

An Enhanced LSTM Method to Improve the Accuracy of the Business Process Prediction

Mohammad Hasan Adalat^{*}, Reza Azmi^{},
Jafar Bagherinejad^{***}**

Abstract

Prediction of the process behavior plays a key role in business process management. This research benefits from recent development in the field of deep learning to predict the next event in business processes. The proposed method uses Long Short-Term Memory (LSTM) as a promising architecture of recurrent neural networks. This architecture is implemented using a number of configurations with the aim of investigating how each of them affects the performance of the prediction models. In order to build and evaluate our prediction models, we used two publicly available datasets (BPI 2012 and BPI 2017). After developing 300 prediction models, the results indicated that the proposed method outperforms the state-of-the-art methods in terms of precision. The best result in terms of Accuracy (0.907) was achieved through “one-hidden” layer LSTM architecture and by using “Big” configuration in the absence of “feedback”.

Keywords: Process Management; Prediction Model; Machine Learning; LSTM Architecture.

Received: Dec. 24, 2018, Accepted: June 07, 2020.

* Department of Computer Engineering, Islamic Azad University Qom, Qom, Iran.

** Department of Industrial Engineering, Alzahra University.

*** Department of Computer Engineering, Alzahra University.

چشم‌انداز مدیریت صنعتی

شاپای چاپی: ۹۸۷۴-۲۲۵۱، شاپای الکترونیکی: ۴۱۶۵-۲۶۴۵

سال دهم، شماره ۳۹، پاییز ۱۳۹۹، صص ۷۱-۹۷ (نوع مقاله: پژوهشی)

DOI: [10.52547/JIMP.10.3.71](https://doi.org/10.52547/JIMP.10.3.71)

بهبود دقت پیش‌بینی فرآیندها در مدیریت فرآیندهای کسب‌وکار با به‌کارگیری معماری LSTM

محمدحسن عدالت*، رضا عزمی**، جعفر باقری‌نژاد***

چکیده

پیش‌بینی رفتار فرآیندهای سازمانی، نقش مهمی در مدیریت فرآیندهای کسب‌وکار ایفا می‌کند. این مهم با توسعه به‌کارگیری الگوریتم‌های یادگیری ماشین در جنبه‌های مختلف آینده‌پژوهی افق‌های نوینی در برابر پیش‌بینی رویدادها و فرآیندها در فضای کسب‌وکار گشوده است. یکی از روش‌های یادگیری ماشین، به‌کارگیری الگوریتم‌های یادگیری عمیق به‌عنوان شاخه‌ای از شبکه‌های عصبی است که توانسته دقت پیش‌بینی را به میزان زیادی افزایش دهد؛ از این‌رو در پژوهش حاضر از معماری LSTM (حافظه طولانی کوتاه‌مدت) شبکه عصبی برای پیش‌بینی فرآیندهای کسب‌وکار استفاده شده است. برای انجام آزمایش، الگوریتم LSTM بر روی مجموعه داده BPI2012 و BPI2017 اعمال شد. نتایج حاصل از ساخت مدل پیش‌بینی نشان داد که در مجموعه داده BPI2017 از مجموع آزمایش‌های انجام‌شده بیشترین دقت ۰/۹۰۷ است که این مقدار دقت از مقادیر دقت به‌دست‌آمده در پژوهش‌های مشابه بالاتر است. این دقت با اجرای الگوریتم LSTM با معماری یک‌لایه و مدل داده‌بزرگ و بدون بازخورد به‌دست آمده است.

کلیدواژه‌ها: مدیریت فرآیندها؛ مدل پیش‌بینی؛ یادگیری ماشین؛ معماری LSTM.

۱. مقدمه

تاریخ دریافت مقاله: ۱۳۹۷/۱۰/۰۳، تاریخ پذیرش مقاله: ۱۳۹۹/۰۳/۱۸.
* دانشجوی دکتری، دانشکده فنی مهندسی، دانشگاه آزاد اسلامی قم.
** دانشیار، دانشگاه الزهرا (نویسنده مسئول).

E-mail: azmi@alzahra.ac.ir

*** دانشیار، دانشگاه الزهرا.

مدیریت فرآیند کسب‌وکار^۱ (BPM) حیطة‌ای از علم مدیریت است که شامل روش‌ها، ابزارها و فنونی می‌شود که طراحی، تجزیه‌وتحلیل، بهینه‌سازی و مدیریت اجرایی فرآیندهای کسب‌وکار را پشتیبانی می‌کند [۱]. مدیریت کسب‌وکار پیوسته در حال ارزیابی فرآیندها برای بهبود اثربخشی و کارایی آن‌ها درون یک سازمان است تا به سه هدف عمده دست یابد: ۱. کاهش هزینه‌ها در سازمان؛ ۲. بهبود کیفیت خدمات‌دهی و ۳. افزایش بهره‌وری و توان رقابتی در مقایسه با سایر سازمان‌های رقیب در حوزه‌های مشترک عملکردی. از این رو سازمان‌ها در تکاپو برای به‌کارگیری نرم‌افزارهای BPM به‌عنوان یک ابزار بهبود مستمر در درون سازمان خود هستند.

یکی از ویژگی‌های به‌کارگیری روش‌های نوین مدیریت کسب‌وکار، توانمندسازی سازمان در پیش‌بینی فرآیندهای آینده است که می‌تواند سازمان را در مواجهه با سناریوهای از پیش شناخته‌شده آماده کند و در تعدیل ریسک کسب‌وکار مؤثر باشد [۵]. روش‌های متداول در پیش‌بینی فرآیندهای سازمانی از داده‌های مربوط به فرآیندهای گذشته و رویدادهای ثبت‌شده برای استخراج اطلاعاتی درباره فرآیندهای آینده استفاده می‌کند [۱۵].

برای نیل به این مقصود، الگوریتم‌های گوناگون داده‌کاوی بر روی تاریخچه فرآیندهای سازمانی اجرا می‌شود تا اطلاعات سودمندی درباره آینده آن‌ها استخراج شود. از میان الگوریتم‌های داده‌کاوی، درخت‌های تصمیم‌گیری، روش‌های خوشه‌بندی و شبکه‌های عصبی برای پیش‌بینی فرآیندهای کسب‌وکار بیشتر مورد توجه قرار گرفته‌اند [۱۴، ۷، ۲۱]. در این پژوهش از کاربرد شبکه‌های عصبی عمیق در پردازش زبان طبیعی، استفاده خواهد شد تا بتوان با استفاده از رویدادهای متن پایه‌ای که از سیستم‌های مدیریت فرآیندهای کسب‌وکار به‌دست می‌آید، رویدادهایی که در آینده به‌وقوع خواهند پیوست را پیش‌بینی کرد.

یکی از مشکلات شبکه‌های عصبی بازگشتی، ناتوانی آن‌ها در یادگیری وابستگی بلندمدت است. شبکه‌های عصبی بازگشتی قادر نیستند تا بین الگوهایی که در فاصله زمانی دورتری از داده کشف کرده‌اند و کاری که اکنون در حال انجام آن هستند، رابطه برقرار کنند. اگر بتوان میان یافته‌های یک شبکه عصبی در گذشته دور و فعالیت جاری آن پل زد، آنگاه هم‌افزایی ایجاد می‌شود و دقت و کارایی کشف الگوهای تازه افزایش خواهد یافت. برای نیل به این هدف، هاکریتز و همکاران (۱۹۹۷)، معماری را که برای شبکه‌های عصبی بازگشتی پیشنهاد کردند به معماری حافظه طولانی کوتاه‌مدت^۲ معروف شد و قادر بود تا نقصان فراموشی ذکر شده را جبران کند [۱۱]. در این پژوهش سعی شده است با به‌کارگیری شبکه عصبی با معماری حافظه طولانی کوتاه‌مدت، مدل پیش‌بینی فرآیندهای کسب‌وکار بر اساس مجموعه داده‌های BPI 2012 و

1. Business Process Management

2. Long short-term memory

BPI 2017 تولید شود. در این پژوهش با انجام چندین آزمایش، بررسی‌هایی پیرامون پیکربندی‌ها مختلف (۵ پیکربندی)، تعداد لایه‌های مختلف (یک لایه و دو لایه) و بازخورد (فرآیندهای ورودی دارای بازخورد و بدون بازخورد) بر روی تولید مدل پیش‌بینی فرآیندها انجام شد. در گزارش آزمایش‌های صورت‌گرفته، دقت بالاتر نسب به کارهای پیشین که در جدول ۱، آمده است، به دست آمد.

افزایش دقت در تشخیص فرآیندهای سازمانی موجب افزایش بهره‌وری سازمان، کسب آمادگی‌های حداکثری برای مواجهه با بحران‌های پیش‌رو، بهبود اثربخشی و کارایی مدیریت فرآیندهای سازمانی، بهبود رضایت خدمت‌گیرندگان و در نهایت توانمندسازی کسب‌وکار می‌شود. این مقاله شامل ۵ بخش است: بخش دوم مبانی نظری و پیشینه پژوهش می‌باشد که به مبانی شبکه‌های عصبی می‌پردازد و ضمن معرفی آن، به سیر تحول شبکه‌های عصبی و معماری‌های آن اشاره دارد. در ادامه به مبانی نظری مروری و پژوهش‌های صورت‌گرفته در حوزه پیش‌بینی فرآیندهای سازمانی پرداخته می‌شود. در این بخش، روش‌های گوناگون آماری نظیر مدل پنهانی مارکوف، و روش‌های داده‌کاوی، نظیر یافتن نزدیک‌ترین همسایه که در پیش‌بینی فرآیندها توسط پژوهشگران مورد استفاده قرار گرفته است، بررسی شده و دقت پیش‌بینی آن‌ها مورد مطالعه قرار گرفته است. بخش سوم به روش‌شناسی پژوهش اختصاص یافته است. بخش چهارم تحلیل داده‌ها و یافته‌های پژوهش می‌باشد. این بخش به تشریح ویژگی‌ها و تنظیمات آزمایش‌های صورت‌گرفته بر روی مجموعه‌های داده اختصاص دارد. در این بخش سعی شده است تا آزمایش‌ها به گونه‌ای طراحی شوند تا بتواند تصویری گویا از عملکرد معماری LSTM^۱ در پیش‌بینی فرآیندها ارائه دهد. بخش پنجم به نتیجه‌گیری و ارائه پیشنهادها می‌پردازد و به تحلیل نتایج ۳۰ آزمایش انجام شده پرداخته خواهد شد تا تأثیر تعداد لایه، اندازه مدل و بازخورد در دقت مدل نهایی و سنجیده شود.

