازه می‌دهد سریع‌تر یاد بگیرد و عملکرد بهتری داشته باشد. ضابطه تابع لیکی رلو به‌صورت $\text{Leaky ReLU}(x) = \max(0.1x, x)$ تعریف می‌گردد (شکل ۴). همچنین در این تابع به سبب در نظر گرفتن ضریبی به‌عنوان $\alpha$، مشکل ایجاد نورون مرده به ازای ورود اعداد منفی برای تابع رفع گردید.

## ۲-۸- مدل‌سازی مسئله با شبکه عصبی

مسئله پیش رو در این مقاله از نوع یادگیری با ناظر و رگرسیون می‌باشد. جهت طراحی شبکه عصبی عمیق این مقاله، ابتدا بانک اطلاعاتی شامل مجموعه داده با ۵۸۱ نمونه که حاصل از ۱۶ سری آزمایش بین سال‌های ۱۹۶۶ تا ۲۰۲۰ بوده و شامل ۸ ویژگی می‌باشد را به‌عنوان داده اصلی در نظر

گرفته شد. تمام ویژگی‌ها از نوع عددی و عاری از داده‌های نول می‌باشند. مجموعه داده به‌صورت فایل CSV با استفاده از کتابخانه پانداز فراخوانی شده است. توزیع آماری داده‌ها، استخراج شاخصه‌های آماری و تشخیص داده‌های پرت توسط کتابخانه پانداز انجام‌گرفته است. شبکه طراحی‌شده از نوع سکوئنشال می‌باشد. تمامی داده‌های ورودی به شبکه با استفاده از کتابخانه سایکیت-لرن در زبان برنامه‌نویسی پایتون استانداردسازی شده‌اند. جهت ذخیره‌سازی و تعیین بهترین وزن‌های شبکه از کال بک modelcheckpoint استفاده‌شده است.



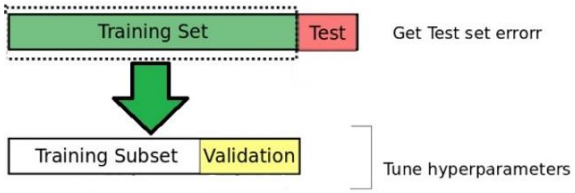

**شکل (۴):** تابع لیکی رلو.

## ۳- نتایج و بحث

در فرآیند پردازش مجموعه‌ی داده، در گام اول تعداد داده‌ها جهت یادگیری مدل ۴۹۳ عدد معادل ۸۵ درصد از کل مجموعه‌ی داده انتخاب شد و ۸۸ داده‌ی باقی‌مانده به‌عنوان داده‌های آزمایشی تعیین شد. سپس ۵۰ داده معادل ۱۰ درصد از داده‌های تعیین‌شده جهت یادگیری شبکه را به‌عنوان داده‌های ارزیاب انتخاب شد. در انتها، ۴۴۳ عدد داده جهت آموزش به مدل به‌عنوان داده‌ی آموزشی، وارد شبکه گردید و به‌وسیله‌ی ۸۸ داده‌ی آزمایشی عملکرد مدل ایجادشده ارزیابی شد (جدول **۱** و شکل **۵**).

### ۳-۱- معماری شبکه

معماری شبکه عصبی عمیق ارائه‌شده نتیجه بررسی تعداد زیادی از گونه‌های مختلف توابع فعال‌ساز، اعمال شیوه‌های مختلف در وزن دهی اولیه به لایه‌ها، تغییر در میزان نرخ یادگیری الگوریتم بهینه‌ساز، شیوه‌های مختلف مقیاس کردن داده‌ها، بررسی داده‌های پرت و همبستگی میان ویژگی‌های داده، جلوگیری از ایجاد بیش برازش، تغییر در تعداد لایه‌های مخفی، نورون‌ها و تعداد تکرار شبکه است که درمجموع بر اساس جمع‌بندی نتایج مدل ارائه‌شده شامل مشخصات جدول **۲** می‌باشد.




**شکل (۵):** شبیه‌سازی نحوه تقسیم‌بندی داده‌ها.

**جدول (۲):** معماری مدل شبکه‌ی عصبی.

| نوع فراپارامتر | تعداد یا نوع |
|---|---|
| **تعداد** لایه‌های **مخفی** | ۱۰ |
| **تعداد** نورون‌های مورداستفاده **در** لایه‌های **مخفی** | ۱۰۰، ۵۰، ۲۰ و ۱۰ |
| **نوع تابع** فعال‌ساز | لیکی رلو |
| **ضریب** لیک **در تابع** فعال‌ساز | ۰/۱ |
| **تابع هدف یا خطای مدل** | میانگین مربعات خطا |
| شاخص‌های **ارزیابی مدل** | خطای جذر میانگین مربعات، ضریب تعیین و زمان حل |
| **تعداد تکرار** | ۷۰۰ |
| **الگوریتم** بهینه‌ساز | آدام |
| **نرخ یادگیری الگوریتم** بهینه‌ساز | ۰/۰۰۰۱ |

### ۳-۲- فرآیند آموزش و ارزیابی مدل

عموماً در شبکه‌های عصبی عمیق فرآیند آموزش و ارزیابی به صورتی به کار می‌رود که به ما اجازه دهد در هر مرحله از فرآیند آموزش، میزان تمایز بین مقدار پیش‌بینی مدل با مقدار واقعی را مشخص کرد. یکی از متداول‌ترین روش‌ها به این صورت است که قسمتی از داده را به‌عنوان داده تست در نظر گرفته می‌شود و مدل تا انتهای آموزش به آن دسترسی ندارد. داده‌های آموزشی در هر مرحله به چند قسمت تقسیم می‌شوند و یک قسمت آن به‌عنوان داده‌ی ارزیابی و مابقی

