



## پیشگویی میزان موفقیت روش تلقیح اسپرم داخل سیتوپلاسم تخمک (ICSI) با استفاده از شبکه عصبی مصنوعی و لوژستیک رگرسیون

سیده فضا هاشمی کروی: کارشناسی ارشد، دانشکده علوم زیستی، دانشگاه تربیت مدرس، تهران، ایران، و مرکز تحقیقات باروری و ناباروری حضرت فاطمه الزهرا (س)، دانشگاه علوم پزشکی بابل، بابل، ایران

پرویز عبدالملکی: دانشیار، دانشکده علوم زیستی، دانشگاه تربیت مدرس، تهران، ایران (\*نویسنده مسئول) [parviz@modares.ac.ir](mailto:parviz@modares.ac.ir)  
سید غلامعلی جوهرسرای: استاد، مرکز تحقیقات باروری و ناباروری حضرت فاطمه الزهرا (س)، دانشگاه علوم پزشکی بابل، بابل، ایران

### چکیده

#### کلیدواژه‌ها

تزریق داخل سیتوپلاسمی تخمک،  
لوژستیک رگرسیون،  
شبکه عصبی مصنوعی

**زمینه و هدف:** تزریق اسپرم به داخل تخمک (ICSI) در زوجین نابارور، تحولی شگرف در درمان این بیماران ایجاد نموده است. متأسفانه علی‌رغم هزینه بالای انجام ICSI، میزان موفقیت آن چشم‌گیر نبوده و شکست در ایجاد حاملگی، استرس سنگینی به آن‌ها تحمیل می‌کند. هدف از این تحقیق کوششی برای استخراج بهترین پارامترهای پیشگویی‌کننده در میزان موفقیت ICSI و ارتقای دقت، حساسیت و ویژگی پیشگویی به کمک مدل ریاضی لوژستیک رگرسیون و شبکه عصبی مصنوعی می‌باشد.

**روش کار:** اطلاعات لازم از پرونده ۳۴۵ بیمار تحت درمان ICSI و در قالب ۵۴ پارامتراسمی و عددی استخراج گردید. داده‌ها به صورت تصادفی به دو دسته تخمین و یا آموزش ( $n=276$ ) و اعتبارسنجی یا آزمون ( $n=69$ ) تقسیم شدند. مدل‌های بکاربرده شده در این تحقیق باینری لوژستیک رگرسیون و شبکه عصبی با الگوی یادگیری لونیبرگ مارکوارد بودند. نهایتاً مدل‌ها بوسیله شاخص‌هایی مانند دقت، حساسیت و ویژگی ارزیابی گردیدند.

**یافته‌ها:** بهترین خروجی بدست آمده با مدل باینری لوژستیک و با استفاده از ۵۴ متغیر ۹۷٪ دقت، ۸۶٪ حساسیت و ۹۴٪ ویژگی را بدست داد. بهترین خروجی برای شبکه عصبی با استفاده از الگوی آموزشی لونیبرگ مارکوارد با مدل کاهش یافته شامل تعداد ۲۹ متغیر، دقت ۸۲٪، حساسیت ۹۲٪ و ویژگی ۷۶٪ بدست داد.

**نتیجه‌گیری:** مدل باینری لوژستیک رگرسیون دارای قدرت زیادی در پیشگویی میزان موفقیت تزریق داخل سیتوپلاسمی تخمک هنگامی که خروجی به صورت باینری است می‌باشد.

**تعارض منافع:** گزارش نشده است.

**منبع حمایت‌کننده:** حامی مالی نداشته است.

### شیوه استناد به این مقاله:

Hashemi Karouei SF, Abdolmaleki P, Joorsaraei SG. Prediction the success rate of intracytoplasmic sperm injection (ICSI) using artificial neural networks and logistic regression. Razi J Med Sci. 2019;26(6):118-126.

\*انتشار این مقاله به صورت دسترسی آزاد مطابق با [CC BY-NC-SA 3.0](https://creativecommons.org/licenses/by-nc-sa/3.0/) صورت گرفته است.



Original Article

## Prediction the success rate of intracytoplasmic sperm injection (ICSI) using artificial neural networks and logistic regression

**Seyedeh Fezeh Hashemi Karouei**, MSc, Faculty of Science, Tarbiat Modares University, Tehran, Iran, & Fatemeh Zahra Infertility and Reproductive Health Research Center, Medical Science University of Babol, Babol, Iran

**Parviz Abdolmaleki**, Associate Professor, Faculty of Bioscience, Tarbiat Modares University, Tehran, Iran (\*Corresponding author) [parviz@modares.ac.ir](mailto:parviz@modares.ac.ir)

**Syed Gholamali Joorsaraei**, Professor, Fatemeh Zahra Infertility and Reproductive Health Research Center, Medical Science University of Babol, Babol, Iran

### Abstract

**Background:** Intra-cytoplasmic sperm injection (ICSI) in infertile couples has created a great improvement for treatment of these patients. Unfortunately, despite of high cost of doing ICSI the rate of success is not acceptable and failing pregnancy put a heavy anxiety to couples. This study is aimed to make an effort in order to extract the best predictors for predict the success rate of intra cytoplasmic sperm injection and promote accuracy, sensitivity and specificity by the use of artificial neural network and logistic regression.