۲. مبانی نظری و پیشینه پژوهش

یادگیری عمیق برای پیش‌بینی فرآیندها

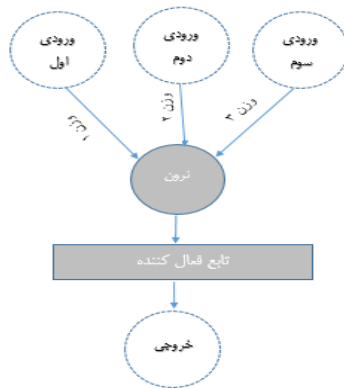
مقدمه‌ای بر شبکه عصبی. شبکه عصبی نوع خاصی از نمودار محاسباتی غیرمدور است که شامل سلول‌هایی می‌شود که این سلول‌ها بر روی یک لایه ورودی، یک یا چند لایه میانی (به آن لایه مخفی نیز می‌گویند) و یک لایه خروجی توزیع شده‌اند. سلول‌ها که نقش پردازش‌گر را بازی می‌کنند در هر لایه توسط رابط‌های وزن داری به سلول‌های لایه‌های پیش و پس خود متصل شده‌اند که این امر باعث می‌شود معماری‌های مختلفی از یک شبکه به وجود بیاید. شبکه عصبی

1. Long Short Time Memory

یک روش یادگیری با نظارت^۱ است [۱۹]. در این شبکه خروجی هر سلول تابعی از مجموع وزن‌های ورودی‌های آن سلول است که به صورت زیر محاسبه می‌شود:

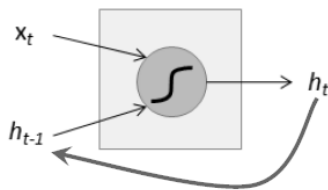
$$f(X, W) = \phi(X \cdot W) = \phi\left(\sum_{i=1}^p x_i \cdot w_i\right) \quad \text{رابطه (۱)}$$

در رابطه بالا x_i ها ورودی‌های یک نرون و w_i ها وزن‌های متناظر با هر یک از این ورودی‌ها است. ϕ تابع فعال‌کننده نرون نام دارد که به صورت پیش‌فرض می‌تواند یک تابع تانژانت هایپربولیک یا یک تابع سیگموئید باشد. شکل ۱، ساختار یک نرون با سه ورودی را در شبکه عصبی نشان می‌دهد.، ایجاد مدل‌های پیش‌بینی هزینه و زمان نهایی پروژه‌ها، پیش‌بینی تقاضاهای متلاطم در زنجیره تأمین و به‌کارگیری آن در فناوری‌های گروهی و مقایسه‌ای از کاربردهای شبکه‌های عصبی است [۱، ۲۳، ۱۸].



شکل ۱. ساختار یک نرون در شبکه عصبی

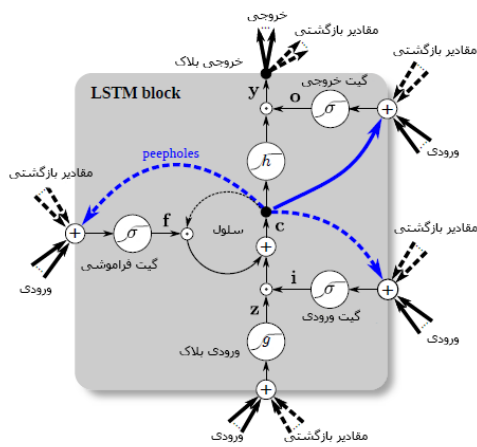
شبکه عصبی بازگشتی. شبکه عصبی بازگشتی یک شبکه عصبی است که خروجی هر سلول نه تنها به لایه‌های پیشین مربوط می‌شود، بلکه خروجی هر سلول ورودی خودش را نیز تغذیه می‌کند؛ به عبارت دیگر، شبکه بازگشتی با استفاده از یک حلقه بازگشتی این امکان را به وجود می‌آورد تا اطلاعات به دست آمده (آخرین وضعیت) توسط شبکه در پایان هر محاسبه برای محاسبات آتی مورد استفاده قرار گیرد. شکل ۲، معماری یک نرون در شبکه‌های عصبی بازگشتی را نشان می‌دهد که دارای یک ورودی در لحظه t ، x_t و یک ورودی دیگر h_{t-1} است. h_{t-1} در حقیقت خروجی این سلول شبکه عصبی در لحظه $t-1$ (آخرین خروجی نرون) بوده است.



شکل ۲. ساختار یک نرون در شبکه عصبی بازگشتی

یادگیری عمیق. اگر تعداد لایه‌های میانی در یک شبکه عصبی بازگشتی زیاد شود در کنار افزایش پیچیدگی، این شبکه قادر خواهد بود با بازدهی بالا و دقت زیاد به کشف الگوهای مخفی شده در داده برای مسائلی که پیچیده و غیرخطی هستند، پردازد. واژه «عمیق» در عبارت «یادگیری عمیق» به تعداد لایه‌های میانی اشاره دارد که هر چه بیشتر باشد، شبکه از عمق بیشتری برخوردار خواهد بود.

معماری LSTM. یکی از مشکلات شبکه‌های عصبی بازگشتی، ناتوانی آن‌ها در یادگیری وابستگی بلندمدت است. برای رفع این مشکل هاگرایتر و همکاران (۱۹۹۷)، یک معماری برای شبکه‌های عصبی بازگشتی پیشنهاد کردند که به معماری حافظه طولانی کوتاه‌مدت معروف شد و قادر بود تا نقصان فراموشی ذکر شده را جبران کند [11].



شکل ۳. ساختار یک بلاک LSTM (K. Greff & et.al, 2017)

از زمان معرفی معماری LSTM تاکنون بهبودهایی در راستای افزایش کارایی بر روی آن انجام شده است. متداول‌ترین معماری LSTM معماری موسوم به وانیلی است که در این پژوهش از آن استفاده شده است [۱۰]. شکل ۳، ساختار یک بلاک LSTM که شامل سه

دروازه ورودی، خروجی و فراموشی می‌شود را نشان می‌دهد. خروجی بلاک LSTM به صورت بازگشتی به ورودی بلاک و ورودی سه گیت نام‌برده شده متصل می‌شود. توابع فعال‌سازی در گیت‌های ورودی و خروجی معمولاً تابع تانژانت هذلولی است و تابع فعال‌سازی در گیت فراموشی تابع سیگموئید است. اتصالات غیرنقطه‌چین شامل اتصالات غیروزن‌دار هستند و اتصالات نقطه‌چین‌دار همگی وزن‌دار هستند.

پیشینه پژوهش. بیشتر پژوهش‌ها در زمینه «پیش‌بینی فرآیند» معطوف به پیش‌بینی نتایج حاصل از یک فرآیند بوده است؛ از این رو پژوهش درباره پیش‌بینی فرآیندها و رویداد بعدی در یک سازمان و یا کسب‌وکار هنوز به بلوغ لازم نرسیده است. در میان پژوهش‌های صورت‌گرفته درباره پیش‌بینی فرآیند بعدی، تاکنون رویکردهای زیر پیشنهاد شده است:

گابریز و همکاران (۲۰۱۲) با به‌کارگیری یک مدل ترکیبی شامل مدل جست‌وجوی نزدیک‌ترین همسایه^۱ و مدل پنهانی مارکوف، روشی را برای پیش‌بینی فرآیند بعدی پیشنهاد کرده‌اند [۱۳]. این رویکرد هر دنباله قبلی فرآیند را به‌عنوان یک حالت در یک ماتریس انتقال حالت بررسی می‌کند به این ترتیب از دنباله قبلی مشاهده‌شده و رویدادهای بعدی آن، یک ماتریس انتقال حالت ساخته می‌شود. هنگامی که یک مورد در حال اجرا به یک حالت که در ماتریس انتقال حالت موجود نیست، برخورد می‌کند شباهت آن به دنباله‌های قبلی مشاهده‌شده با استفاده از فاصله ویرایشی محاسبه می‌شود. پیش‌بینی از موارد مشابه بیشتر مشاهده‌شده در مورد‌های قبلی به‌دست می‌آید. ارزیابی رویکرد در دو مجموعه داده از یک شرکت، دقت در پیش‌بینی رویداد بعدی را ۲۵ و ۷۵ درصد برای دو مجموعه داده‌های خود گزارش می‌دهند. به‌کارگیری مدل پنهانی مارکوف^۲ (HMM) برای پیش‌بینی «فرآیند بعدی» توسط لاکشمن و همکاران (۲۰۱۵) و یونوار و همکاران (۲۰۱۶)، پیشنهاد شده است [۱۲، ۲۲]. در این روش، یک مدل فرآیند از logs موجود استخراج می‌شود؛ آنگاه برای هر انشعاب XOR در مدل، درخت تصمیم‌گیری از داده‌های دنباله فرآیند، استخراج می‌شود؛ سپس این درخت‌ها برای محاسبه احتمال انتقال حالت برای HMM به‌کار می‌روند. پس از آن HMM برای پیش‌بینی احتمال وقایع استفاده می‌شود. این رویکرد در داده‌های شبیه‌سازی‌شده ارزیابی می‌شود و پس از آن آموزش درخت‌های تصمیم‌گیری برای HMM در نیمی از دنباله‌های فرآیندها (مجموعه آموزش)، استفاده خواهد شد. هر دنباله فرآیند در مجموعه آزمون به صورت یک پیشوند و پسوند در یک نقطه تصادفی از هم جدا شده و برای هر یک از پیشوندها، احتمال مشاهده پسوند متناظر مربوطه گزارش می‌شود.

1. K-Nearest Neighbors (KNN)
2. Hidden Markov Model (HMM)

سسی و همکاران (۲۰۱۴) با به‌کارگیری روش تحلیل رگرسیون و اجرای درخت تصمیم توانستند با استفاده از داده‌کاوای دنباله فرآیندها پیشوندهای دنباله فرآیند را پیش‌بینی کنند [۶]. در این روش برای هر پیشوند یک مدل رگرسیون برای پیش‌بینی زمان باقی‌مانده برای تکمیل، ایجاد می‌شود و یک درخت تصمیم‌گیری برای پیش‌بینی رویداد بعدی آموزش داده شده است. در ارزیابی تجربی از دو مجموعه داده استفاده شده نشان می‌دهد که مقدار دقت پیش‌بینی برای رویداد بعدی تقریباً ۶۵ درصد در یک مجموعه داده است و بسته به نوع درخت تصمیم‌گیری و آستانه فرکانس برای پیشوند، دقت پیش‌بینی تقریباً ۵۰ درصد حاصل شده است.

