

## پیش بینی غیبت در محل کار توسط شبکه عصبی پس خور

بنیامین دهجوریان<sup>۱</sup>، محمد کمالی مقدم<sup>۲\*</sup>

۱- دانشجوی کارشناسی ارشد، سبزوار، ایران،

Darkbabikool@gmail.com

۲- استادیار، دانشگاه فناوری های نوین سبزوار، ایران،

m.kamali1982@gmail.com

### چکیده

امروزه مهمترین و به عبارت دیگر، عامل اصلی خیلی از معضلات هر جامعه ریشه در مسائل مالی دارد. به طور مثال، تحقیق درباره ی یک جامعه و مناسبات اجتماعی نشان می دهد، آنانی که از نظر اقتصادی در وضعیت مطلوبی هستند از نظر اجتماعی نیز دارای موقعیت و شرایط بهتری هستند و برعکس آنانی که از نظر اقتصادی در وضعیت نابسامانی بسر می برند. یکی از مهم ترین عوامل تعیین کننده وضعیت اقتصاد یک کشور صنایع، کارخانه ها و شرکت های در حال فعالیت می باشند و مهم ترین عنصر داخلی این بخش ها نیز نیرو انسانی شناخته شده است. نیرو انسانی در بخش های مختلف باعث بهبود و یا تضعیف اقتصاد می شود. یکی از عواملی که تاثیر منفی در این فرایند می گذارد غیبت نیرو انسانی و عدم حضور در محل کار است. در این پژوهش با استفاده از اطلاعات بدست آمده، به کمک الگوریتم شبکه عصبی غیبت در محل کار پیش بینی می شود. برای انجام این کار از الگوریتم پس خور برای پیش بینی استفاده شده و همچنین از پارامتر میانگین مربعات خطا برای اندازه گیری استفاده می شود که در نهایت به خطای ۴,۰۰۵۴۶۳ رسیده ایم.

کلمات کلیدی: پیش بینی غیبت در محل کار، پس خور، داده کاوی.

### ۱- مقدمه

در جهان امروز کشورهایی که دارای قدرت اقتصادی بیشتری هستند نقش مهم تری را در تغییر و تحول اقتصاد کشور های دیگر ایفا می کنند. اقتصاد در تمام ابعاد یک جامعه (سیاسی، فرهنگی، علمی و...) از اهمیت بالایی برخوردار است، زیرا استقلال جامعه و قدرت مقاومت آن در مقابل حوادث طبیعی و اجتماعی، شکوه و عظمت یک جامعه در مقابل جوامع دیگر و ثبات سیاسی آن، بستگی به رفح کمبود های اقتصادی دارد. از این رو صنایع و شرکت ها نقش بزرگی را در توسعه اقتصادی کشور برعهده دارند. رقابت، سهم بازار، توسعه حرفه ای و حمایت شخصی از فعالیت های اجتماعی، بهداشت، آموزش و پرورش به یک بازار جدید امیدوار کننده مرتبط است. در کنار توسعه سازمان ها، فشار برای دستیابی به اهداف بیشتر باعث کاهش کارآمدی کارکنان به شکل فزاینده است، و نهایتا موجب اختلال در فعالیت های مرتبط با سلامت می شود که غیبت در محل کار نیز یکی از این اختلالات می باشد. با توجه به تمامی مشکلات بوجود آمده به دلایل ذکر شده و تاثیر آن بر روی اهمیت یک کشور و جامعه، تحقیق و بررسی در زمینه غیبت در محل کار می تواند نقش مفیدی در رابطه با کاهش چنین اختلالاتی را داشته باشد. در این مقاله توسط الگوریتم پس خور مدل پیش بینی غیبت در محل کار ارائه شده است.

