



دانشگاه علوم کشاورزی و منابع طبیعی گیلان

نشریه پژوهش در نشخوارکنندگان

جلد سوم، شماره چهارم، ۱۳۹۴

<http://ejrr.gau.ac.ir>

بررسی ارتباط بین فاصله گوساله‌زایی با صفات اقتصادی گاوهای شیری در تلقیح‌های مختلف با استفاده از شبکه عصبی مصنوعی

*مصطفی قادری زفره‌ایی^۱، محمدرضا بحرینی بهزادی^۱، محمد فیاض^۲ و سمیه شریفی^۳
^۱استادیار ژنتیک و اصلاح دام، گروه علوم دامی، دانشکده کشاورزی، دانشگاه یاسوج، ^۲دانشجوی دکتری آمار زیستی، گروه
زیست‌شناسی، دانشکده علوم پایه، دانشگاه شهید بهشتی و ^۳دانشجوی دکتری ژنتیک و اصلاح دام، گروه علوم دامی،
دانشکده کشاورزی، دانشگاه صنعتی اصفهان

تاریخ دریافت: ۱۳۹۴/۹/۳۰؛ تاریخ پذیرش: ۱۳۹۴/۱۲/۲۷

چکیده

سابقه و هدف: فاصله گوساله‌زایی از جمله صفاتی است که به میزان قابل توجهی تحت تأثیر اثرات محیطی و مدیریتی قرار می‌گیرد. استفاده از الگوریتم‌های هوشمند روش‌های یادگیری ماشین در بررسی سامانه‌های پیچیده رو به افزایش است و این روش‌ها نیز می‌تواند رهیافت مناسبی برای تحلیل داده‌های صنعت گاو‌شیری به حساب آیند. شبکه‌های عصبی مصنوعی بخشی از هوش مصنوعی بوده و معمولاً از این الگوریتم‌ها بهره می‌جویند. هدف از انجام این مطالعه بررسی ارتباط بین صفات تولیدی و تولیدمثلی با فاصله گوساله‌زایی در تلقیح‌های متفاوت بود. چنین فرض گردید که استخراج این ارتباط می‌تواند در مدیریت بهتر این صفت نقش بهتری ایفا کند.

مواد و روش‌ها: در این مطالعه، از داده‌های تولیدی و تولیدمثلی شرکت کشت و دامداری فکا وابسته به شرکت تعاونی وحدت گاوداران اصفهان استفاده شد. داده‌های مربوط به فاصله گوساله‌زایی (روز)، طول دوره‌ی خشکی (روز)، تعداد تلقیح، کل شیر تولیدی (کیلوگرم)، کل چربی تولیدی (کیلوگرم)، کل پروتئین تولیدی (کیلوگرم)، روزهای شیردهی (روز)، شیر تصحیح شده (کیلوگرم)، چربی تصحیح شده (کیلوگرم)، پروتئین تصحیح شده (کیلوگرم) مورد بررسی قرار گرفت. داده‌ها مربوط به ۱۵۴۶۵ رأس گاو بود که تاریخ زایش آن‌ها در بازه بین

*مسئول مکاتبه: mosmos741@yahoo.com

سال‌های ۱۳۶۸ تا ۱۳۹۳ قرار داشت. جهت کاهش ابعاد داده‌های مورد استفاده در آموزش شبکه عصبی مصنوعی از دو رویکرد حذف همبستگی بالا و تحلیل مؤلفه اصلی استفاده شد. سپس به ازای هر تلقیح، یک شبکه عصبی آموزش داده شد. برای بررسی کارایی شبکه عصبی از معیارهای ضریب تعیین و جذر میانگین مربعات خطا استفاده شد. میزان رابطه خطی و غیرخطی بین متغیرهای ورودی با فاصله گوساله‌زایی نیز با معیار ضریب اطلاعات بیشین بررسی گردید.

یافته‌ها: مقادیر مشابه معیارهای ضریب تعیین و جذر میانگین مربعات خطا حاصل از شبکه عصبی مصنوعی نشان داد که ارتباط بین متغیرهای ورودی و فاصله گوساله‌زایی در تلقیح‌های مختلف از روند تقریباً یکسانی پیروی می‌کند. مقادیر بزرگتر جذر میانگین مربعات خطا حاصل از رویکرد تحلیل مؤلفه اصلی نسبت به رویکرد حذف همبستگی بالا نشان داد که استفاده از رویکرد حذف همبستگی بالا مناسبتر می‌باشد. معیار ضریب اطلاعات بیشین روند تقریباً یکسانی در میزان رابطه‌ی خطی و غیرخطی متغیرهای مختلف و فاصله گوساله‌زایی در تلقیح‌های مختلف نشان داد. در این راستا روزهای شیردهی و تولید شیر بیشترین ارتباط را با فاصله گوساله‌زایی داشتند.