بروکر و همکاران (۲۰۱۶)، برای پیش‌بینی فرآیند بعدی از یک ماشین اتوماتای احتمالاتی متناهی^۱ (PFA) استفاده کردند؛ به این ترتیب حالت پنهان بعدی می‌تواند از حالت پنهان قبلی و رویداد مشاهده شده قبلی (که خود نتیجه احتمالی از حالت پنهان قبلی است) حاصل شود [۴]. در این روش همچنین از یک الگوریتم بیشینه‌کردن امید ریاضی برای تخمین پارامترهای مدل PFA استفاده می‌شود. برای ارزیابی روش پیشنهادی از مجموعه داده رقابت BPI در سال ۲۰۱۲ و ۲۰۱۳، استفاده شده است.

اورمن و همکاران (۲۰۱۷)، با استفاده از الگوریتم یادگیری عمیق روشی را پیشنهاد کردند که قادر است رفتار یک فرآیند را پیش‌بینی کند [۹]. این روش که از مجموعه داده رقابت BPI در سال ۲۰۱۲ و ۲۰۱۳ برای یادگیری استفاده می‌کند، توانسته است با دقت ۸۰ درصد رفتار آتی یک فرآیند را پیش‌بینی کند. در ادامه پژوهش‌های صورت‌گرفته در جدول 1 به صورت مختصر ارائه می‌شود.

در این پژوهش کوشیده می‌شود تا با استفاده از معماری LSTM در شبکه‌های عصبی، روشی پیشنهاد شود که بتواند فرآیند آتی در یک کسب‌وکار و یا در یک سازمان را با دقت بالایی پیش‌بینی کند. با توجه به ویژگی‌های معماری LSTM انتظار می‌رود تا دقت پیش‌بینی صورت‌گرفته در مقایسه با سایر روش‌های یادگیری عمیق به گونه معناداری بالاتر باشد.

جدول ۱. پژوهش‌های انجام‌شده توسط سایر پژوهشگران به تفکیک مجموعه داده استفاده‌شده، معیار ارزیابی، مقدار ارزیابی، روش و نوع پیش‌بینی

پژوهشگر	سال	مقدار ارزیابی	معیار ارزیابی	مجموعه داده	روش	فرد پیش‌بینی
گابریز و همکاران [۱۳]	۲۰۱۲	۷۰	دقت	داده‌های یک شرکت	نزدیک‌ترین همسایه ^۱ و مدل پنهانی مارکوف	رویداد بعدی
سی سی و همکاران [۶]	۲۰۱۴	۶۵	دقت	-	رگراسیون و درخت تصمیم	رویداد بعدی و زمان اتمام فرآیند
اورمان و همکاران [۹]	۲۰۱۷	۸۰	دقت	BPI ۲۰۱۲, ۲۰۱۳	شبکه عصبی LSTM	رویداد بعدی
پولاتو و همکاران [۱۷]	۲۰۱۸	۶/۰۶	خطای جذر میانگین مربعات ^۴	Road fines log	شبکه بی‌زی ^۲ و بردار پشتیبان تصمیم ^۳	رویداد بعدی و زمان اتمام فرآیند
لکش مانان و همکاران [۱۲]	۲۰۱۵	-	-	۲,۰۰۰ فرآیند نمونه	زنجیره مارکوف ^۵ و مدل فرآیند احتمالاتی ^۶	رویداد بعدی به صورت احتمالی
بکر و همکاران [۲]	۲۰۱۴	۷۱/۴	دقت ^۸	۳/۷۷۷ فرآیند نمونه	اتوماتای احتمالاتی متناهی ^۷	رویداد بعدی
بروکر و همکاران [۳]	۲۰۱۵	۷۳/۵	دقت	BPI ۲۰۱۲	حداکثر انتظار ^۹	رویداد بعدی
بروکر و همکاران [۴]	۲۰۱۶	۸۱/۱	دقت	BPI ۲۰۱۲, ۲۰۱۳	اتوماتای احتمالاتی متناهی و شبکه پتری ^{۱۰}	رویداد بعدی
تکس و همکاران [۲۱]	۲۰۱۷	۷۱/۲	میانگین خطای مطلق ^{۱۱}	BPI ۲۰۱۲	شبکه عصبی LSTM	رویداد بعدی و زمان اتمام فرآیند
-	پژوهش جاری	۰/۸۳۹	دقت	BPI ۲۰۱۲	شبکه عصبی LSTM	رویداد بعدی
-	جاری	۰/۹۰۷	دقت	BPI ۲۰۱۷	شبکه عصبی LSTM	رویداد بعدی

۳. روش‌شناسی پژوهش

1. K-Nearest Neighbors (KNN)
2. Bayesian network
3. Support vector machines
4. Root of the mean square error
5. Markov chain
6. probabilistic process model
7. Probabilistic Finite Automata
8. accuracy
9. expectation maximization
10. Petri net
11. Mean absolute error

برای انجام این پژوهش از معماری LSTM در شبکه‌های عصبی بازگشتی استفاده شده است. اجرای این الگوریتم با استفاده از کتابخانه قدرتمند و متن‌باز تنسورفلو^۱ و در زبان برنامه‌نویسی پایتون انجام شده است. آزمایش‌های این پژوهش با به‌کارگیری سخت‌افزار قدرتمند شرکت انویدیا^۲ انجام شده است. در این پژوهش برای انجام فرایند یادگیری شبکه عصبی دو مجموعه داده BPI 2012^۳ و BPI 2017^۴ به‌کار گرفته شده است. مجموعه داده ۲۰۱۲ و ۲۰۱۷ مربوط به رویدادهای فرآیندهای برنامه کاربردی درخواست وام از مؤسسه مالی هلندی است.

جدول ۲. ویژگی‌های مجموعه داده BPI 2017 و BPI 2012

مجموعه داده‌ها	تعداد رویدادهای منحصر به فرد	تعداد رویدادهای	تعداد فرآیندها	تعداد کمترین رویدادهای فرآیند	تعداد بیشترین رویدادهای فرآیند	نسبت تعداد رویداد به تعداد فرآیند
BPI 2017	۲۷	۱۲۰۲۲۶۷	۳۱۵۰۹	۱۰	۱۸۰	۳۸/۱۵۶
BPI 2012	۲۴	۲۶۲۲۰۰	۱۳۰۸۷	۳	۱۷۵	۲۰/۰۳۵

رویدادهای مجموعه داده ۲۰۱۷ مربوط به سال ۲۰۱۶ تا فوریه ۲۰۱۷ است. این مجموعه داده دارای ۳۱۵۰۹ فرآیند است. کوتاه‌ترین فرآیند دارای ۱۰ رویداد و بزرگ‌ترین فرآیند دارای ۱۸۰ رویداد بوده که به‌طور متوسط رویدادهای فرآیندهای این مجموعه داده دارای ۳۸ رویداد است. تعداد رویدادهای کل مجموع داده ۲۰۱۷ برابر ۱۲۰۲۲۶۷ است. رویدادهای مجموعه داده ۲۰۱۲ مربوط به سال ۲۰۱۱ تا مارس ۲۰۱۲ است. این مجموعه داده دارای ۱۳۰۸۷ فرآیند می‌باشد. کوتاه‌ترین فرآیند دارای ۳ رویداد و بزرگ‌ترین فرآیند دارای ۱۷۵ رویداد بوده که به‌طور متوسط رویدادهای فرآیندهای این مجموعه داده دارای ۲۰ رویداد است. تعداد رویدادهای کل مجموع داده ۲۰۱۲ برابر ۲۶۲۲۰۰ است. جدول ۲، ویژگی‌های این دو مجموعه داده را خلاصه کرده است. در انجام آزمایش‌ها از پنج مدل برای آموزش داده‌های آموزش استفاده شد تا تأثیر اندازه مدل در دقت مدلی نهایی سنجیده شود. ویژگی‌های مربوط به این مدل‌ها در جدول ۴، آورده شده است. تنظیماتی که برای پیکربندی شبکه‌های عصبی به‌کار رفت، ۵ پیکربندی است که به‌ترتیب به نام‌های «خیلی خیلی کوچک»، «خیلی کوچک»، «کوچک»، «متوسط» و «بزرگ» نام‌گذاری شده است. هر یک از این پیکربندی‌ها دارای تنظیماتی است که در جدول ۴، ارائه شده است تا در صورتی که سایر پژوهشگران بخواهند نتایج این پژوهش را راستی آزمایی کنند، بدینوسیله

1. TensorFlow
 2. NVIDIA GEFORCE GTX 1080 TI
 3. <https://data.4tu.nl/repository/uuid:3926db30-f712-4394-aebc-75976070e91f>
 4. <https://data.4tu.nl/repository/uuid:5f3067df-f10b-45da-b98b-86ae4c7a310b>

بتوانند شبیه نتایج این پژوهش را تولید کنند. همان‌طور که در پیکربندی خیلی خیلی کوچک مطابق شکل‌های ۵ و ۶ مشاهده می‌شود، تعداد مراحل مورد استفاده توسط سلول‌های LSTM برای دریافت ورودی فرآیندها ۵ مرحله است؛ بنابراین خروجی متناظر برای پیش‌بینی نیز پنج‌تایی است. نرخ یادگیری در شبکه عصبی برای تمامی پیکربندی‌ها مقدار یک و نرخ کاهش یادگیری شبکه عصبی به‌ازای هر پیکربندی مقادیر ۰/۵ تا ۰/۸۵ است. تعداد نرون‌ها در لایه پنهان برای هر پیکربندی از ۳۲ شروع می‌شود و تا ۱۵۰۰ عدد افزایش می‌یابد. برای جلوگیری از اشیاء شبکه عصبی با داده‌های آموزش از روش حذف احتمالی استفاده شده است. حذف احتمالی روشی است که جلوی اشیاء شدن شبکه عصبی با داده‌های آموزش را می‌گیرد. همان‌طور که از اسم آن مشخص است، هنگام یادگیری، برخی از نرون‌ها به‌صورت احتمالی حذف می‌شوند. این بدین معنا است که یادگیری بر روی معماری‌های متفاوت با مجموعه نرون‌های متفاوت صورت می‌گیرد.