به‌عنوان داده آموزشی انتخاب می‌شوند و پس از آموزش مدل، شاخص‌های ارزیابی محاسبه می‌گردد. در مرحله‌های بعدی آموزش، این داده‌های ارزیابی تغییر می‌کنند و بخش دیگری از داده‌ی آموزشی به‌عنوان داده ارزیابی انتخاب‌شده و این روند پیگیری می‌شود تا مشاهده گردد که فرآیند آموزش به چه سمتی پیش می‌رود و مدل دچار بیش برازش در آموزش نگردد. در فرآیند آموزش مدل، از محیط برنامه‌نویسی پایتون با استفاده از کتابخانه تنسور فلو و رابط برنامه‌نویسی کراس [۳۶] استفاده گردیده است. در این مدل با توجه به اینکه تعداد داده‌های موجود در ادبیات تحقیق زیاد نمی‌باشد، ما تمامی داده‌ها را به‌صورت یک باره به مدل ارائه و از ریز دسته کردن داده‌ها اجتناب شده است.

نتایج ارائه‌شده در این پژوهش به دو بخش کلی تقسیم می‌گردد. بخش اول به چالش جستجوی معماری شبکه می‌پردازد که در آن موارد مختلف یک شبکه اعم از تعداد لایه‌ها، تعداد نورون‌ها، توابع فعال‌ساز و توابع بهینه‌ساز و سایر فراپارامترها موردبررسی قرارگرفته و با توجه به شاخص‌های ارزیابی بهترین معماری مدل انتخاب می‌گردد. در بخش دوم با توجه به اینکه آزمایش و جمع‌آوری داده‌های تجربی در سالیان مختلف و توسط افراد متفاوت انجام پذیرفته است، این نتایج با پیش‌بینی‌های مدل مورد مقایسه قرارگرفته است. شایان‌ذکر است، روش‌های متفاوتی اعم از روابط تحلیلی و کلاسیک ریاضیاتی نیز اقدام به این پیش‌بینی‌ها نموده است. در اینجا بهبود عملکرد این مدل را در مقایسه با این تلاش‌ها سنجیده و میزان دقت بالای آن موردبررسی قرار داده‌شده است.

## ۳-۳- شناسایی فراپارامترها

قاعدتاً طراح شبکه برای ایجاد شبکه‌ای مناسب و بهینه نیاز به بررسی و مشاهده شبکه‌های ایجادشده از مجموعه مختلفی از فراپارامترها همچون نرخ یادگیری، تعداد نورون‌ها، تعداد لایه‌های مخفی و... را دارد تا بر اساس آن‌ها گزینه‌ی مناسب برای مجموعه داده‌ها را ارائه دهد. گریدسرچ سی‌وی از کتابخانه سایکیت-لرن در دسترس است و با تلفیق آن به‌وسیله رپرز با کتابخانه کراس می‌توان برای مدل طراحی‌شده کلاسی ایجاد و فراپارامترهای وابسته به عملکرد شبکه را در آن تعریف کرد. گریدسرچ سی‌وی با جستجوی تمام شبکه‌های قابل‌تعریف با فراپارامترهای تعیین‌شده مناسب‌ترین شبکه را با توجه به میزان دقت و خطای داده‌های ارزیاب پیشنهاد می‌دهد و طراح شبکه با مشاهده آن می‌تواند کران و بازه‌ی مناسبی برای تعداد فراپارامترهای بهینه مدل انتخاب کند. در مدل پیش رو با بهره‌مندی از این امکان، بازه‌ی مناسب برای هایپرپارامترهای مدل به‌صورت نرخ یادگیری در بازه اطراف عدد ۰/۰۰۰۱، تعداد نورون‌ها در حدود ۱۰۰ عدد و تعداد لایه‌های مخفی بین ۵ تا ۱۰ عدد ارائه شد. مدل ارائه‌شده‌ی این تحقیق با جستجو از میان فراپارامترهای پیشنهادی گریدسرچ سی‌وی و بررسی ویژگی‌های آماری و پیشینه داده-ها در ادبیات تحقیق انتخاب گردیده است. همچنین در این مدل از کال بک که ابزاری جهت ذخیره‌سازی پارامترهای شبکه در حین اجرا و در مراحل مختلف پردازش شبکه می-باشد، استفاده‌شده است.

یکی از انواع کال بک‌ها Modelcheckpoint است که هدف آن، ذخیره‌سازی وزن‌های مدل بعد از هر تکرار و خواندن داده‌های ورودی می‌باشد. این رویکرد جهت ذخیره‌سازی وزن‌ها و پارامترهای تعیین‌شده در مدل در مقابل مشکلات فنی احتمالی در هنگام فرآیند پردازش شبکه که ممکن است ساعت‌ها یا حتی دقایقی به طول انجامد ارائه گردیده است اما آنچه حائز اهمیت است، استفاده از ابزار پرکاربرد restore_best_weights جهت شناسایی بهترین و بهینه‌ترین پارامترها تا هر مرحله از آموزش شبکه برای دریافت، ذخیره-سازی و فراخوانی مدل آموزش‌دیده می‌باشد.