**Methods:** 345 patients received ICSI treatment and each of which constructed of 54 numerical and nominal records. This database was randomly divided into the estimation (n=276) and validation (n=69) data set. The models were used based on binary logistic regression (BLR) feature selection tools and Levenberg-Marquardth neural network classifier. Finally, models were evaluated using important criteria such as accuracy, sensitivity and specificity.

**Results:** The best output of the B LR model by using 54 variables revealed accuracy (97%), sensitivity (86%) and specificity (%94). The best output of the LMNN model using Reduced dataset consisted of n=29 with a feature vector side yielded the accuracy (82%) and sensitivity (%92) and specificity (%76).

**Conclusion:** Our result demonstrated that BLR model outperformed highlighting the great power of BLR in success rate of ICSI prediction while using binary output.

**Conflicts of interest:** None

**Funding:** None

### Keywords

Intra cytoplasmic sperm injection,  
Logistic regression,  
Artificial neural network

Received: 27/04/2019

Accepted: 03/08/2019

### Cite this article as:

Hashemi Karouei SF, Abdolmaleki P, Joorsaraei SG. Prediction the success rate of intracytoplasmic sperm injection (ICSI) using artificial neural networks and logistic regression. Razi J Med Sci. 2019;26(6):118-126.

\*This work is published under [CC BY-NC-SA 3.0 licence](https://creativecommons.org/licenses/by-nc-sa/3.0/).



کار برده شود (۸-۱۲). مدل لوژستیک رگرسیون در زمانی استفاده می‌شود که متغیرهای وابسته دارای ۲ وجه بوده و متغیرهای مستقل بصورت پیوسته، رتبه بندی، قطعی یا هر سه شکل آن باشند. این تکنیک دارای قابلیت کافی برای جداسازی مجموعه‌های مختلف می‌باشد (۱۳). کاربرد شبکه عصبی و لوژستیک رگرسیون در بسیاری از مقالات گزارش گردیده است. بعضی از محققین با استفاده از شبکه عصبی مصنوعی و داده‌هایی مانند سن، تعداد تخمکهای بدست آمده، تعداد جنینهای منتقل شده یا تعداد جنینهای فریز منتقل شده، نتایج IVF را با دقت ۵۹٪ پیشگویی نمودند (۱۴). عده‌ای دیگر نیز توانستند با استفاده از شبکه عصبی و داده‌هایی مانند سن زن، طول مدت ناباروری، شاخص توده بدنی (BMI)، بارداری قبلی، آندومترئوز، عوامل لوله‌ای، فاکتورهای تخمک‌گذاری، تعداد اسپرم، زنده بودن اسپرم، تعداد تخمکهای بازایی شده و تعداد تخمکهای منتقل شده، نتایج IVF را با دقت ۷۳٪ پیشگویی نمایند (۸). در این مطالعه قصد ما بر این بود تا مدل لوژستیک رگرسیون را به عنوان یک ابزار دقیق غربالگری و همچنین به منظور تغذیه پارامترهای معنی‌دار، به شبکه عصبی مصنوعی با تابع آموزشی لونبرگ مارکووارد مورد استفاده قرار دهیم. ما از یکی از روش‌های (Under sampling)sampling برای بهبود نتایج شبکه عصبی استفاده نمودیم. نتایج مدل سپس بوسیله شاخص‌های متداول آماری شامل دقت، حساسیت و ویژگی ارزیابی گردید. در این مطالعه مدلی انتخاب گردید که بیشترین دقت را برای ارزیابی نتایج داشته باشد. پیشگویی مربوط به نتایج اگر قبل از مراحل درمان ناباروری انجام پذیرد، از انجام آزمایش‌های متعدد و همچنین، به میزان قابل توجهی باعث کاهش هزینه درمان شده و با جلوگیری از اتلاف زمان، اثرات رنج ذهنی و فیزیکی زوجهای نابارور را نیز کاهش می‌دهد.

### روش کار

این مطالعه با داده‌های ۳۴۵ زوج نابارور با روش

در سال‌های اخیر تزریق داخل سیتوپلاسمی اسپرم (ICSI) به ویژه برای درمان ناباروری با علت مردانه کاربرد زیادی یافته است (۴-۱). در روش ICSI میتوان از اسپرم‌هایی که با استفاده از انزال طبیعی یا بیوپسی بافت بیضه بازایی شده‌اند، استفاده نمود (۱، ۵). تحقیقاتی که در مورد ICSI انجام پذیرفت، حاکی از آن است که این روش برای درمان ناباروری می‌تواند مؤثر واقع شود (۱). در دو دهه گذشته میزان ناباروری، به خصوص در مردان به شکل فزاینده‌ای گسترش یافته، ولی درمان زوجین نابارور علی‌رغم استفاده از روش‌های کمک باروری از جمله ICSI، چندان چشمگیر نبوده است. باید توجه داشت که هر دوره از درمان ناباروری علاوه بر اینکه هزینه‌های زیادی را در بر خواهد داشت، شکست در حاملگی نیز استرس زیادی را به زوجین نابارور تحمیل می‌کند (۷، ۶، ۱). در بسیاری از تحقیقات، عوامل زیادی به عنوان فاکتورهای پیشگویی کننده میزان موفقیت ICSI تعریف شده است. بعضی از گزارشات بر این امر تاکید دارند که تفاوت معنی‌داری در نتایج بدست آمده با استفاده از اسپرمهای بازایی شده از جراحی با حالت انزال طبیعی وجود ندارد؛ اما بعضی از گزارشات بیانگر نرخ پایین باروری با استفاده از اسپرم بدست آمده از بیماران با آروسپرمی انسدادی و غیر انسدادی است (۷، ۶، ۱). البته دستیابی به باروری می‌تواند ارتباط معنی‌داری با سن زن و تخمک‌های بدست آمده داشته باشد، لذا شکل دستیابی به اسپرم، سن و میزان FSH (hormone Follicle-stimulating) سرم مرد، ممکن است تاثیر زیادی در میزان موفقیت بیمارانی با نمای آروسپرم انسدادی و غیر انسدادی نداشته باشد (۶، ۱).