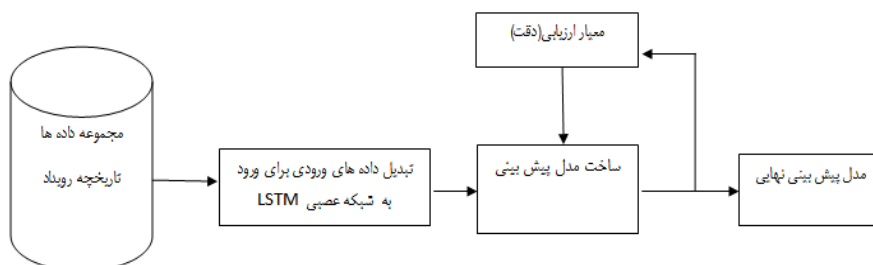
جدول ۴. تنظیمات مدل‌های آموزش داده به‌کارگرفته‌شده در پژوهش

تنظیمات اجرای شبکه عصبی با به‌کارگیری معماری LSTM	چون تعداد	چون تعداد	چون تعداد	چون تعداد	تعداد
مقداردهی اولیه به وزن‌ها و بایاس شبکه عصبی ^۱	۰/۱	۰/۱	۰/۱	۰/۱	۰/۰۴
نرخ یادگیری اولیه ^۲	۱/۰	۱/۰	۱/۰	۱/۰	۱/۰
حداکثر گرادیان مجاز ^۳	۵	۵	۵	۵	۱۰
تعداد گام‌ها ^۴	۵	۲۰	۱۰	۵	۳۵
تعداد نرون‌های لایه پنهان ^۵	۳۲	۲۰۰	۱۰۰	۳۲	۱۵۰۰
حذف تصادفی ^۶	۰/۲	۰/۲	۰/۲	۰/۲	۰/۲
میزان کاهش نرخ یادگیری ^۷	۰/۵	۰/۵	۰/۵	۰/۵	۰/۸۶
اندازه دسته ^۸	۲۰	۲۰	۲۰	۲۰	۲۰
تعداد اپیک‌های مرحله یادگیری ^۹	۱۰	۱۰	۱۰	۱۰	۱۰

1. Initialization scale
2. learning rate
3. max gradian norm
4. Step Number
5. hidden layer size
6. Dropout probability
7. Learning rate decay
8. Batch size
9. Number of training epochs

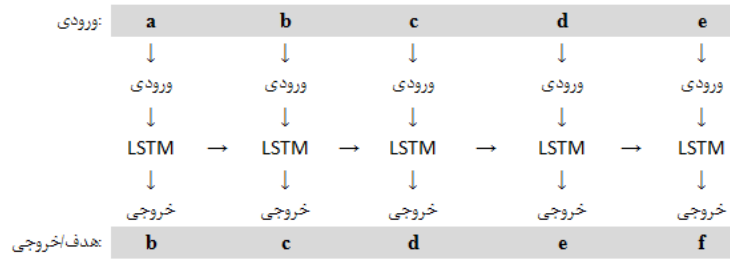
گرایان کاهش‌ی یک الگوریتم بهینه‌سازی برای کاهش خطاها در شبکه عصبی است که کمک می‌کند در تکثیر رو به عقب، خطای به‌روزرسانی وزن‌های گره‌های شبکه عصبی، کمینه شود. مقداردهی اولیه وزن‌ها و بایاس شبکه عصبی با استفاده از توزیع نرمال بین بازه $0/1$ تا $-0/1$ است؛ همچنین با توجه به اینکه برخی از فرآیندها ممکن است دارای بازخورد باشند، عامل «بازخورد» نیز به‌عنوان یکی از پارامترها به آزمایش‌های مربوط به این پژوهش اضافه شد و تمامی آزمایش‌ها با عامل «با بازخورد» و «بدون بازخورد» انجام شد تا تأثیر بازخورد بر دقت مدل اندازه‌گیری شود.

ایجاد مدل. همان‌طور که شکل ۴، نشان می‌دهد، ابتدا داده‌های تاریخچه رویداد تبدیل به داده‌های مناسب برای ورود به شبکه عصبی‌های می‌شود؛ بعد با استفاده از شبکه عصبی LSTM و به‌کارگیری الگوریتم یادگیری و معیارهای ارزیابی، مدل پیش‌بینی نهایی تولید می‌شود.

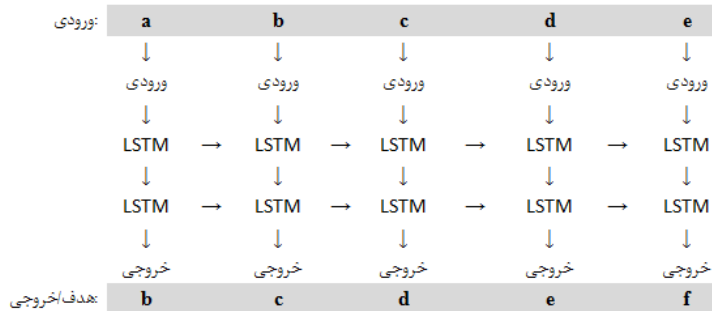


شکل ۴. شمای کلی ایجاد مدل

برای ایجاد مدل پیش‌بینی فرآیند بر اساس تاریخچه رویداد، معماری شبکه عصبی بازگشتی یک‌لایه (شکل ۵) و دولایه (شکل ۶) با به‌کارگیری نرون LSTM ارائه شده است. تعداد گام‌های تعریف‌شده با توجه به جدول ۴، می‌تواند متغیر باشد. در شکل‌های ۵ و ۶، تنظیمات شبکه عصبی بازگشتی LSTM شامل پنج گام است. در طراحی مدلی که باعنوان «بزرگ» نامگذاری شده است، تعداد گام‌ها ۳۵ قدم است.



شکل ۵. معماری شبکه عصبی بازگشتی یک‌لایه با به‌کارگیری نرون LSTM



شکل ۶. معماری شبکه عصبی بازگشتی دو‌لایه با به‌کارگیری نرون LSTM

ورودی‌ها. ورودی الگوریتم یک فایل تاریخچه رویداد است. فرض شده است که ثبت رویدادها به‌صورت متوالی، به‌طوری که هر رویداد به یک فعالیت اشاره داشته و مرتبط با یک نمونه فرآیند خاص است، امکان‌پذیر است. تاریخچه رویداد ممکن است حاوی اطلاعات اضافی پیرامون رویدادها نیز باشد. درحقیقت، الگوریتم می‌تواند از اطلاعات اضافی نظیر منابع (شخص یا وسیله) که در یک فعالیت مصرف یا مقداردهی اولیه می‌شوند، برچسب زمانی رویدادها یا عناصر داده‌ای که با یک رویداد ثبت می‌شوند (نظیر مقدار یک سفارش) و غیره استفاده کند.

جدول ۴. نمونه‌ای از تاریخچه رویدادها

شناسه فرآیند	شناسه رویداد	زمان	رویداد	منابع	هزینه
۱	۱۰۲۰۳۰	۱۲-۱۲-۲۰۱۷:۱۲.۱۵	A	Name1	۱۰۰۰
	۱۰۲۰۳۱	۱۳-۱۲-۲۰۱۷:۱۲.۱۵	B	Name2	۲۰۰۰
	۱۰۲۰۳۲	۱۴-۱۲-۲۰۱۷:۱۲.۱۵	C	Name3	۱۰۰۰
	۱۰۲۰۳۳	۱۵-۱۲-۲۰۱۷:۱۲.۱۵	D	Name4	۲۰۰۰
۲	۱۰۲۰۴۴	۱۹-۱۲-۲۰۱۷:۱۲.۱۵	A	Name1	۱۰۰۰
	۱۰۲۰۴۵	۲۰-۱۲-۲۰۱۷:۱۲.۱۵	F	Name5	۳۰۰۰
	۱۰۲۰۴۶	۲۱-۱۲-۲۰۱۷:۱۲.۱۵	C	Name3	۱۰۰۰

معیار ارزیابی. در این پژوهش از معیارهای دقت، آنتروپی و آشفتگی برای مقایسه مدل‌هایی که از فایل‌های تاریخچه رویدادها در مجموعه داده‌های BPI2012 و BPI 2017 به دست آمده است، استفاده خواهد شد. در ادامه این معیارهای به صورت مختصر توضیح داده می‌شود.

دقت^۱: نشان‌دهنده تعداد پیش‌بینی‌های صحیح انجام شده توسط دسته‌بند^۲ تقسیم بر تعداد کل پیش‌بینی‌های انجام شده توسط همان دسته‌بند است.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{رابطه (۲)}$$

آنتروپی^۳: در علوم فیزیک و در نظریه سیستم‌ها به معنای (بی‌نظمی) و در علوم ارتباطات و اطلاعات نمایانگر اطمینان موجود در اطلاعات مشخصی است. آنتروپی نخستین بار توسط شانون (۱۹۴۸) بیان شد که می‌توان آن را معادل بی‌نظمی دانست. هرچه نظم سیستمی بیشتر شود، آنتروپی آن کاهش می‌یابد و برعکس کاهش نظم باعث افزایش آنتروپی می‌شود. شانون تئوری آنتروپی را به صورت کمی بیان کرد. آنتروپی، یکی از رویکردهای مدیریتی است که به منظور برخورد با بی‌نظمی و بی‌ثباتی در یک سیستم استفاده می‌شود. آنتروپی متغیر تصادفی گسسته X با تابع جرم احتمال $P(X)$ را با نماد $H(X)$ نمایش می‌دهند و به صورت زیر تعریف می‌شود:

$$H(X) = \sum_{i=1}^n p(x_i) I(x_i) = \sum_{i=1}^n P(x_i) \log_b P(x_i) \quad \text{رابطه (۳)}$$

آشفتگی^۴: یک معیار اندازه‌گیری در نظریه اطلاعات برای بررسی میزان مناسب بودن یک توزیع احتمالی یا پیش‌بینی مدل احتمالی یک نمونه است. می‌توان از آشفتگی برای مقایسه مدل‌های احتمالاتی استفاده کرد. $H(p)$ میزان آنتروپی تابع جرم احتمال $P(X)$ است.

$$2H(p) = 2^{-\sum_x p(x) \log_b P(x)} \quad \text{رابطه (۴)}$$

1. Accuracy
2. classification
3. entropy
4. perplexity

بازخورد^۱ در فرآیند: یک فرآیند شامل چندین فعالیت است و ممکن است در یک فرآیند فعالیت‌هایی پشت سر هم تکرار شده باشند. برای مثال، فرآیند نمونه‌ای با توالی ABBBBBCDFFBRTB را در نظر بگیرید. اگر بازخورد از فرآیند نمونه که گفته شد، برداشته شود، این ABCDFBRTB توالی تولید خواهد شد؛ یعنی به جای ۴ تا B از یک B و به جای ۲ تا F از یک F استفاده شده است. در این پژوهش بررسی می‌شود که اگر بازخورد از روی داده‌های تاریخچه رویداد حذف شود، آیا بر دقت پیش‌بینی مدل تأثیرگذار خواهد بود یا خیر؟

برای انجام آزمایش‌ها معماری شبکه عصبی یک بار به صورت یک لایه (شکل ۵) و بار دیگر به صورت دو لایه (شکل ۶) طراحی شد تا تأثیر تعداد لایه‌ها بر روی دقت مدل سنجیده شود. با توجه به اینکه در آزمایش‌های مربوط به این پژوهش از ۱۰ ایپاک استفاده شد، تعداد کل آزمایش‌های انجام‌شده برای این پژوهش ۳۰ آزمایش است که به تفکیک مجموعه داده به شرح جدول ۳، خواهد بود.