## ۴-۳- کارایی مدل

فرآیند اجرا و تعلیم شبکه با استفاده از پردازنده گرافیکی در حدود کمتر از ۳ دقیقه به طول انجامید. مدل ارائه‌شده‌ی تحقیق از میان ۸۸ داده‌ی آزمایشی، ۷۶ درصد از داده‌ها معادل ۶۷ عدد را در بازه خطای کمتر از ۱۰ درصد پیش‌بینی کرد که با توجه به محدودیت تعداد داده‌های موجود در پایگاه داده و ترکیب ۴ نوع فلز مختلف قابل‌توجه است. همچنین این مدل توانست ۸۸ درصد از داده‌ها معادل ۷۸ عدد را در بازه خطای کمتر از ۲۰ درصد پیش‌بینی کند. لازم به ذکر است داده‌های این تحقیق توسط الگوریتم‌های سنتی یادگیری

ماشین نیز بررسی شد که نتیجه مناسب را در برنداشت و استفاده از شبکه عصبی عمیق به دلیل عملکرد نسبتاً ضعیف مدل‌های سنتی یادگیری ماشین و کاربرد آن در پردازش داده‌های با حجم بالا در نظر گرفته شد.

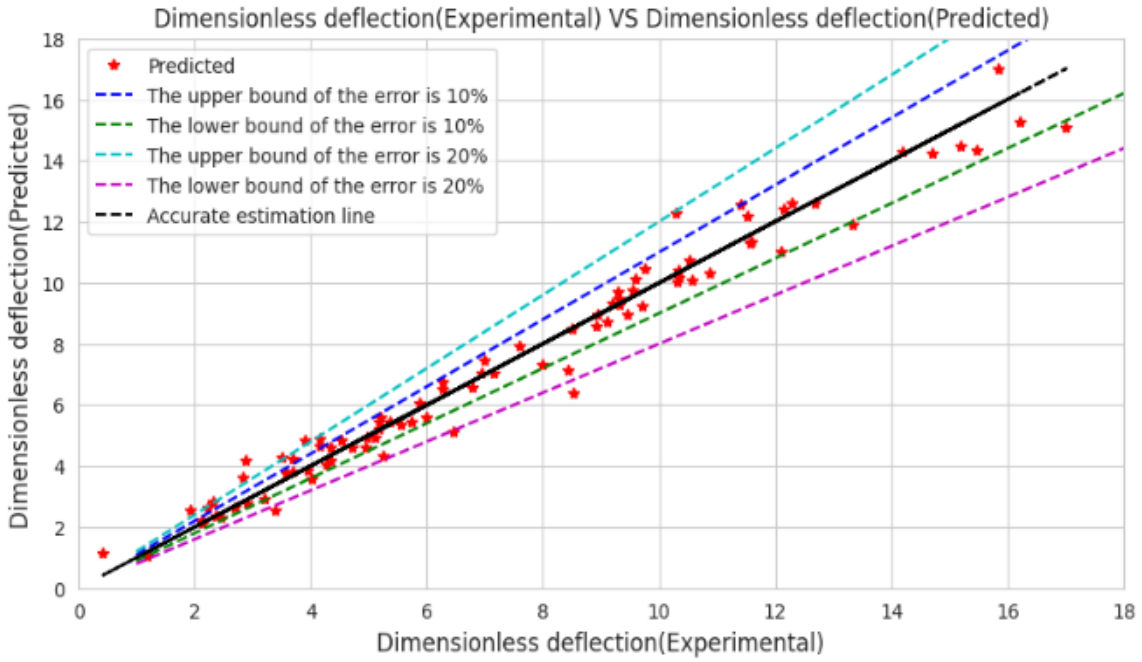
در شکل **۶** میزان خیز دائمی ورق بر اساس مقدار واقعی موجود حاصل از انجام آزمایش‌های تجربی و مقدار پیش‌بینی‌شده توسط مدل نمایش داده‌شده است که داده‌های قرارگرفته روی خط سیاه‌رنگ نمایانگر برابری مقدار واقعی و پیش‌بینی‌شده توسط مدل می‌باشد و پراکندگی داده‌های اطراف خط سیاه‌رنگ میزان اختلاف میان مقدار پیش‌بینی‌شده و مقدار واقعی را نشان می‌دهد. گاهی عملکرد یک شبکه و میزان دقت آن در پیش‌بینی‌ها می‌تواند وابسته به خطای موجود در انجام آزمایش‌های تجربی و تولید داده‌ها باشد که اجتناب‌ناپذیر است. در این‌گونه موارد می‌توان یا تفکیک داده‌های خروجی شبکه و بررسی مراجع و شیوه‌ی تولید آن داده تحلیلی دقیق‌تر از عملکرد و دقت شبکه داشت (شکل **۷**).

داده‌های آموزشی و آزمایشی در این مدل دارای توزیع همسان بوده و تجمیع داده‌های آزمایشی در ناحیه‌ی خاصی از نمونه‌های آموزشی صورت نگرفته است. همچنین طیف خاصی از فلزهای یکسان برای سنجش عملکرد مدل به کار نرفته و از میان انواع فلزات موجود در فرآیند آموزش شبکه نمونه آزمایشی در میان داده‌های تست وجود دارد. شکل **۸** گویای این دو ادعا است.