در سال‌های اخیر کاربرد شبکه عصبی مصنوعی در علوم وسیعی مانند پزشکی، بیولوژی، مهندسی، رباتیک و تحلیل تقاضاهای اقتصادی و اجتماعی مشاهده شده است. شبکه عصبی مصنوعی می‌تواند به صورت موثر و با دقت زیاد جایگزین روشهای سنتی پیشگویی آماری شده و در پیشگویی میزان باروری بطور موفقیت آمیز به

دارای ارزش بین صفر و مثبت بی نهایت می باشد (۸). لگاریتم این نسبت که یک متغیر پیوسته است، ارزشی بین منفی بینهایت و مثبت بینهایت می دهد. با استفاده از این معادله ارزش خروجی بدست آمد. به دلیل نیاز به یک آستانه به داده های خروجی، کد دهی مجدد انجام گرفت. برای متغیرهای مستقل، در ۲ وضعیت، ارزش این آستانه بهینه سازی شد. در این مطالعه مدل لوجستیک رگرسیون با روش Backward-wald و در محیط نرم افزاری SPSS 16 اجرا شد. این روش، ابتدا توصیف کننده های اصلی را انتخاب می کند و مدل لوژیستیک رگرسیون را بهبود می بخشد. مرحله آموزش در ۵ مرتبه تکرار گردید. در هر مرتبه بخشی از داده های آموزش از کل داده جدا شده و به عنوان آزمون در نظر گرفته شد. با استفاده از متد backward-wald و تکرار بیشتر در هر مرحله آموزش  $-2\loglikelihood$  بهینه گردید. نهایتاً مدل ضرایبی را برای متغیر مستقل به عنوان ثابت در هر  $-2\loglikelihood$  پیشنهاد داد. این پژوهش با توجه به اینکه، شبکه های عصبی مصنوعی مدل های غیرخطی قدرتمندی هستند و به طور فزاینده ای برای طبقه بندی انواع گوناگون داده ها استفاده می شوند، نیزبکار گرفته شد (۵). ساختار شبکه عصبی شامل چندین لایه است: لایه لایه های ورودی که بردار ورودی را دریافت مینماید، لایه پنهان یا لایه میانی و لایه خروجی که بردار خروجی از لایه قبل را دریافت می نماید. حاصل ضرب بردارهای ورودی در وزنهای خاص آنها که تحت تاثیر تابع فعالسازی قرار میگیرد، بردار خروجی را ایجاد مینماید. این وزنها به صورت تصادفی انتخاب شده و سپس تغییر می یابند به این ترتیب در حالیکه مرحله آموزش در حال انجام است، مجموع مربع خطا (Sum-square error (SSE) بهینه میگردد. شبکه مورد استفاده در این پژوهش، سه لایه ای بود. لایه ورودی آن به طور اتوماتیک با توجه به تعداد پارامترهای انتخاب شده توسط الگوریتم پس انتشار خطا و تابع آموزش (train LM) به منظور آموزش شبکه تعیین گردید (۲۵). به منظور بالا بردن دقت شبکه از مرحله کد دهی برای متغیرهایی مانند تعداد اسپرم، سن زن، و اسپرم با تحرک بالا استفاده نمودیم.

ارزیابی مدل: همه مدلها با استفاده از بعضی از

نمونه گیری آسان و به صورت گذشته نگر که در فاصله زمانی ۱۳۹۱-۱۳۸۹، در مرکز ناباروری حضرت فاطمه الزهرا (س) دانشگاه علوم پزشکی بابل، تحت پروتکل درمانی ICSI قرار گرفته بودند، انجام گردید. اطلاعات بیماران به صورت انتخابی و به روش پرونده خوانی بود. میانگین سنی زنان ۳۰ سال (۱۸-۴۲) و میانگین سنی مردان ۴۲ سال (۶۱-۲۲) بود (۵،۱۵). بارداری (میزان  $HCG/ml \geq U 5I$  در روزهای ۱۰-۱۲ بعد از انتقال جنین به عنوان تست مثبت بارداری در نظر گرفته شد) به عنوان متغیر وابسته انتخاب گردید. ۸۲ ویژگی نیز بعنوان متغیرهای غیر وابسته از اطلاعات پرونده بیماران استخراج گردید (۱۶). در این مرحله عملیات پیش پردازش و پاکسازی داده ها با استفاده از ابزار تحلیل در نرم افزار SPSS صورت گرفت. مرحله پیش پردازش به منظور بهبود داده ها انجام گردید و فرایندهایی مثل تصحیح و یا حذف داده های بدون مقدار، تعیین محدوده مجاز و تصحیح مقادیر غیر مجاز، انجام محاسبات مجدد برای برخی از ویژگی ها و تبدیل آن ها به ویژگی های دیگر مد نظر قرار گرفت. نهایتاً ۵۴ ویژگی بعنوان متغیرهای غیر وابسته انتخاب گردید.