جدول ۵. محاسبه تعداد آزمایش‌های صورت‌گرفته به تفکیک مجموعه داده مورد استفاده

مجموعه داده	تعداد آزمایش‌ها = (تعداد لایه * تعداد مدل * تعداد بازخورد)
۲۰۱۷	۲ * ۵ * ۲ = ۲۰
۲۰۱۲	۱ * ۵ * ۲ = ۱۰
مجموع آزمایش‌ها	۳۰ = ۳ * (آزمایش) * ۱۰ (ایپاک)

۴. تحلیل داده‌ها و یافته‌های پژوهش

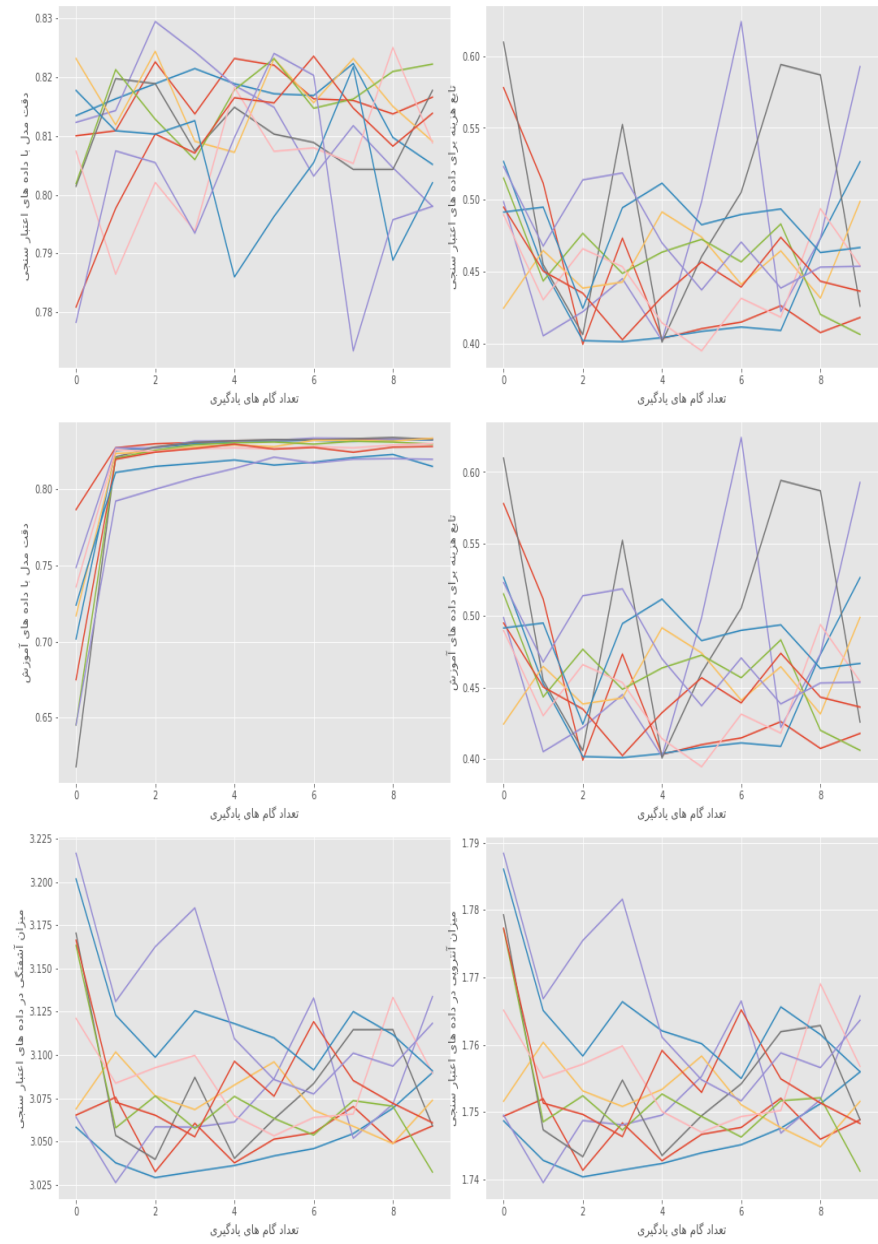
نتایج آزمایش‌ها بر روی مجموعه داده BPI 2012. نتایج حاصل از ۱۰۰ آزمایش انجام‌شده بر روی مجموعه داده BPI 2012 با توجه به تعداد لایه‌ها و تنظیمات پیکربندی در شکل‌های ۴ و ۵، نشان داده شده است. تنظیمات مدل طبق جدول ۲، نیز آزمایش شده است. جدول ۴، نتایج این آزمایش‌ها را به تفکیک تعداد لایه و اندازه مدل نشان می‌دهد. لازم به توضیح است که مجموعه داده BPI 2012 دارای بازخورد نیست.

جدول ۶. نتایج آزمایش‌های مربوط به دقت مدل با داده‌های اعتبارسنجی در مجموعه داده‌های BPI 2012 و نرون LSTM به تفکیک لایه و مدل

تعداد لایه	مدل	تعداد	کمترین	بیشترین	میانگین	انحراف معیار
یک لایه	خیلی خیلی کوچک	۱۰	۰/۷۸۶	۰/۸۲۲	۰/۸۰۵۲۰	۰/۰۱۱۹۷۰
	خیلی کوچک	۱۰	۰/۷۸۶	۰/۸۲۵	۰/۸۰۶۱۰	۰/۰۱۰۹۵۹
	کوچک	۱۰	۰/۸۰۷	۰/۸۲۴	۰/۸۱۶۱۰	۰/۰۰۶۷۳۴
	متوسط	۱۰	۰/۷۹۸	۰/۸۲۹	۰/۸۱۳۱۰	۰/۰۰۹۴۸۰
	بزرگ	۱۰	۰/۸۱۰	۰/۸۲۳	۰/۸۱۶۶۰	۰/۰۰۴۷۱۹
دو لایه	خیلی خیلی کوچک	۱۰	۰/۷۷۳	۰/۸۲۴	۰/۸۰۰۴۰	۰/۰۱۶۴۳۳
	خیلی کوچک	۱۰	۰/۷۸۱	۰/۸۲۴	۰/۸۰۸۹۰	۰/۰۱۱۹۹۵
	کوچک	۱۰	۰/۸۰۲	۰/۸۲۳	۰/۸۱۵۷۰	۰/۰۰۷۰۲۵
	متوسط	۱۰	۰/۸۰۱	۰/۸۲۰	۰/۸۱۰۷۰	۰/۰۰۶۸۹۷
	بزرگ	۱۰	۰/۸۰۵	۰/۸۲۲	۰/۸۱۵۹۰	۰/۰۰۵۲۳۸

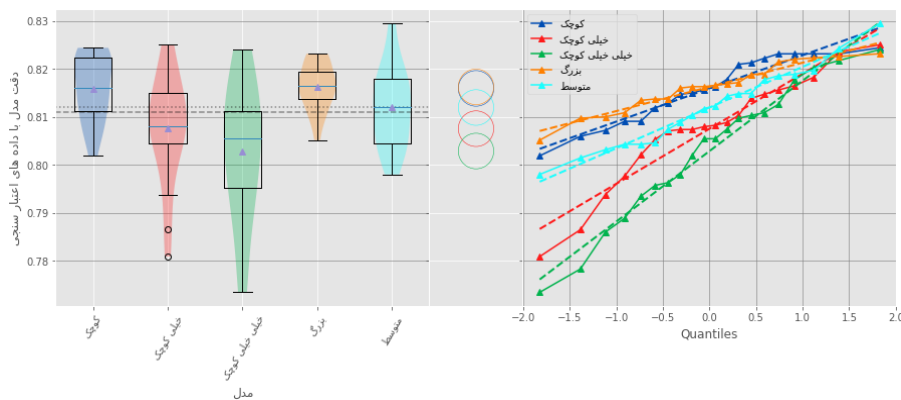
در شکل ۷، ردیف اول سمت چپ، مربوط به نمودار معیار دقت مدل با داده‌های اعتبارسنجی و نمودار سمت راست، ردیف اول تابع هزینه است. در ردیف دوم، سمت چپ مربوط به نمودار معیار دقت مدل با داده‌های آموزش و نمودار سمت راست ردیف دوم، تابع هزینه است. در ردیف سوم، سمت چپ میزان آشفستگی مدل پیش‌بینی با توجه به داده‌های اعتبارسنجی را نشان می‌دهد و ردیف سوم سمت راست، نمودار آنتروپی مدل با توجه به داده‌های اعتبارسنجی را نشان می‌دهد. هر خط رنگی در نمودار که در بالای شکل ۷، آورده شده است مربوط به نوع تنظیمات شبکه عصبی، تعداد لایه و بازخورد/عدم‌بازخورد فرآیندها در داده‌های ورودی است.

— (یک لایه) تعداد لایه (بزرگ): تنظیمات مدل
— (دو لایه) تعداد لایه (بزرگ): تنظیمات مدل
— (یک لایه) تعداد لایه (متوسط): تنظیمات مدل
— (یک لایه) تعداد لایه (کوچک): تنظیمات مدل
— (یک لایه) تعداد لایه (کوچک): تنظیمات مدل
— (یک لایه) تعداد لایه (جلب کوچک): تنظیمات مدل
— (دو لایه) تعداد لایه (جلب کوچک): تنظیمات مدل
— (دو لایه) تعداد لایه (جلب کوچک): تنظیمات مدل
— (یک لایه) تعداد لایه (کوچک): تنظیمات مدل



شکل ۷. نرون LSTM و مجموعه داده BPI 2012

نتایج به‌دست‌آمده از اجرای نرون LSTM با استفاده از مجموعه داده BPI 2012 در دو بخش زیر قابل تحلیل است:
نخست: بررسی تأثیر تنظیمات مدل (خیلی خیلی کوچک، خیلی کوچک، کوچک، متوسط، بزرگ) بر روی دقت مدل.