## ۲- کارهای گذشته

در سال ۲۰۰۶ مارک لو وایلانت، جرومی دینت و ماری آن ماسوئل در زمینه پیش بینی هزینه های ناشی از بیخوابی بر روی غیبت در محل کار و تجزیه و تحلیل این هزینه ها و پرداختی ها فعالیت خود را انجام دادند. این فعالیت در یک دوره دوساله در پاریس و منطقه ایل د فرانسه انجام شد. داده های مورد استفاده شامل اطلاعات دو دسته از کارکنان همسان که تعداد ۳۶۹ نفر بی خواب و ۳۶۹ نفر دارای خواب خوب بودن را می شد. این داده ها به وسیله آزمون ویلکاکسون که یک آزمون آماری می باشد مورد بررسی و تجزیه و تحلیل قرار گرفت [۱].

در سال ۲۰۱۲ آقای مارتینانو و همکارانش بر روی پیش بینی غیبت در محل کار توسط یک شبکه عصبی فازی تحقیقات خود را انجام دادند. برای انجام این تحقیقات از پایگاه داده مربوط به اطلاعات غیبت از کار که در طول سه سال (۲۰۰۷ - ۲۰۱۰) توسط شرکت بیمه جمع آروی شده بود استفاده شد. داده ها به صورت فایل های اکسل جدول بندی شد و به وسیله یک شبکه عصبی چند لایه با الگوریتم بازگشت به عقب چند پرسپترونی تحلیل شد [۲].

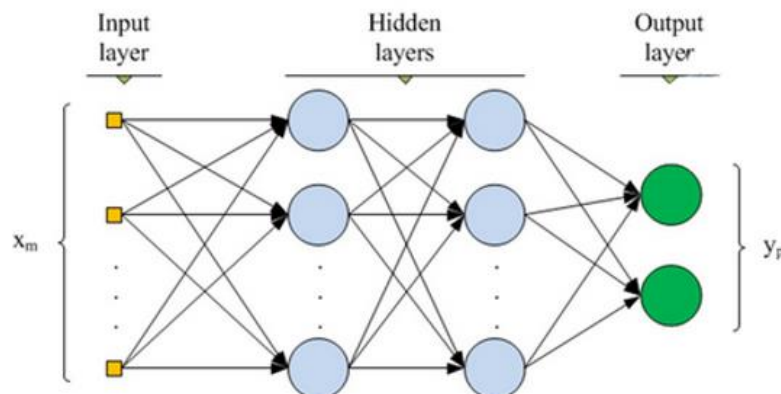
### ۱-۳- الگوریتم شبکه عصبی

فلسفه اصلی شبکه عصبی مصنوعی، مدل کردن ویژگی های پردازشی مغز انسان برای تقریب زدن روش های معمول محاسباتی با روش پردازش زیستی می باشد. به بیان دیگر، شبکه عصبی مصنوعی روشی است که دانش ارتباط بین چند مجموعه داده را از طریق آموزش فراگرفته و برای استفاده در موارد مشابه ذخیره می کند. این پردازنده از دو جهت مشابه مغز انسان عمل می کند:

۱. یادگیری شبکه عصبی از طریق آموزش صورت می گیرد.

۲. وزن دهی مشابه با سیستم ذخیره سازی اطلاعات در شبکه عصبی مغز انسان انجام می گیرد.

یک شبکه عصبی مصنوعی، از سه لایه ورودی، خروجی و پردازش تشکیل می شود. هر لایه شامل گروهی از سلول های عصبی (نورون) است که عموماً با کلیه نورون های لایه های دیگر در ارتباطند، مگر این که کاربر ارتباط بین نورون ها را محدود کند؛ ولی نورون های هر لایه با سایر نورون های همان لایه، مرتبط نیستند.