از مدل باینری لوژیستیک رگرسیون و برنامه نرم افزاری SPSS نسخه ۱۶ و مدل شبکه عصبی مصنوعی با تابع آموزشی لونیبرگ مارکووارد (LM) و نرم افزار MATLAB نسخه ۲۰۱۳ استفاده شد (۱۳). داده ها به صورت تصادفی به ۲۷۶ داده آموزش و ۲۹ داده آزمون مجزا و مستقل تقسیم گردیدند (۵). برای تخمین دقت با استفاده از روش یادگیری ماشین (Machine learning algorithms) از روش ارزیابی متقاطع (Cross validation) استفاده شد. با انتخاب مدل لوژیستیک رگرسیون، ارقام صفر (۰) و یک (۱) به ترتیب جهت شکست یا موفقیت درمان در نظر گرفته شد.

[ معادله ۱ ]

$$OUT = \log \left[ \frac{P}{1-P} \right] \\ = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n$$

در اینجا  $p$  احتمال  $Y=1$  را نشان می دهد و  $\beta_0$  یک مقدار ثابت است.  $\beta_1 - \beta_n$  ضریب تاثیر مرتبط با متغیرهای غیر وابسته  $x_1$  و  $x_n$  می باشد. نسبت  $p/(1-p)$

حل پیدا کردن یک هماهنگی قابل قبول بین نتایج مثبت حقیقی و منفی حقیقی در طبقه بندی میباشد. دو روش مهم *sampling* شامل *over sampling* که نمونه های طبقه کمتر را دو برابر می نماید و *under sampling* که بعضی از نمونه ها در طبقه بزرگتر را حذف مینماید، مورد استفاده قرار گرفت (۱۳). کارایی اجرای استراتژی *sampling* در تصمیم گیری های پزشکی بر پایه یادگیری با ماشین (*machine learning*) به صورت کاربردی مشخص شده است (۲۰-۱۷). در این تحقیق ما از روش *under sampling* استفاده نمودیم و بمنظور ارائه طبقه بندی، داده های ICSI را پس از موازنه مجدد بررسی نمودیم (۱۷).

### یافته‌ها

در مجموع ۵۴ متغیر شامل، ۳۲ متغیر عددی و ۲۲ متغیر اسمی انتخاب گردید. ۴۱ متغیر در جدول ۱ جهت مشاهده گزارش شده‌اند.

جدول ۲ مشخصات شبکه عصبی بکار رفته در پروژه ما را نشان می دهد.

شبکه عصبی مصنوعی ما خصوصیات زیر را دارا بود: لایه ورودی، شامل ۲۹ پارامتر معنی دار دریافت شده از ۵ مرحله ارزیابی متقاطع (*Cross validation*) آموزش و آزمون توسط لوژستیک رگرسیون بود (جدول ۳).

لایه پنهان: شامل سه نرون محاسباتی بود که تعداد آن با آزمون و خطا بهینه شده بود.

لایه خروجی: دارای یک نرون محاسباتی که خروجی در دامنه صفر تا یک تولید شد، که نشان دهنده میزان احتمال موفقیت درمان تزریق داخل سیتوپلاسمی تخمک است (۸).

کار با شبکه عصبی را با ۲۹ توصیف کننده بدست آمده از مدل لوژستیک رگرسیون آغاز نمودیم. برای لایه‌های پنهان و خروجی از تابع انتقال سیگموئید و تابع آموزش لونیبرگ مارکووارد استفاده نمودیم (۱۳).

نتایج حاصل از پیشگویی مدل شبکه عصبی مصنوعی با استفاده از ۲۹ توصیف کننده در جدول ۴ گزارش شده است:

در این مطالعه ۳۰۰ تکرار (*Epoch*) بطور تقریبی به منظور آموزش شبکه در نظر گرفته شد و بهترین

شاخصهای آماری ارزیابی می شوند. در تحقیق حاضر ما از شاخص وابسته به آستانه برای ارزیابی مدل لوژستیک رگرسیون و شبکه عصبی مصنوعی استفاده نمودیم.

شاخص وابسته به آستانه: ما از سه مقیاس متفاوت برای ارزیابی کارایی مدل استفاده کردیم. این ۳ مقیاس از چهار شاخص زیر استخراج شده‌اند:

**True positive (TP<sub>i</sub>)** با عنوان تعداد نتایج مثبت تست بارداری (*βHCG result*) که به درستی توسط مدل نتیجه‌گیری شده‌اند، تعریف می‌شود.

**True negative (TN<sub>i</sub>)** با عنوان تعداد نتایج منفی تست بارداری (*βHCG result*) که به درستی توسط مدل نتیجه‌گیری شده‌اند، تعریف می‌شود.

**False positive (FP<sub>i</sub>)** با عنوان تعداد نتایج مثبت تست بارداری (*βHCG result*) که به اشتباه توسط مدل نتیجه‌گیری شده‌اند، تعریف می‌شود.

**False negative (FN<sub>i</sub>)** با عنوان تعداد نتایج منفی تست بارداری (*βHCG result*) که به اشتباه توسط مدل نتیجه‌گیری شده‌اند، تعریف می‌شود.