شکل ۸. نمودار جعبه‌ای و نمودار استاندارد شده داده‌ها که مربوط به دقت مدل با داده‌های اعتبار سنجی بر اساس تنظیمات مدل می‌باشد را با مجموع داده‌های BPI 2012 و نرون LSTM نشان می‌دهد.

بر اساس نمودار جعبه‌ای که میزان پراکندگی دقت را برای مدل‌ها رسم کرده است، توزیع میزان دقت برای مدل‌ها برابر نیست. این مطلب در نمودار مقادیر استاندارد شده داده‌ها نیز قابل مشاهده است به نحوی که مقادیر استاندارد شده داده‌ها نشان می‌دهد مدل‌ها بر یکدیگر منطبق نیستند (شکل ۸).

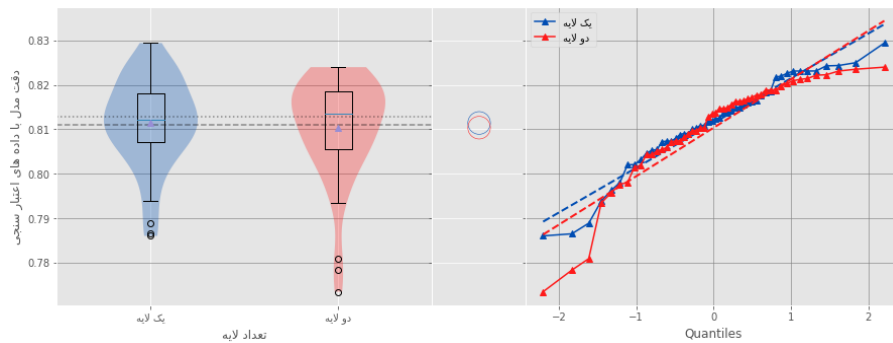
جدول ۷. نتایج آزمون همبستگی اسپیرمن

دقت مدل با داده‌های اعتبار سنجی		مدل	
۰/۳۶۲**	۱/۰۰۰	Correlation Coefficient	دقت مدل با داده‌های اعتبار سنجی
۰/۰۰۰	.	Sig. (2-tailed)	
۱۰۰	۱۰۰	N	
۱/۰۰۰	۰/۳۶۲**	Correlation Coefficient	مدل
.	۰/۰۰۰	Sig. (2-tailed)	
۱۰۰	۱۰۰	N	

Spearman's rho

با توجه به اینکه توزیع داده‌ها غیرنرمال و داده‌ها رتبه‌ای است از آزمون اسپیرمن استفاده شد. مقدار ضریب همبستگی اسپیرمن ۰/۶۸۹۵ به‌دست آمد که مقدار معناداری آزمون برای آن

۰/۰۰۰۰ است و چون از ۰/۰۵ کمتر است، فرض صفر تأیید نمی‌شود. به این معنی که بین متغیر مدل و دقت همبستگی وجود دارد؛ به عبارت دیگر با تغییر مدل مقدار دقت تغییر می‌یابد (جدول ۷).
دوم: بررسی تأثیر تعداد لایه‌ها در دقت مدل.



شکل ۹. نمودار جعبه‌ای و نمودار استاندارد شده داده‌ها که مربوط به دقت مدل با داده‌های اعتبارسنجی بر اساس لایه‌ها است را با مجموع داده‌های BPI 2012 و نرون LSTM نشان می‌دهد.

بر اساس نمودار جعبه‌ای که میزان پراکندگی دقت را برای لایه‌ها رسم کرده است، توزیع میزان دقت برای لایه‌ها تقریباً برابر است. این مطلب در نمودار استاندارد شده داده‌ها نیز قابل مشاهده است به نحوی که مقادیر استاندارد شده داده نشان می‌دهد لایه‌ها تقریباً بر یکدیگر منطبق است و اختلاف قابل توجهی را نشان نمی‌دهد (شکل ۹). با توجه به اینکه توزیع داده‌ها غیرنرمال و داده‌ها رتبه‌ای است، برای تحلیل داده از آزمون اسپیرمن نیز استفاده شد. مقدار ضریب همبستگی اسپیرمن برابر با $-0/023$ به دست آمد. مقدار معناداری آزمون برای آن برابر با $0/824$ بوده و چون از $0/05$ بیشتر است، فرض صفر تأیید می‌شود؛ به این معنی که بین دو متغیر لایه و دقت مدل، همبستگی ثابت نمی‌شود و معناداری وجود ندارد (جدول ۸).

جدول ۸. نتایج آزمون همبستگی اسپیرمن

تعداد لایه		دقت مدل با داده‌های اعتبارسنجی		Spearman's rho
تعداد لایه	دقت مدل با داده‌های اعتبارسنجی	تعداد لایه	دقت مدل با داده‌های اعتبارسنجی	
۱	$-0/023$	۱	$0/824$	دقت مدل با داده‌های اعتبارسنجی
۲	$-0/023$	۲	$0/824$	
Correlation Coefficient		Sig. (2-tailed)		تعداد لایه
۱		۱		
Correlation Coefficient		Sig. (2-tailed)		دقت مدل با داده‌های اعتبارسنجی
$-0/023$		$0/824$		
N		N		
۱		۱		

نتایج آزمایش‌ها بر روی مجموعه داده **BPI 2017**. نتایج حاصل از ۲۰۰ آزمایش انجام شده بر روی مجموعه داده BPI 2017 در شکل ۵، نشان داده شده است. جدول ۹، نتایج این آزمایش‌ها را به تفکیک تعداد لایه و اندازه مدل فهرست کرده است.

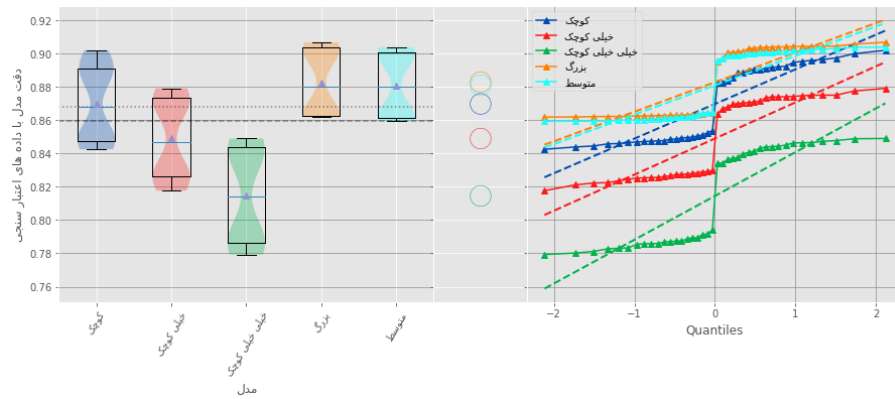
جدول ۹. نتایج آزمایش‌های مربوط به بانک اطلاعاتی BPI 2017 و نرون LSTM به تفکیک لایه، مدل و بازخورد

تعداد لایه	مدل	بازخورد	تعداد	بازخورد	بازخورد	انحراف معیار
یک لایه	خیلی خیلی	بدون بازخور	۱۰	۰/۸۳۷	۰/۸۴۹	۰/۰۴۲۵۴
	کوچک	دارای بازخور	۱۰	۰/۷۸۱	۰/۷۹۲	۰/۰۳۳۲۷
	خیلی کوچک	بدون بازخور	۱۰	۰/۸۶۷	۰/۸۷۹	۰/۰۳۶۲۷
		دارای بازخور	۱۰	۰/۸۲۱	۰/۸۲۹	۰/۰۲۵۳۰
	کوچک	بدون بازخور	۱۰	۰/۸۸۲	۰/۸۹۷	۰/۰۴۷۶۶
		دارای بازخور	۱۰	۰/۸۴۳	۰/۸۵۱	۰/۰۳۴۹۷
	متوسط	بدون بازخور	۱۰	۰/۸۹۶	۰/۹۰۴	۰/۰۲۸۰۷
		دارای بازخور	۱۰	۰/۸۵۹	۰/۸۶۵	۰/۰۱۷۵۱
	بزرگ	بدون بازخور	۱۰	۰/۸۹۵	۰/۹۰۷	۰/۰۳۳۰۰
		دارای بازخور	۱۰	۰/۸۶۲	۰/۸۶۵	۰/۰۱۱۵۵
دو لایه	خیلی خیلی	بدون بازخور	۱۰	۰/۸۳۴	۰/۸۴۹	۰/۰۵۵۹۹
	کوچک	دارای بازخور	۱۰	۰/۷۷۹	۰/۷۹۴	۰/۰۴۵۹۶
	خیلی کوچک	بدون بازخور	۱۰	۰/۸۶۴	۰/۸۷۸	۰/۰۳۸۶۰
		دارای بازخور	۱۰	۰/۸۱۸	۰/۸۳۰	۰/۰۳۴۲۵
	کوچک	بدون بازخور	۱۰	۰/۸۸۲	۰/۹۰۲	۰/۰۵۷۳۶
		دارای بازخور	۱۰	۰/۸۴۵	۰/۸۵۴	۰/۰۲۹۶۱
	متوسط	بدون بازخور	۱۰	۰/۸۹۹	۰/۹۰۳	۰/۰۱۳۳۷
		دارای بازخور	۱۰	۰/۸۵۹	۰/۸۶۴	۰/۰۱۸۹۷
	بزرگ	بدون بازخور	۱۰	۰/۸۹۷	۰/۹۰۵	۰/۰۲۵۵۸
		دارای بازخور	۱۰	۰/۸۶۲	۰/۸۶۴	۰/۰۰۶۶۷

نتایج به‌دست‌آمده از اجرای نرون LSTM با استفاده از مجموعه داده BPI 2017 در سه

بخش زیر قابل تحلیل است:

نخست: بررسی تأثیر تنظیمات مدل (خیلی خیلی کوچک، خیلی کوچک، کوچک، متوسط، بزرگ) بر روی دقت مدل.