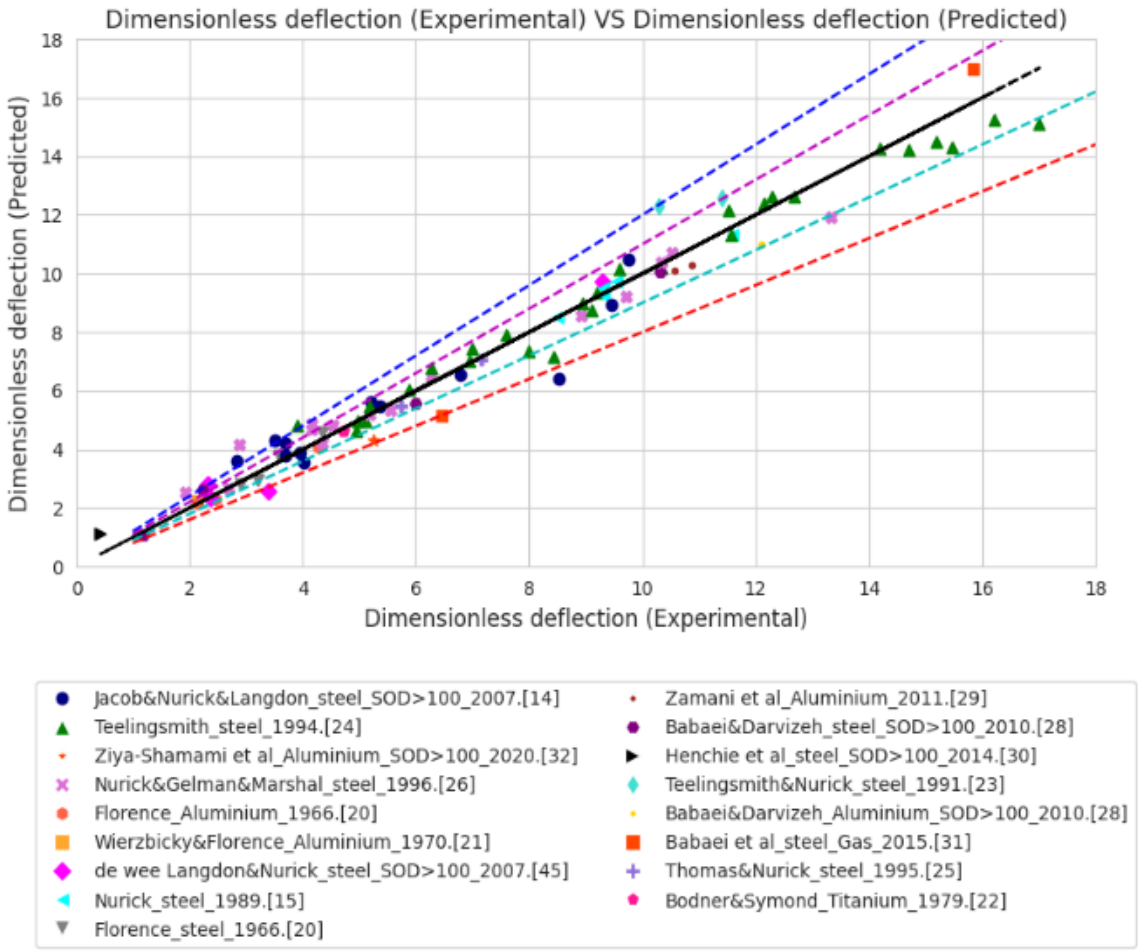
در شبکه طراحی‌شده نشانه و شاخصی مبنی بر وقوع بیش برازش و کم برازش مدل دریافت نکردیم و سیر نزولی و مشابه دو نمودار خطای آموزشی و آزمایشی در شکل **۹** گویای این ادعا است.

همچنین مقایسه‌ی تصویری میزان دقت و خطای شبکه در پیش‌بینی میزان خیز به ازای ضخامت آن ورق در شکل **۱۰** نشان داده‌شده است.

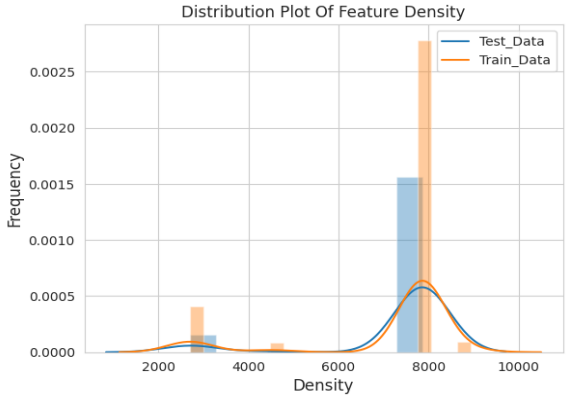
میزان پراکندگی و توزیع داده‌های آزمایشی و آموزشی به ازای خیز ورق نیز به‌صورت دو شکل **۱۱** و **۱۲** قابل‌مشاهده است که تصادفی بودن انتخاب داده‌های آزمایشی نسبت به داده‌های آزمایشی را با وضوح بیشتری نشان می‌دهد.