با استفاده از فرمول‌هایی که قبلاً در گزارشات علمی چاپ شده است، دقت پیشگویی، حساسیت، ویژگی برای خروجی مدل چنین محاسبه می‌شود:

دقت پیشگویی: درصدی از نتایج تست بارداری که به درستی توسط مدل پیشگویی شده‌است [معادله ۲] (۱۳ و ۱۷):

$$\text{Prediction Accuracy}(i) = \frac{TP_i + TN_i}{\text{total}} \times 100$$

حساسیت: درصد نتایج مثبت تست بارداری که به طور صحیحی توسط مدل پیشگویی شده‌است [معادله ۳] (۱۳ و ۱۷):

$$\text{Sensitivity}(i) = \frac{TP_i}{TP_i + FN_i} \times 100$$

ویژگی: درصد نتایج منفی تست بارداری که بدرستی توسط مدل پیشگویی شده‌اند [معادله ۴].

$$\text{Specificity}(i) = \frac{TN_i}{TN_i + FP_i} \times 100$$

نمونه گیری: یکی از مشکلات داده های نامتوازن (*imbalance*) در طبقه بندی دستیابی به حداکثر دقت میباشد. داده های نامتوازن دارای نتایج پیشگویی غیر متقاعد کننده میباشد (۱۷). بهترین راه

جدول ۱- متغیر تحقیق و نوع و مقادیر بازه

نام ویژگی	بازه مقادیر	نوع
سن زن	بازه (۱۸-۴۲)	عددی
تعداد تخمکهای بدست آمده	بازه (۱-۳۷)	عددی
تعداد جنینهای منتقل شده	بازه (۰-۱۰۰)	عددی
تعداد جنینهای فریز شده	بازه (۰-۱۶)	عددی
تعداد تخمکهای بارور شده	بازه (۱-۱۸)	عددی
طول مدت ناباروری (سال)	بازه (۰-۲۲)	عددی
طول مدت ناباروری (ماه)	بازه (۰-۱۴)	عددی
ICSI دفعات	بازه (۰-۶)	عددی
تعداد دفعات انتقال جنینهای فریز شده	بازه (۰-۴)	عددی
تعداد جنینهای فریز انتقال یافته بار اول و دوم	بازه (۰-۶)	عددی
تعداد جنینهای فریز انتقال یافته بار سوم	بازه (۰-۴)	عددی
تعداد جنینهای فریز انتقال یافته بار چهارم	بازه (۰-۲)	عددی
تعداد کلی اسپرم	بازه $10^6 \times (1-100)$	عددی
درصد اسپرمهای با تحرک بالا	بازه $(1-50)\%$	عددی
درصد اسپرمهای با تحرک بسیار بالا	بازه $(1-40)\%$	عددی
تعداد سقط	بازه (۰-۵)	عددی
تعداد جنینهای منتقل شده Grade A	بازه (۰-۴)	عددی
تعداد جنینهای منتقل شده Grade B	بازه (۰-۵)	عددی
تعداد جنینهای منتقل شده Grade C	بازه (۰-۴)	عددی
تعداد جنینهای فریز Grade A منتقل شده بار اول	بازه (۰-۴)	عددی
تعداد جنینهای فریز Grade B منتقل شده بار اول	بازه (۰-۳)	عددی
تعداد جنینهای فریز Grade C منتقل شده بار اول	بازه (۰-۵)	عددی
تعداد جنینهای فریز Grade A منتقل شده بار دوم	بازه (۰-۴)	عددی
تعداد جنینهای فریز Grade B منتقل شده بار دوم	بازه (۰-۳)	عددی
تعداد جنینهای فریز Grade C منتقل شده بار دوم	بازه (۰-۶)	عددی
تعداد جنینهای فریز Grade A منتقل شده بار سوم	بازه (۰-۳)	عددی
تعداد جنینهای فریز Grade B منتقل شده بار سوم	بازه (۰-۲)	عددی
تعداد جنینهای فریز Grade C منتقل شده بار سوم	بازه (۰-۳)	عددی
تعداد جنینهای فریز Grade A, B منتقل شده بار چهارم	بازه (۰-۲)	عددی
تعداد جنینهای فریز Grade C منتقل شده بار چهارم	.	عددی
کیفیت انتقال جنین	(Good&Bloody)	اسمی
نوع ناباروری	اولیه، ثانویه	اسمی
میزان هورمونهای تستوسترون و FSH مرد	(Unknown, Low, Normal, High)	اسمی
تولد زنده	مثبت و منفی	اسمی
نتایج تست حاملگی	مثبت و منفی	اسمی
نژاد	مازنی، ترکمن، غیره	اسمی
نوع سلولهای جنینهای منتقل شده Grade A, B, C	(None, 2, 3cells, 4cells, 5or6cells)	اسمی
نوع سلولهای جنینهای فریز Grade A, B, C منتقل شده برای بار اول	(None, 2, 3cells, 4cells, 5or6cells)	اسمی
نوع سلولهای جنینهای فریز Grade A, B, C منتقل شده برای بار دوم	(None, 2, 3cells, 4cells, 5or6cells)	اسمی
نوع سلولهای جنینهای فریز Grade A, B, C منتقل شده برای بار سوم	(None, 2, 3cells, 4cells, 5or6cells)	اسمی
نوع سلولهای جنینهای فریز Grade A, B, C منتقل شده برای بار چهارم	(None, 2, 3cells, 4cells, 5or6cells)	اسمی

مرحله ارزیابی متقاطع (cross validation) استفاده شد. در هر مرحله در حین تکرار  $2\text{-loglikelihood}$  کاهش یافت. نتایج حاصل از مدل باینری لوژیستیک با استفاده از ۵۴ متغیر نشان دهنده دقت ۰.۹۷٪، حساسیت ۰.۸۶٪ و ویژگی ۰.۹۴٪ بود. با استفاده از توانایی مدل لوژیستیک رگرسیون به منظور تشخیص پارامترهای معنی دار از

عملکرد آموزش و شبیه سازی در تکرار ۸ بدست آمد (شکل ۱).