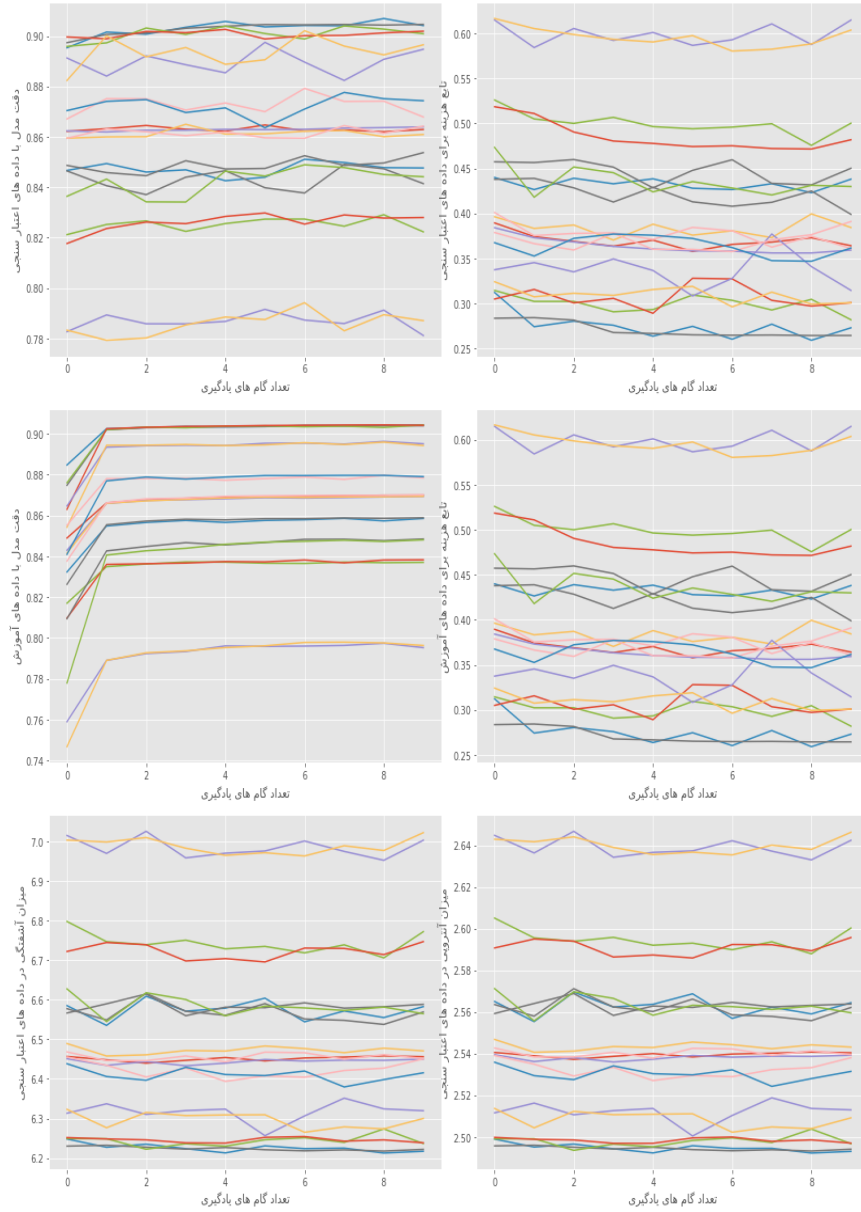
شکل ۱۰. نمودار جعبه‌ای و نمودار استاندارد شده داده‌ها که مربوط به دقت مدل با داده‌های اعتبارسنجی بر اساس تنظیمات مدل است را با مجموع داده‌های BPI 2017 و نرون LSTM نشان می‌دهد.

بر اساس نمودار جعبه‌ای که میزان پراکندگی دقت را برای مدل‌ها رسم کرده است، توزیع میزان دقت برای مدل‌ها برابر نیست. این مطلب در نمودار مقادیر استاندارد شده داده‌ها نیز قابل مشاهده است به نحوی که مقادیر استاندارد شده داده‌ها نشان می‌دهد مدل‌ها بر یکدیگر منطبق نیستند (شکل ۱۰).

جدول ۱۰. نتایج آزمون همبستگی اسپیرمن

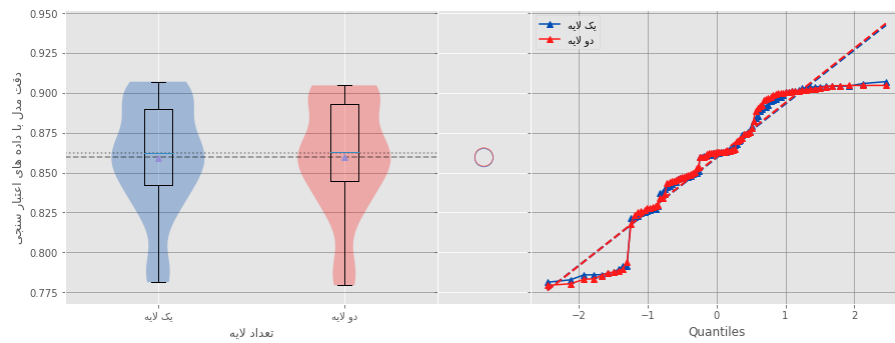
دقت مدل با داده‌های اعتبارسنجی		مدل	
۰/۶۹۰**	۱/۰۰۰	Correlation Coefficient	دقت مدل با داده‌های اعتبارسنجی
۰/۰۰۰	.	Sig. (2-tailed)	
۲۰۰	۲۰۰	N	
۱/۰۰۰	۰/۶۹۰**	Correlation Coefficient	Spearman's rho
.	۰/۰۰۰	Sig. (2-tailed)	مدل
۲۰۰	۲۰۰	N	

با توجه به اینکه توزیع داده‌ها غیرنرمال و داده‌ها رتبه‌ای است از آزمون اسپیرمن استفاده شد و مقدار ضریب همبستگی اسپیرمن ۰/۶۹۰ به دست آمد که مقدار معناداری آزمون برای آن ۰/۰۰۰۰ است و چون از ۰/۰۵ کمتر است، فرض صفر تأیید نمی‌شود؛ به این معنا که بین متغیر مدل و دقت همبستگی وجود دارد؛ به عبارت دیگر همبستگی مثبت بین مدل و دقت وجود دارد. نتایج آزمون همبستگی اسپیرمن در جدول ۱۰، مشاهده می‌شود.



شکل ۱۱. نرون LSTM و مجموعه داده BPI 2017

دوم: بررسی تأثیر تعداد لایه‌ها در دقت مدل.



شکل ۱۲. نمودار جعبه‌ای و نمودار استاندارد شده داده‌ها که مربوط به دقت مدل با داده‌های اعتبارسنجی بر اساس لایه‌ها است را با مجموع داده‌های BPI 2017 و نرون LSTM نشان می‌دهد.

بر اساس نمودار جعبه‌ای که میزان پراکندگی دقت را برای لایه‌ها رسم کرده است، توزیع میزان دقت برای لایه‌ها تقریباً برابر است. این مطلب در نمودار استاندارد شده داده‌ها نیز قابل مشاهده است به نحوی که مقادیر استاندارد شده داده نشان می‌دهد لایه‌ها تقریباً بر یکدیگر منطبق است و اختلاف قابل توجهی را نشان نمی‌دهد (شکل ۱۲).

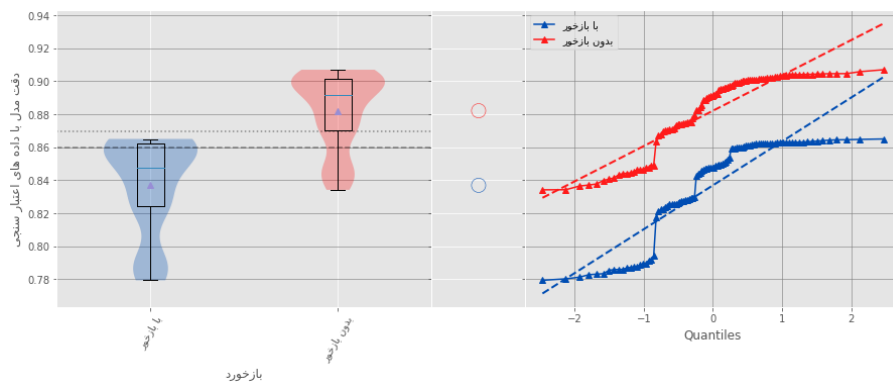
با توجه به اینکه توزیع داده‌ها غیرنرمال و داده‌ها رتبه‌ای است، آزمون اسپیرمن نیز به کار رفت و مقدار ضریب همبستگی اسپیرمن معادل ۰/۱۱۰ به دست آمد که مقدار معناداری آزمون برای آن ۰/۸۷۶ است و چون از ۰/۰۵ بیشتر است، فرض صفر تأیید می‌شود؛ به این معنا که بین دو متغیر لایه و دقت مدل همبستگی ثابت نمی‌شود و معناداری وجود ندارد (جدول ۱۱).

جدول ۱۱. نتایج آزمون همبستگی اسپیرمن

تعداد لایه	دقت مدل با داده‌های اعتبارسنجی		
۰/۰۱۱	۱/۰۰۰	Correlation Coefficient	دقت مدل با داده‌های اعتبارسنجی
۰/۸۷۶	.	Sig. (2-tailed)	
۲۰۰	۲۰۰	N	
۱/۰۰۰	۰/۰۱۱	Correlation Coefficient	تعداد لایه
.	۰/۸۷۶	Sig. (2-tailed)	
۲۰۰	۲۰۰	N	

Spearman's rho

سوم: بررسی تأثیر بازخورد بر روی دقت مدل.



شکل ۱۳. نمودار جعبه‌ای و نمودار استاندارد شده داده‌ها که مربوط به دقت مدل با داده‌های اعتبارسنجی بر اساس تأثیر بازخورد است را با مجموع داده‌های BPI 2017 و نرون LSTM نشان می‌دهد.

بر اساس نمودار جعبه‌ای که میزان پراکندگی دقت را برای بازخورد رسم کرده است، توزیع میزان دقت برای بازخورد برابر نمی‌باشد. این مطلب در نمودار مقادیر استاندارد شده داده‌ها نیز قابل مشاهده می‌باشد به نحوی که مقادیر استاندارد شده داده‌ها نشان می‌دهد مدل‌ها بر یکدیگر منطبق نمی‌باشند (شکل ۱۳).

جدول ۱۲. نتایج آزمون همبستگی اسپیرمن

دقت مدل با داده‌های اعتبارسنجی		بازخورد		Spearman's rho
Correlation Coefficient	۱/۰۰۰	Correlation Coefficient	-۰/۶۶۹**	
Sig. (2-tailed)	.	Sig. (2-tailed)	۰/۰۰۰	
N	۲۰۰	N	۲۰۰	
Correlation Coefficient	-۰/۶۶۹**	Correlation Coefficient	۱/۰۰۰	بازخورد
Sig. (2-tailed)	۰/۰۰۰	Sig. (2-tailed)	.	
N	۲۰۰	N	۲۰۰	

با توجه به اینکه توزیع داده‌ها غیرنرمال و داده‌ها رتبه‌ای است از آزمون اسپیرمن استفاده شد و مقدار ضریب همبستگی اسپیرمن معادل -۰/۶۶۹ است که مقدار معناداری آزمون برای آن ۰/۰۰۰۰ است و چون از ۰/۰۵ کمتر است، فرض صفر تأیید نمی‌شود؛ به این معنا که بین متغیر بازخورد و دقت، همبستگی وجود دارد؛ به عبارت دیگر با تغییر داده‌ها از حالت بدون بازخورد به حالت بازخورد دقت کاهش می‌یابد (جدول ۱۲).