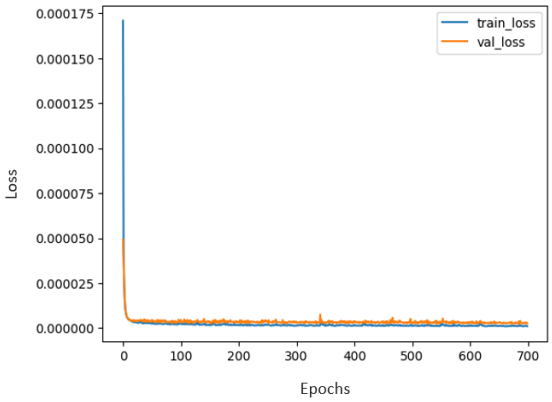

**شکل (۶):** پراکندگی میزان خیز ورق در محدوده خطای ۱۰ و ۲۰ درصد.



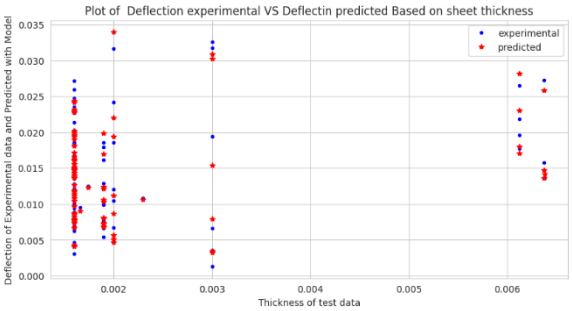

**شکل (۷):** پراکندگی میزان خیز ورق در محدوده خطای ۱۰ و ۲۰ درصد با ذکر مرجع.



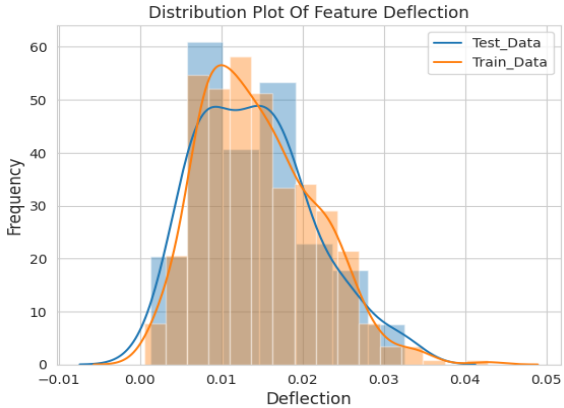

**شکل (۸):** فراوانی داده‌های آزمایشی و آموزشی برحسب چگالی.



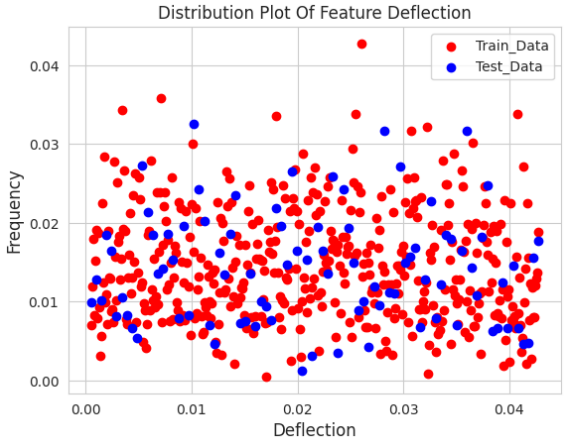

**شکل (۹):** نمودار خطای شبکه به ازای داده‌های آموزشی و ارزیاب برحسب تعداد تکرار الگوریتم.




**شکل (۱۰)**: پیش‌بینی میزان خیز به ازای ضخامت ورق.




**شکل (۱۱)**: فراوانی داده‌های آزمایشی و آموزشی برحسب میزان خیز.




**شکل (۱۲)**: پراکندگی داده‌های آموزشی و آزمایشی برحسب میزان خیز.

### ۳-۵- مقادیر شاخص‌های ارزیابی

جهت ارزیابی و سنجش عملکرد مدل‌های مبتنی بر یادگیری ماشین در مسائل رگرسیون سه شاخص مهم آماری که پیشتر شیوه‌ی محاسبه‌ی آن ارائه شد، وجود دارد و در مدل طراحی‌شده در این تحقیق مقادیر جدول **۳** را به خود اختصاص داده‌اند. جدول **۴** مقایسه‌ی میزان شاخص خطای خطای جذر میانگین مربعات را به ازای نتایج حاصل از مدل-های موجود در ادبیات تحقیق نشان می‌دهد که در مطالعه‌ی حاضر این شاخص بیش از ۱۰۰ برابر نسبت به آخرین میزان خطای به‌دست‌آمده در سال ۲۰۲۰ کاهش پیدا کرد.

**جدول (۳)**: شاخص‌های ارزیابی مدل.

| | |
|---|---|
| ضریب تعیین | ۰/۹۶ |
| $D_2$ | ۰/۸۳ |
| خطای جذر میانگین مربعات | ۰/۰۰۱۲ |
| زمان حل | ۱۴۴ ثانیه |
| پردازشگر | GPU |
| پارامترها | ۶۷۶۹۱ |

**جدول (۴)**: مقایسه مقدار شاخص میانگین مربعات خطا با روابط پیشین.

| مرجع | محققان | سال | RMSE |
|---|---|---|---|
| [۳۷] | جونز | ۱۹۷۱ | ۷/۷ |
| [۳۸] | گودس سوارز | ۱۹۸۱ | ۶/۹۴ |
| [۳۹] | کالادین | ۱۹۸۳ | ۵/۴۲ |
| [۴۲] | پرون و باهادرا | ۱۹۸۴ | ۲/۱ |
| [۳۷] | جونز | ۱۹۷۱ | ۱/۷ |
| [۴۳] | کولوت و نوریک | ۲۰۱۴ | ۱/۵۱ |
| [۴۸] | بابایی و درویزه | ۲۰۱۰ | ۱/۲۷ |
| [۴۴] | لیپمن | ۱۹۷۴ | ۱/۲۶ |
| [۱۵] | نوریک و مارتین | ۱۹۸۹ | ۰/۹۹ |
| [۴۷] | بابایی و درویزه | ۲۰۱۲ | ۰/۸۹ |
| [۸] | یوان و همکاران | ۲۰۱۷ | ۰/۹۷ |
| [۳] | مستوفی و همکاران | ۲۰۲۰ | ۰/۸۴ |
| | مطالعه‌ی حاضر | ۲۰۲۳ | ۰/۰۰۱۲ |