برای بهبود مدل لوژیستیک رگرسیون به عنوان یک نتیجه، از متد Backward binary logistic استفاده گردید. این روش ابتدا توصیف کننده‌های با اهمیت را پیدا می کند، سپس مدل را بهبود می بخشد. از ۵

## جدول ۲- مشخصات شبکه عصبی

مشخصات شبکه عصبی	
تعداد نرون ها در لایه ورودی	۲۹ لایه
تعداد نرونها در لایه میانی	۳ عدد
تعداد نرونها در لایه خروجی	۱ عدد
نوع تابع فعال سازی	تابع فعالسازی سیگموئید
نوع تابع آموزش	تابع آموزش LM
تعداد تکرار	۸ عدد
مجموع مربعات خطا	(1e-5)
نوع شبکه	سه لایه تغذیه شونده به جلو

## جدول ۳- پارامترهای انتخابی تغذیه شده به شبکه عصبی

شرح	
سن زن	تعداد جنینهای منتقل شده
تعداد جنینهای منتقل شده	تعداد جنینهای فریز شده
تعداد تخمکهای بارور شده	طول مدت ناباروری (سال)
طول مدت ناباروری (ماه)	دفعات ICSI
تعداد دفعات انتقال جنینهای فریز شده	تعداد جنینهای فریز انتقال یافته بار اول
تعداد جنینهای فریز انتقال یافته بار دوم	تعداد جنینهای فریز انتقال یافته بار سوم
تعداد کلی اسپرم	درصد اسپرمهای با تحرک بالا
کیفیت انتقال جنین	نوع ناباروری
میزان هورمون تستوسترون مرد	میزان هورمون FSH مرد
تولد زنده	تعداد سقط
نژاد	تعداد جنینهای منتقل شده Grade A, B, C
نوع سلولهای جنین های منتقل شده Grade B, C	تعداد جنینهای فریز Grade A منتقل شده بار اول
نوع سلولهای جنینهای فریز Grade B منتقل شده بار اول	تعداد جنینهای فریز Grade B منتقل شده بار دوم
نتایج تست حاملگی	

تعداد اسپرم و اسپرم با تحرک بالا کد دهی شدند. ۱۱۵ داده به ۹۵ داده آموزش و ۱۵ داده آزمون تقسیم شده و طی ۶ مرحله به شبکه عصبی تغذیه گردیدند. نتایج دقت، حساسیت و ویژگی بدست آمده برای مدل کاهش یافته به ترتیب ۰.۸۲٪، ۰.۹۲٪ و ۰.۷۶٪ حاصل گردید.