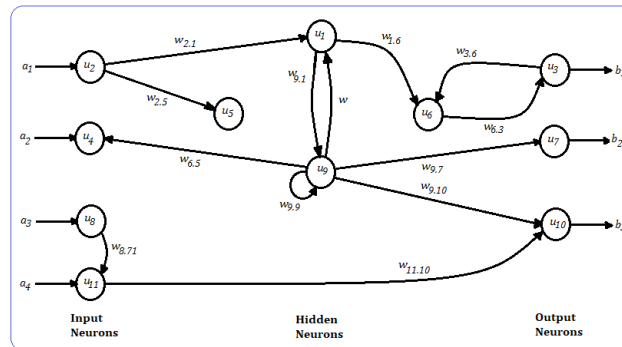
شکل ۱: شمای کلی شبکه عصبی

۲-۳- الگوریتم شبکه عصبی پس خور

<sup>1</sup> Neurons

## سومین همایش ملی مهندسی کامپیوتر، داده کاوی و داده های حجیم

این شبکه ها، شبکه های عصبی با یک حلقه برگشتی (پس خور) یا بیشتر هستند. برگشت ها می توانند محلی یا کلی باشند. خاصیت اصلی آنها این است که رفتار آن ها هم به وسیله خودشان و هم به وسیله ورودی های خارجی به شبکه تعیین می گردد.



شکل ۲: یک نمونه الگوریتم پس خور

### ۴- داده های مورد استفاده

داده های مقاله مربوط به سایت UCI می باشد که توسط آندریا مارتینانو، ریکاردو پینتو فریرا و رانتو خوزه ساسی جمع آوری شده [۳]. این داده ها اطلاعات مربوط به غیبت در محل کار طی ۳ سال (جولای ۲۰۰۷ تا جولای ۲۰۱۰) در برزیل را شامل می شود و دارای ۷۴۰ نمونه و ۲۱ ستون ویژگی می باشد که شامل هیچگونه مقدار گمشده نیست. مقادیر بازه  $\{0, 1\}$  تبدیل شده تا نرمال سازی شوند [۴]. در رابطه (۱) طریقه این کار نمایش داده شده است.

$$x_{new} = \frac{x_{old} - \min}{\max - \min} \quad (1)$$

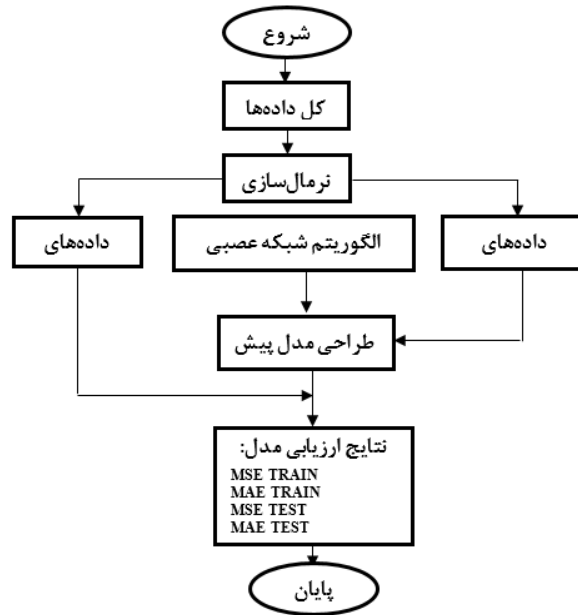
ستون آخر این مجموعه میزان زمان غیبت به ساعت می باشد که نرمال سازی شده.

### ۵- اجرا الگوریتم پس خور

تفاوت شبکه های پس خور با شبکه های پس خور در آن است که در شبکه های برگشتی حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرون های همان لایه یا نرون های لایه های قبیل وجود دارد و اگر نرونی دارای فیدبک باشد بدین مفهوم است که خروجی نرون در لحظه حال نه تنها به ورودی در آن لحظه بلکه به مقدار خروجی خود نرون در لحظه گذشته نیز وابسته می باشد.

در شکل (۳) فلوچارت شبکه های پس خور آمده است.