## ۴- نتیجه‌گیری

در این مقاله، با ایجاد یکی از مدل‌های مبتنی بر یادگیری عمیق، یک پیش‌بینی کاربردی از میزان خیز ورق‌های فلزی دایره‌ای تحت بار انفجاری یکنواخت انجام شده است که از مهم‌ترین نتایج آن می‌توان به موارد زیر اشاره کرد:

۱) کاهش میزان خطا و زمان آموزش مدل و افزایش دقت و سرعت آموزش مدل نسبت به روابط سنتی و تحلیلی پیشین. اگرچه تولید داده‌های مصنوعی با استفاده از روابط آماری از روی داده‌های تجربی موجود جهت آموزش بیشتر و بهتر مدل با ویژگی‌های نهفته و جزئی موجود در دل داده‌ها روشی مرسوم جهت افزایش کارایی مدل و قدرت تحلیل و پیش‌بینی آن است، اما در تحقیق پیش رو به جهت امکان بروز مشکلات علمی و رعایت انواع بارگذاری‌ها جهت مقایسه و استناد و ارجاع به مقالات پیشینه پژوهش از آن صرف‌نظر گردید اما در پژوهش‌های آینده جهت ایجاد یک مدل جامع و مرجع می‌توان از این شیوه بهره برد.

۲) افزایش سرعت پردازش و ساخت مدل برای داده‌های با تعداد بالا. دسترسی به پردازنده گرافیکی در ساخت این مدل الزامی نبود و زمان محاسبه تفاوت زیادی با پردازنده‌های مرکزی نداشت اما در صورت افزایش مجموعه داده استفاده از پردازنده‌های گرافیکی رایگان از طریق دسترسی آنلاین پیشنهاد می‌گردد که مناسب و کاربردی است.

۳) استفاده از ویژگی نرخ کرنش در مدل‌سازی و فرآیند تعلیم شبکه.

۴) ارزیابی پاسخ فلزات مختلف به‌وسیله یک مدل واحد و یکسان که به ازای ۴ نوع فلز مختلف در مجموعه داده توانستیم ۷۶ درصد از داده‌های آزمایشی را در محدوده خطای کمتر از ۱۰ درصد و ۸۸ درصد را در محدوده خطای کمتر از ۲۰ درصد پیش‌بینی می‌کند.

۵) عدم وقوع بیش‌برازش یا کم‌برازش در حین آموزش شبکه. همچنین در این مدل به دلیل عدم دسترسی به مجموعه داده با حجم زیاد و کاهش تعداد داده‌ها، حذف داده‌های پرت باعث ایجاد تأثیر منفی در یادگیری شبکه شد که از حذف داده‌های پرت صرف‌نظر گردید.

۶) ایجاد تعادل میان انواع داده‌ها یکی از موارد تأثیرگذار در بهبود عملکرد شبکه است. در مدل ارائه‌شده‌ی این تحقیق میزان داده‌های مربوط به فلزات تیتانیوم و مس نسبت به دو فلز دیگر فولادی و آلومینیوم کمتر بود که در میان داده‌های تست نیز سهم کمتری به خود اختصاص دادند. افزودن یا حذف داده‌ها جهت ایجاد تعادل به‌عنوان یک راهکار در میان طراحان شبکه‌های عصبی مرسوم است اما در این تحقیق ازآنجاکه سهم ناچیزی از داده‌های تست را فلزات کم تعداد در برگرفته‌اند از انجام آن صرف‌نظر شد هرچند که در فرآیند تعلیم شبکه حضور داشتند و مدل بر اساس آن‌ها تعلیم‌دیده است اما حذف آن‌ها از فرآیند آموزش شبکه نیز تغییر محسوسی در بهبود شاخص‌های ارزیابی ایجاد نکرد.

## ۵- مراجع

[1] Sakaridis E, Karathanasopoulos N, Mohr D. Machine-learning based prediction of crash response of tubular structures. International Journal of Impact Engineering. 2022;166:104240. **DOI** https://doi.org/10.1016/j.ijimpeng.2022.104240.

[2] Mostofi TM, Babaei H, Alitavoli M. Theoretical analysis on the effect of uniform and localized impulsive loading on the dynamic plastic behaviour of fully clamped thin quadrangular plates. Thin-Walled Structures. 2016;109:367-76. **DOI** https://doi.org/10.1016/j.tws.2016.10.009.

[3] Mirzababaie Mostofi T, Sayah Badkhor M, Babaei H. A study on plastic response of circular plates under uniformly and locally distributed dynamic loading. Journal of Solid and Fluid Mechanics. 2020;10(2):11-27. **DOI** https://doi.org/10.22044/JSFM.2020.9264.3091.

[4] Babaei H, Mostofi TM, Alitavoli M, Darvizeh A. Empirical modelling for prediction of large deformation of clamped circular plates in gas detonation forming process. Experimental Techniques. 2016;40(6):1485-94. **DOI** https://doi.org/10.1007/s40799-016-0063-3.