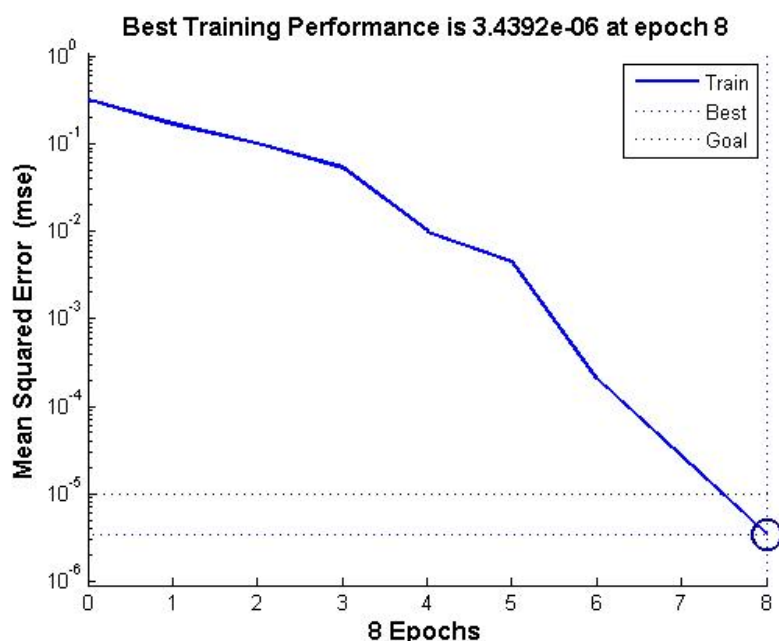
## بحث و نتیجه گیری

متاسفانه علی رغم هزینه بالایی که برای انجام مراحل ICSI صورت می پذیرد، میزان موفقیت چندان زیاد نیست. با توجه به اینکه در بسیاری از اوقات زوجهای نابارور ابتدا به جستجوی هزینه درمان در کلینیکهای ژنیکولوژی پرداخته و نقش بارز آنها در روند زندگی خود حس می کنند، لذا دست نیافتن به نتیجه مثبت و تجربه نکردن باروری، استرس سنگینی را به آنها تحمیل می کند (۱۴). اگرچه ICSI یک روش بسیار پیشرفته برای درمانهای کمک باروری محسوب می شود، ولی ارزش پیشگویی فاکتورهای تاثیر گذار بر میزان موفقیت آن، از اهمیت بالایی برخوردار است (۱، ۲۱). به منظور پیشگویی باروری در چرخه های ICSI، فاکتورهای چند گانه ای پیشنهاد شده است (۲۴-۲۲). در سالهای اخیر بسیاری از تحقیقات بر روی بهبود تکنیکهای مختلف کامپیوتری و مدلهایی ریاضی برای طبقه بندی IUI، ICSI و ET متمرکز شده است و میزان پیشگویی و موفقیت IVF/ICSI توسط این مدل ها مورد پذیرش قرار گرفته است. البته باید در نظر داشت که هیچکدام از این مدلها نمی توانند میزان موفقیت را تا ۱۰۰٪ پیشگویی کنند، اما اگر روشهای پیشگویی کننده بر اساس یاد گیری ماشین به عنوان یک ناظر دوم در کنار جنین شناس قرار گیرند، می توان میزان موفقیت ناباروری را تا حدودی بهبود بخشیده و از درمانهای غیر ضروری که مستلزم تحمیل هزینه برای بیمار و اشغال تخت بیمارستان است ممانعت بعمل آورد. عملکرد مدل لوژیستیک رگرسیون در این پروژه، بسیار سریع بوده و به طور دقیق و صحیح، بیشتر توصیف کننده های موثر را انتخاب نمود. یکی از قابلیت های لوژیستیک رگرسیون ارائه ضرایبی برای متغیرهای مستقل می باشد که به نوعی بیانگر میزان اهمیت هر متغیر در پیشگویی میزان موفقیت درمان می باشد. در حالیکه شبکه عصبی همانند یک جعبه سیاه بوده و

لحاظ آماری و به منظور بهبود دقت شبکه عصبی، ۲۹ پارامتر معنی دار بدست آمده از ۵ مرحله ارزیابی متقاطع بوسیله لوژیستیک رگرسیون به عنوان ورودی به شبکه عصبی تغذیه شدند. با توجه به اینکه پایگاه دادهها در این پروژه، ۳۴۵ داده شامل ۲۷۶ داده آموزش و ۶۹ داده آزمون بودند، لذا با استفاده از متد undersampling (undersampling)، ۲۷۶ داده به ۱۱۵ داده تقلیل یافت. سپس بعضی از پارامترها مانند سن زن،

جدول ۴- توزیع طبقات و نتایج پیشگویی پس از Under sampling بر روی داده های آموزش

تعداد پایگاه داده ها	۱	۲	۳	۴	۵	۶
تعداد موارد مثبت	۴۵	۴۷	۴۵	۴۵	۴۵	۴۸
تعداد موارد منفی	۵۰	۴۸	۵۰	۵۰	۵۰	۵۲
نرخ مثبت حقیقی	۷	۸	۱۱	۱۰	۸	۷
نرخ مثبت کاذب	۱	۳	۳	۴	۳	۲



شکل ۱- بهترین شبیه سازی

دیگر استخراج نماید. در این مطالعه ما موفق به پیشگویی میزان موفقیت ICSI با استفاده از شبکه عصبی مصنوعی و مدل باینری لوژستیک رگرسیون شدیم. ما در یافتیم مدل باینری لوژستیک رگرسیون نتایج دقیقی را در ارزیابی میزان موفقیت ICSI هنگامی که داده ها به صورت باینری باشد، بدست می دهد.

### تقدیر و تشکر

از زحمات سرکار خانم دکتر صدیقه اسماعیل زاده ریاست محترم مرکز درمانی تخصصی فاطمه الزهرا (س) و سرکار خانم دکتر مولود آقاجانی دلاور و همچنین سرکار خانم مریم عبدالله زاده که در مراحل انجام این پروژه ما را یاری نمودند، کمال تشکر و قدردانی را داریم.

### References

1. HashemianNaeni E, AkbariAsbagh F,

کارایی و قابلیت هر یک از متغیرهای مستقل نسبت به متغیر وابسته را به تنهایی تشخیص نمی دهد. بعلاوه شبکه عصبی مستعد به دام افتادن در مینیممهای محلی (Local minima) و در نتیجه طولانی شدن فرآیند آموزش می گردد. نتایج پروژه ما نشان می دهد که طراحی یک شبکه عصبی مصنوعی پیشخور ( feed forward) (۳:۱) و همراهی آن با مدل لوژستیک رگرسیون قادر می باشد، پارامترهای مهم را از میان داده های ورودی استخراج نماید. این امر موجب افزایش دقت، حساسیت و ویژگی شبکه عصبی مصنوعی می گردد. همچنین نتایج بدست آمده حاکی از آن است که قابلیت شبکه عصبی با ۲۹ پارامتر معنی دار بیشتر از زمانی است که آموزش شبکه با ۵۴ پارامتر ورودی انجام پذیرد. در این تحقیق تلاش شده است تا برخی از مشکلات شبکه عصبی، با کمک لوژستیک رگرسیون بهبود یابد. این پروژه مدلی را طراحی نموده که قادر است پارامترهای معنی دار را از میان تمام پارامترهای