## سومین همایش ملی مهندسی کامپیوتر، داده کاوی و داده های حجیم



شکل ۳- فلوجارت الگوریتم پس خور

### ۶- ارزیابی

داده های ما دارای ۷۴۰ رکورد و ۲۱ ستون که هر کدام مربوط به یک ویژگی از داده می باشد. ۲۰ ستون اول متغیرهای وابسته و ستون آخر مستقل است. این داده ها به دو قسمت آموزش و تست تقسیم می شود. ۴۹۳ نمونه اول برای آموزش و ۲۴۷ نمونه آخر برای تست انتخاب شده است. از پارامترهای MSE و MAE برای همه ی زیرمجموعه ها استفاده خواهد شد. فرمول (۲) نیز مربوط به میانگین مربعات خطا [۵] (MSE) و فرمول (۳) مربوط به میانگین خطای مطلق [۶] (MAE) است. توجه داشته باشید که تعداد نوروں چه برای داده های آموزشی و چه تست به صورت پیش فرض ۲ در نظر گرفته شده است.

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (۲) \quad \text{میانگین خطای مربعات: MSE}$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (۳) \quad \text{میانگین خطای مطلق: MAE}$$

این دو پارامتر هم برای ارزیابی مجموعه آموزشی و هم مجموعه تست، استفاده می شوند. برای ارزیابی عملکرد شبکه عصبی از روش نمونه گیری چندباره تصادفی استفاده شده است. این روش بر روی داده های مربوط به غیبت در محل کار اجرا شده و بهترین نتایج به دست آمده در جدول (۲) درج شده است.

جدول ۱: ارزیابی الگوریتم پس خور

کل متغیرها		نام داده		
MAE	MSE	آموزشی	تست	آموزشی
4.3482e-10 1.0536e-05	0.00012632	4.5463e-07	N=2 N=2	N=2
		غیبت در محل کار		

مقادیر جدول بالا بصورت غیرریاضی به شرح زیر می باشد.

MSE (آموزش):  $4.3482e-10 = 0.00000000043482$

MSE (تست):  $4.5463e-07 = 0.00000045463$

MAE (تست):  $1.0536e-05 = 0.00001536$

## ۷- نتیجه گیری

امروزه قدرت اقتصادی نقش بسیار مهمی را در شرایط رفاهی داخلی و بین المللی کشورها ایفا می کند، که این امر بستگی به فعالیت نیروی کار سازمان ها و صنایع آن ها بستگی دارد. به دلیل مهم بودن نقش کار در بالا رفتن قدرت اقتصادی کشور ها در این مطالعه روشی برای پیش بینی غیبت در محل کار با استفاده از شبکه عصبی پس خور ارائه شده است. تکنیک های داده کاوی به خوبی می توانند در امر پیش بینی استفاده شوند که از این میان می توان از الگوریتم های رگرسیون درخت تصمیم، رگرسیون چندگانه، نزدیک ترین همسایگی و غیره نام برد که در اینجا ما برای انجام این کار از شبکه عصبی پس خور استفاده کرده ایم و نتایج بدست آمده را مورد بررسی و تجزیه و تحلیل قرار داده ایم.

## ۸- مراجع

- [1]. Godet-Cayré, Virginie, Nathalie Pelletier-Fleury, Marc Le Vaillant, Jérôme Dinet, Marie-Anne Massuel, and Damien Léger. "Insomnia and Absenteeism at Work. Who Pays the Cost?" Sleep 29, no. 2 (2006): 179-184.
- [2]. Martiniano, A., R. P. Ferreira, R. J. Sassi, and C. Affonso. "Application of a Neuro Fuzzy Network in Prediction of Absenteeism at Work." In 7th Iberian Conference on Information Systems and Technologies (CISTI 2012), 1-4. IEEE, 2012.
- [3]. "UCI Machine Learning Repository: Absenteeism at Work Data Set." Accessed June 14, 2019. <https://archive.ics.uci.edu/ml/datasets/Absenteeism+at+work>.
- [4]. Normalization (statistics). In: Wikipedia [Internet]. 2018 [cited 2018 Dec 18].
- [5]. Mean Square Error: Definition & Examples [Internet]. Study.com. [cited 2018 Dec 21].
- [6]. Stephanie. Absolute Error & Mean Absolute Error (MAE) [Internet]. Statistics How To. 2016 [cited 2018 Dec 21].