جدول ۱۳. خلاصه کلیه آزمایش‌های همبستگی به تفکیک مجموعه داده‌ها و متغیرهای لایه، مدل و بازخورد

توضیحات	آزمون همبستگی اسپیرمن			آزمون ناپارامتری Kruskal-Wallis H			متغیر	مجموعه داده‌ها
	N	Sig. (2-tailed)	Correlation Coefficient	Asymp. Sig.	Df	Kruskal-Wallis H		
بین دو متغیر لایه و دقت مدل همبستگی ثابت نمی‌شود.	۱۰	۰/۸۲۴	-۰/۰۲۳	۰/۸۲۳	۱	۰/۰۵۰	لایه	BPI 2012
بین متغیر مدل و دقت همبستگی وجود دارد.	۱۰	۰/۰۰۰	۰/۳۶۲	۰/۰۰۱	۴	۱۹/۳۷۴	مدل	
بازخورد فرآیندهای مجموعه داده BPI 2012 فاقد بازخورد است.								
بین دو متغیر لایه و دقت مدل همبستگی ثابت نمی‌شود.	۲۰	۰/۸۷۶	۰/۰۱۱	۰/۸۷۶	۱	۰/۰۲۴	لایه	BPI 2017
بین متغیر مدل و دقت همبستگی وجود دارد.	۲۰	۰/۰۰۰	۰/۶۹۰	۰/۰۰۰	۴	۹۹/۹۲۶	مدل	
با تغییر فرآیندها ورودی از حالت بدون بازخورد به حالت بازخورد دقت کاهش می‌یابد.	۲۰	۰/۰۰۰	-۰/۶۶۹	۰/۰۰۰	۱	۸۹/۱۲۸	بازخورد	

۵. نتیجه‌گیری و پیشنهادها

یافته‌های مهم حاصل از پژوهش. پس از بررسی نتایج حاصل از ۳۰ آزمایش با نرون LSTM بر روی مجموعه داده BPI2012 و BPI2017 موارد زیر به‌عنوان یافته‌های این پژوهش معرفی می‌شوند:

در بانک اطلاعاتی ۲۰۱۲ از مجموع ۱۰ آزمایش انجام شده، بیشترین دقت ۰/۸۳ است که مربوط به حالت تک‌لایه در مدل متوسط است. کمترین مقدار، ۰/۷۷۳ است که مربوط به حالت دولایه و مدل خیلی خیلی کوچک است. بیشترین میانگین ۰/۸۱۵۳ است که مربوط به حالت دولایه و مدل بزرگ است.

در بانک اطلاعاتی ۲۰۱۲ هر چه مدل بزرگ‌تر می‌شود، دقت مدل افزایش می‌یابد. در حالت خیلی خیلی کوچک، میانگین دقت ۰/۸۰ و در مدل بزرگ، میانگین دقت ۰/۸۳ است. در بانک اطلاعاتی ۲۰۱۲ تعداد لایه‌ها بر روی دقت تأثیری ندارد و میانگین دقت در یک‌لایه و دولایه ۰/۸۱ است.

در بانک اطلاعاتی ۲۰۱۷ از مجموع ۲۰ آزمایش انجام‌شده، بیشترین دقت ۰/۹۰۷ است که مربوط به حالت یک‌لایه در مدل بزرگ و بازخورد بدون بازخورد است. کمترین مقدار ۰/۷۷۹ به‌دست آمد که مربوط به حالت دولایه و مدل خیلی‌خیلی کوچک و بازخورد «دارای بازخورد» است. بیشترین میانگین ۰/۹۰۳ است که مربوط به حالت یک‌لایه، مدل بزرگ و بازخورد «بدون بازخورد» است.

در بانک اطلاعاتی ۲۰۱۷ حالت بدون بازخورد دارای دقت میانگین ۰/۸۸ است که دقت میانگین بیشتری نسبت به حالت دارای بازخورد دارد که دقت میانگین آن ۰/۸۴ است. در بانک اطلاعاتی ۲۰۱۷ هر چه مدل بزرگ‌تر می‌شود، دقت مدل بیشتر می‌شود. در حالت خیلی خیلی کوچک میانگین دقت ۰/۸۱ و در مدل بزرگ میانگین دقت ۰/۸۸ است. در بانک اطلاعاتی ۲۰۱۷ تعداد لایه‌ها بر روی دقت تأثیری ندارد و میانگین دقت در یک‌لایه و دولایه ۰/۸۶ است.

موارد مشترک در هر دو مجموعه داده عبارت‌اند از:

۱. در هر دو بانک اطلاعاتی تعداد لایه بر روی دقت مدل تأثیرگذار نیست؛

۲. تغییر تنظیمات مدل بر روی دقت تأثیرگذار است.

پیشنهادها. در این پژوهش با بهره‌گیری از معماری LSTM و با استفاده از مجموعه داده BPI 2012 و BPI2017 دقت پیش‌بینی فرآیندها را در مدیریت فرآیندهای کسب‌وکار به‌میزان زیادی افزایش یافت. بالاترین دقت میانگین به‌دست‌آمده در میان آزمایش‌های صورت‌گرفته ۰/۹۰۳ است که در مقایسه با بالاترین دقت میانگین به‌دست‌آمده در پژوهش‌های مشابه (۰/۸۱) به‌طور معناداری بالاتر است. امکان پیش‌بینی رفتار آینده فرآیند با چنین دقتی این توانمندی را به هر سازمانی می‌دهد تا در مواجهه با سناریوهای از قبل پیش‌بینی‌شده، ریسک چالش‌های مدیریت فرآیندهای کسب‌وکار را تعدیل کنند. با توجه به دقت به‌دست‌آمده در مدل نهایی، پیشنهاد می‌شود در پژوهش‌های آینده با استفاده از معماری‌های دیگری از شبکه عصبی بازگشتی مانند GRU، بر روی مدل‌های جدید پیش‌بینی فرآیند مطالعه شود تا ویژگی‌های آن معماری‌ها در بهبود پیش‌بینی فرآیندهای کسب‌وکار مورد بررسی قرار گیرد.

منابع

1. Alem Tabriz, A., Farrokh, M., Ahmadi, E. (2014). A Comparison of the Neural Network Approach and the Earned Value Management in Predicting Final Cost and Duration of Projects. *Journal of Industrial Management Perspective*, 4(13), 51-65 (In Persian).
2. Becker, J., Breuker, D., Delfmann, P., Matzner, M. (2014) *Designing and Implementing a Framework for Event-based Predictive Modelling of Business Processes*.
3. Breuker, D., Delfmann, P., Matzner, M., Becker, J. (2015). Designing and evaluating an interpretable predictive modeling technique for business processes. *Lecture Notes in Business Information Processing*, Springer Verlag, 541-553.
4. Breuker, D., Delfmann, P., Matzner, M., Becker, J. (2016). Comprehensible Predictive Models for Business Processes. *MIS Quarterly*, 40(4), 1009-1034.
5. Camunda (2016). *An open source platform for workflow and business process management*. camunda.org
6. Ceci, M., Lanotte, P. F., Fumarola, F., Cavallo, D. P., Malerba, D. (2014). Completion time and next activity prediction of processes using sequential pattern mining. *International Conference on Discovery Science*, pp. 49-61.
7. Di Francescomarino, C., Dumas, M., Maggi, F. M., Teinemaa, I. (2016). Clustering-Based Predictive Process Monitoring. *IEEE Transactions on Services Computing*.
8. Dorri, B., & Mazaheri S. (2013). Project Portfolio Selection Based on Performance Assessment: using an Artificial Neural Network. *Journal of Industrial Management Perspective*, 3(11), 39-61. (In Persian)
9. Evermann, J., Rehse, J. R. and Fettke, P. (2017) 'Predicting process behaviour using deep learning', *Decision Support Systems*. Elsevier B.V., 100, 129-140.
10. Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *Ieee Transactions On Neural Networks And Learning Systems*, 28(10), 2222-2232.
11. Hochreiter, S., Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
12. Lakshmanan, G. T., Shamsi, D., Doganata, Y. N., Unuvar, M., Khalaf, R. (2015). Markov prediction model for data-driven semi-structured business processes. *Knowledge and Information Systems*, Springer, 42, 97-126.
13. Le, M., Gabrys, B., & Nauck, D. (2012). A hybrid model for business process event. *International Conference on Artificial Intelligence*, Springer.
14. Maggi, F. M., Di Francescomarino, C., Dumas, M., Ghidini, C. (2014). Predictive Monitoring of Business Processes. *International Conference on Advanced Information Systems Engineering*, Springer.
15. Márquez-Chamorro, A E., Resinas, M., Ruiz-Cortés, A., Toro, M. (2017). Run-time prediction of business process indicators using evolutionary decision rules. *Expert Systems With Applications*, 87, 1-14.
16. Meidan, J.A. García-García, (2017). A survey on business processes management suites. *Computer Standards & Interfaces*, 51.
17. Polato, M., Sperduti, A., Burattin, A., & de Leoni, M. (2018). Time and activity sequence prediction of business process instances. *Computing*. Springer-Verlag Wien, 100(9), 1005-1031.

18. Pooya, A.R., Javan Rad, E.(2014). Mplementation of Neural Networks in Group Technology and Its Comparison to the Results of K-means, Similarity Coefficient Method and Rank Order Clustering. *Journal of Industrial Management Perspective*, 3(12), 39-62. (In Persian)
19. Schmidhuber, J (2015). Deep learning in neural networks: An overview. *Neural Networks*, 16.
20. Shannon, C. E. (1948). A Mathematical Theory of Communication', *Bell System Technical Journal*, 27(3), 379-423.
21. Tax, N., Verenich, I., La Rosa, M., Dumas, M. (2017). Predictive Business Process Monitoring with LSTM Neural Networks. *International Conference on Advanced Information Systems Engineering*, 477-492.
22. Unuvar M., Lakshmanan G. T., & Doganata, Y. N. (2016). Leveraging path information to generate predictions for parallel business processes. *Knowledge and Information Systems, Springer*, 47, 433-461.
23. Yousefi Zenouz, R., & Menhaj, M.B. (2011). Impact of Lumpy Demand Forecasting System on Bullwhip Effect in Supply Chain: A Comparative Approach. *Journal of Industrial Management Perspective*, 1(3), 29-41. (In Persian)