- ZareNiyestanak M, Mehrcay A, Foroozandeh E, Shokoohi M, et al. Factors affecting pregnancy and abortion rates following intra cytoplasmic sperm injection in the obstructive and non-obstructive azoospermicmen. *Med Sci*; 2012.22(3):211-215.
2. Speroff L, Fritz M. Assisted reproductive technologies. In: Weinberg RW, Editor. *Clinical Gynecologic Endocrinology and Infertility*. 7<sup>th</sup> ed. Philadelphia: Lippincot Williams and Wilkins; 2005:1215-38.
3. Manoli V, Dandekar S, Desai S, Mangoli R. The outcome of ART in males with impaired spermatogenesis. *J Hum Reprod Sci*; 2008.1:73-76.
4. Hollingsworth B, Harris A, Mortimer D. The cost effectiveness of intracytoplasmic sperm injection (ICSI). *J Assist Reprod Genet*; 2007.24:571-77.
5. Wald M, Speaker AET, Sandlow J, Van-Vorhis B, Syrop CH, Niederberger CS. Computational models for prediction of IVF/ICSI outcomes with surgically retrieved spermatozoa. *Reproduct Biomed Online*; 2005.11(3):325-331.
6. Tournaye H, Devroey P, Liu J, Nagy Z, Lissens W, Van Steirteghem A. [Microsurgical epididymal sperm aspiration and intracytoplasmic sperm injection: a new effective approach to infertility as a result of congenital bilateral absence of the vas deferens]. *Fertile Steril*; 1994.61:1045-51.
7. Zarinara A, Zeraati H, Kamali K, Mohammad K, Shahnazari P, Akhondi MM. [Models predicting success of infertility treatment: A systematic review]. *J Reprod Infertil*; 2016.17(2):66-81.
8. Durairaj M, Thamilselvan P. [Applications of artificial neural network for IVF data analysis and prediction]. *J Engineer Comput Appl Sci*; 2013.2(9).
9. Durairaj M, Meena K. [Application of artificial neural network for predicting fertilization potential of frozen spermatozoa of cattle and buffalo]. *Int J Comput Sci Syst Analyz*; 2008:1-10.
10. Larsson H, Rodrigues M. Can we use in vitro fertilization tests to predict semen fertility? *Animal Reprod Sci*; 2000. 61:327-336.
11. Malinowski P, Milewski R, Ziniewics P, Milewska AJ, Czerniecki J, Wołczyński S. [The use of data mining methods to predict the result of infertility treatment using the IVF ET method]. *Stud Log Grammar Rhetoric*; 2014.39(52):67-74.
12. Milewska AJ, Jankowska D, Cwalina U, Więsak T, Citko D, Morgan A, et al. [Analyzing outcomes of intrauterine insemination treatment by application of cluster analysis or kohonen neural network]. *Studies In Logic, Grammar And Rhetoric. Log Stat Comput Methods Med*; 2013.35(48):7-25.
13. Hayatshahi SHS, Abdolmaleki P, Safarian Sh, Khajeh Kh. [Non-linear quantitative structure-activity relationship for adenine derivatives as competitive inhibitors of adenosine deaminase]. *Biochem Biophys Res Commun*; 2005.1137-1142.
14. Kaufman SJ, Eastaugh L, Snowden S, Swye Sw, Sharma V. [The application of neural networks in predicting the outcome of in-vitro fertilization]. *Hum Reprod*; 1997.12(7):1454-1457.
15. Farsi MM, Jorsaraei A, Hajiahmadi M, Esmaelzadeh S. [Role of embryo morphology in intracytoplasmic sperm injection cycles for prediction of pregnancy]. *Iran J Reprod Med*; 2007. 5(1):23-27.
16. Milewski R, Milewska AJ, Więsak T, Morgan A. [Comparison of artificial neural network and logistic regression analyze in pregnancy prediction using the in vitro fertilization treatment]. *Stud Log Grammar Rhetoric*; 2013.35(48):39-48.
17. Uyar A, Bener A, Ciray HN, Bahceci M. [Handling the imbalance problem of IVf implantation prediction]. *Int J Comput Sci*; 2010.37(2).
18. Cohen G, Hilario M, Sax H, Hugonnet S, Geissbuhler A. [Learning from imbalanced data in surveillance of nosocomial infection]. *Artific Intellig Med*; 2006.37:7-18.
19. Mazurowski HP, Zuranda MA, Lob J, Baker J, Tourassi G. [Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance]. *Neural Networks*; 2008.21:427-436.
20. Namee BM, Cunningham P, Byrne S, Corrigan O. [The problem of bias in the data in regression problems in medical decision support]. *Artific Intellig Med*; 2002.24:51-70.
21. Kozarov G, Stosic L. [Use of semen quality to predict pregnancy in couples undergoing ICSI]. *J Women's Health*; 2016.1(1):001-004.
22. Van Steirteghem AC, Nagy Z, Joris H, Lio J, Strssen C, Smits J, et al. [High fertilization and implantation rates after ICSI]. *Hum Reprod*; 1993.8:1061-66.
23. Friedler S, Raziell A, Strassburger D, Schachter M, Soffer Y, Ron-El R. [Factors influencing the outcome of ICSI I patients with obstructive and non-obstructive azoospermia L: a comparative study]. *Hum Reprod*; 2002.17:3114-21.
24. Mansour RT, Kamal A, Fahmy I, Tawab N, Serour GI, Aboulghar MA. [Intre cytoplasmic sperm injection in obstructive and non-obstructive azoospermia]. *Hum Reprod*; 1997.12:1974-79.
25. Gavin HP. [The levenberg-marquardt method for nonlinear least squares curve-fitting problems]. Department of Civil and Environmental Engineering; 2016:1-18.