[5] Mostofi TM, Babaei H, Alitavoli M. The influence of gas mixture detonation loads on large plastic deformation of thin quadrangular plates: Experimental investigation and empirical modelling. Thin-Walled

Structures. 2017;118:1-11. **DOI** https://doi.org/10.1016/j.tws.2017.04.031.

[6] Mostofi TM, Babaei H, Alitavoli M, Lu G, Ruan D. Large transverse deformation of double-layered rectangular plates subjected to gas mixture detonation load. International Journal of Impact Engineering. 2019;125:93-106. **DOI** https://doi.org/10.1016/j.ijimpeng.2018.11.005.

[7] Nurick G, Martin J. Deformation of thin plates subjected to impulsive loading—a review: Part i: Theoretical considerations. International Journal of Impact Engineering. 1989;8(2):159-70. **DOI** https://doi.org/10.1016/0734-743X(89)90015-8.

[8] Yuen SCK, Nurick G, Langdon G, Iyer Y. Deformation of thin plates subjected to impulsive load: Part III–an update 25 years on. International Journal of Impact Engineering. 2017;107:108-17. **DOI** https://doi.org/10.1016/j.ijimpeng.2016.06.010.

[9] Mirzababaie Mostofi T, Sayah Badkhor M, Ghasemi E. Experimental investigation and optimal analysis of the high-velocity forming process of bilayer plates. Journal of Solid and Fluid Mechanics. 2019;9(3):65-80. **DOI** https://doi.org/10.22044/JSFM.2019.8586.2953.

[10] Haghgoo M, Babaei H, Mostofi TM. 3D numerical investigation of the detonation wave propagation influence on the triangular plate deformation using finite rate chemistry model of LS-DYNA CESE method. International Journal of Impact Engineering. 2022;161:104108. **DOI** https://doi.org/10.1016/j.ijimpeng.2021.104108.

[11] Nasiri S, Sadegh-Yazdi M, Mousavi S, Ziya-Shamami M, Mostofi T. Repeated underwater explosive forming: Experimental investigation and numerical modeling based on coupled Eulerian–Lagrangian approach. Thin-Walled Structures. 2022;172:108860. **DOI** https://doi.org/10.1016/j.tws.2021.108860.

[12] Babaei H, Mirzababaie Mostofi T. Modeling and prediction of fatigue life in composite materials by using singular value decomposition method. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications. 2020;234(2):246-54. **DOI** https://doi.org/10.1177/1464420716660875.

[13] Nurick GN, Martin JB. Deformation of thin plates subjected to impulsive loading—a review Part II: Experimental studies. International Journal of Impact Engineering. 1989;8(2):171-86. **DOI** https://doi.org/10.1016/0734-743X(89)90015-8.

[14] Jacob N, Nurick G, Langdon G. The effect of stand-off distance on the failure of fully clamped circular mild steel plates subjected to blast loads. Engineering Structures. 2007;29(10):2723-36. **DOI** https://doi.org/10.1016/j.engstruct.2007.01.021.

[15] Nurick G, Martin J. Deformation of thin plates subjected to impulsive loading—a review part II: experimental studies. International journal of impact engineering. 1989;8(2):171-86. **DOI** https://doi.org/10.1016/0734-743X(89)90015-8.

[16] Goodfellow I, Bengio Y, Courville A. Deep learning: MIT press; 2016.

[17] Bishop CM, Nasrabadi NM. Pattern recognition and machine learning: Springer; 2006.

[18] Bengio Y. Practical recommendations for gradient-based training of deep architectures. InNeural networks: Tricks of the trade: Second edition 2012 May (pp. 437-478). Berlin, Heidelberg: Springer Berlin Heidelberg. **DOI** https://doi.org/10.1007/978-3-642-35289-8_26.

[19] Zhang A, Lipton ZC, Li M, Smola AJ. Dive into deep learning: Cambridge University Press; 2023. **DOI** https://doi.org/10.48550/arXiv.2106.11342.

[20] Florence A. Circular plate under a uniformly distributed impulse. International Journal of Solids and Structures. 1966;2(1):37-47. **DOI** https://doi.org/10.1016/0020-7683(66)90005-9.

[21] Wierzbicki T, Florence A. A theoretical and experimental investigation of impulsively loaded clamped circular viscoplastic plates. International Journal of solids and structures. 1970;6(5):553-68. **DOI** https://doi.org/10.1016/0020-7683(70)90030-2.

[22] Bodner SR, Symonds PS. Experiments on viscoplastic response of circular plates to impulsive loading. Journal of the Mechanics and Physics of Solids. 1979;27(2):91-113. **DOI** https://doi.org/10.1016/0022-5096(79)90013-9.

[23] Teeling-Smith R, Nurick G. The deformation and tearing of thin circular plates subjected to impulsive loads. International Journal of Impact Engineering. 1991;11(1):77-91. **DOI** https://doi.org/10.1016/0734-743X(91)90032-B.

[24] Nurick GN, Teeling-Smith RG. Predicting the onset of necking and hence rupture of thin plates loaded impulsively-an experimental view. InStructures Under Shock and Impact II: Proceedings of the Second International Conference, held in Portsmouth, UK, 16th–18th June, 1992 1992 (pp. 431-445). Thomas Telford Publishing.

[25] Thomas B. The effect of boundary conditions on thin plates subjected to impulsive loads. Dynamic Plasticity and Structural Behaviours. 1995;85.

[26] Nurick G, Gelman M, Marshall N. Tearing of blast loaded plates with clamped boundary conditions. International Journal of Impact Engineering.

1996;18(7-8):803-27. **DOI** https://doi.org/10.1016/S0734-743X(96)00026-7.

[27] Nurick G, Lumpp D. Deflection and tearing of clamped stiffened circular plates subjected to uniform impulsive blast loads. WIT Transactions on The Built Environment. 1970;25. **DOI** https://doi.org/10.2495/SUSI960361.

[28] Gharababaei H, Darvizeh A. Experimental and Analytical Investigation of Large Deformation of Thin Circular Plates Subjected to Localized and Uniform Impulsive Loading#. Mechanics Based Design of Structures and Machines. 2010;38(2):171-89. **DOI** https://doi.org/10.1080/15397730903554633.

[29] Zamani J, Safari K, Ghamsari A, Zamiri A. Experimental analysis of clamped AA5010 and steel plates subjected to blast loading and underwater explosion. The Journal of Strain Analysis for Engineering Design. 2011;46(3):201-12. **DOI** https://doi.org/10.1177/0309324710396601.

[30] Henchie TF, Yuen SCK, Nurick G, Ranwaha N, Balden V. The response of circular plates to repeated uniform blast loads: An experimental and numerical study. International Journal of Impact Engineering. 2014;74:36-45. **DOI** https://doi.org/10.1016/j.ijimpeng.2014.02.021.

[31] Babaei H, Mostofi TM, Sadraei SH. Effect of gas detonation on response of circular plate-experimental and theoretical. Structural Engineering and Mechanics. 2015;56(4):535-48. **DOI** https://doi.org/10.12989/sem.2015.56.4.535.

[32] Ziya-Shamami M, Babaei H, Mostofi TM, Khodarahmi H. Structural response of monolithic and multi-layered circular metallic plates under repeated uniformly distributed impulsive loading: An experimental study. Thin-Walled Structures. 2020;157:107024. **DOI** https://doi.org/10.1016/j.tws.2020.107024.

[33] Souza ML, da Costa CA, de Oliveira Ramos G. A machine-learning based data-oriented pipeline for Prognosis and Health Management Systems. Computers in Industry. 2023;148:103903. **DOI** https://doi.org/10.1016/j.compind.2023.103903.

[34] Ketkar N, Santana E. Deep learning with Python. Berkeley, CA: Apress; 2017. **DOI** https://doi.org/10.1007/978-1-4842-5364-9.

[35] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014. **DOI** https://doi.org/10.48550/arXiv.1412.6980.

[36] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467. 2016. **DOI** https://doi.org/10.48550/arXiv.1603.04467.

[37] Jones N. A theoretical study of the dynamic plastic behavior of beams and plates with finite-deflections. International Journal of Solids and Structures. 1971;7(8):1007-29. **DOI** https://doi.org/10.1016/0020-7683(71)90078-3.

[38] Soares CG. A mode solution for the finite deflections of a circular plate loaded impulsively. Engineering Transactions. 1981;29(1):99-114.

[39] Mostofi TM, Babaei H, Alitavoli M, Hosseinzadeh S. On dimensionless numbers for predicting large ductile transverse deformation of monolithic and multi-layered metallic square targets struck normally by rigid spherical projectile. Thin-Walled Structures. 2017;112:118-24. **DOI** https://doi.org/10.1016/j.tws.2016.12.014.

[40] Duffey TA. Large deflection dynamic response of clamped circular plates subjected to explosive loading. Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); 1967.

[41] Babaei H, Mirzababaie Mostofi T. New dimensionless numbers for deformation of circular mild steel plates with large strains as a result of localized and uniform impulsive loading. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications. 2020;234(2):231-45. **DOI** https://doi.org/10.1177/1464420716654195.

[42] Babaei H, Mirzababaie Mostofi T, Armoudli E. On dimensionless numbers for the dynamic plastic response of quadrangular mild steel plates subjected to localized and uniform impulsive loading. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 2017;231(5):939-50. **DOI** https://doi.org/10.1177/0954408916650713.

[43] Cloete T, Nurick G. On the influence of radial displacements and bending strains on the large deflections of impulsively loaded circular plates. International Journal of Mechanical Sciences. 2014;82:140-8. **DOI** https://doi.org/10.1016/j.ijmecsci.2014.02.026.

[44] Lippmann H. Kinetics of the axisymmetric rigid-plastic membrane subject to initial impact. International Journal of Mechanical Sciences. 1974;16(5):297-303. **DOI** https://doi.org/10.1016/0020-7403(74)90046-0.

[45] C.K. De Wee, G.S. Langdon, G.N. Nurick. The influence of confined and unconfined explosions on the structural response of fully clamped mild steel plates. Undergraduate thesis, University of Cape Town, 2007.

[46] Mirzababaie Mostofi T, Babaei H. Plastic deformation of polymeric-coated aluminum plates subjected to gas mixture detonation loading: Part II: Analytical and empirical modelling. Journal of Solid and Fluid Mechanics. 2019;9(2):15-29. **DOI** https://doi.org/10.22044/JSFM.2019.7816.2778.

[47] Babaei H, Darvizeh A. Analytical study of plastic deformation of clamped circular plates subjected to impulsive loading. Journal of Mechanics of Materials and Structures. 2012;7(4):309-22. **DOI** https://doi.org/10.2140/jomms.2012.7.309.

[48] Gharababaei H, Darvizeh A, Darvizeh M. Analytical and experimental studies for deformation of circular plates subjected to blast loading. Journal of mechanical Science and Technology. 2010;24:1855-64. **DOI** https://doi.org/10.1007/s12206-010-0602-2.