نتیجه‌گیری کلی: نتایج نشان داد که معیارهای ضریب تعیین و جذر میانگین مربعات خطا در تلقیح‌های مختلف تقریباً یکسان هستند. از این رو نباید انتظار داشت که رابطه بین متغیرهای ورودی به کار رفته در این پژوهش با فاصله گوساله‌زایی در تلقیح‌های مختلف متفاوت باشد. بر اساس نتایج این مطالعه مشخص گردید که متغیرهای به کار رفته در این مطالعه پیش‌بینی خوبی از فاصله گوساله‌زایی را نشان می‌دهند. همچنین می‌توان چنین استنباط کرد که شرایط محیطی - مدیریتی در تلقیح‌های مختلف اثر ناچیزی روی ارتباط بین فاصله گوساله‌زایی و سایر متغیرها داشته است.

واژه‌های کلیدی: تحلیل مؤلفه اصلی، شبکه عصبی مصنوعی، ضریب اطلاعات بیشین، فاصله گوساله‌زایی

مقدمه

با توجه به همبستگی ژنتیکی نامطلوب بین صفات تولید شیر و باروری، توجه به صفات باروری در برنامه‌های اصلاح‌نژادی توصیه شده است (۳۲). تولیدمثل از جمله سازه‌های کلیدی در تعیین بازده و سودآوری در صنعت پرورش گاوهای شیرده محسوب می‌شود (۷). فاصله گوساله‌زایی یکی از صفات شاخص باروری در پرورش گاوهای شیری است و از نظر اقتصادی، میزان باروری در گاوهای شیری معمولاً با توجه به فاصله گوساله‌زایی ارزیابی می‌شود. گزارش شده است که افزایش فاصله گوساله‌زایی باعث کاهش سودآوری می‌گردد (۳۲). پهلوان افشار و همکاران (۲۰۱۲) بر اساس یک مدل زیست‌محیطی گزارش کردند که گاوهای با میانگین فاصله گوساله‌زایی ۲/۴۲۰ روز، در طول عمر تولیدی خود تعداد ۰/۳۶ رأس گوساله کمتری نسبت به گاوهای دارای فاصله گوساله‌زایی ۴/۳۷۵ روز تولید می‌کنند (۲۲). در گله‌هایی که با مشکلات تولیدمثلی مواجه هستند، حذف غیراختیاری افزایش می‌یابد و موجب کاهش تعداد گوساله در گله می‌شود. همچنین کاهش تعداد گوساله نیز باعث کاهش تلیسه جایگزین می‌شود. از طرفی درآمد حاصل از فروش گوساله، به ازای هر روز افزایش در فاصله گوساله‌زایی کاهش می‌یابد. در واقع تمرکز روی پیشرفت صفات تولیدی در صنعت پرورش گاو شیری باعث کاهش عملکرد صفات تولیدمثلی شده است. استفاده از شبکه عصبی در تحلیل داده‌های گاو شیری سابقه نسبتاً طولانی دارد (۸، ۳۳). در واکاوی و تحلیل بیماری ورم پستان از شبکه عصبی مصنوعی استفاده شده است (۳، ۶، ۱۲، ۱۳، ۲۰، ۳۱، ۳۴، ۳۵، ۳۶، ۳۷). گزارش‌هایی مربوط به استفاده از شبکه عصبی در تولید شیر و ارزیابی ژنتیکی آن نیز وجود دارد (۱۰، ۱۱، ۱۵، ۲۰، ۲۵، ۲۶، ۲۷، ۲۹). همچنین برای تشخیص فحلی، کنکاش بیماری لنگش و برای پژوهش‌های تغذیه‌ای نیز در گاو شیری مورد استفاده قرار گرفته است (۵، ۴، ۱۶، ۲۳). برای تحلیل داده‌های حس‌گرهای ریزجریان^۱ در برآورد دقیق زمان باروری گاو شیری نیز از شبکه عصبی استفاده شده است (۲). گزارش شده است که در بررسی سازه‌های موثر بر عملکرد تولیدمثلی گاوهای شیری نیز می‌توان از شبکه عصبی استفاده کرد (۱، ۹). ولی استفاده از روش‌های هوش مصنوعی در بررسی عملکرد تولیدمثلی بسیار کم است. از این‌رو، گسترش و کاربرد روش‌های توسعه یافته در سایر علوم مانند شبکه‌های عصبی مصنوعی که به جهت ریاضی و عملی کارا هستند، در بخش پرورش گاو شیری منطقی به نظر می‌رسد. شبکه‌های

1- Microfluidic

عصبی مصنوعی از روش‌های محاسباتی یادگیری ماشین هستند که توانایی پیش‌بینی پاسخ‌های مناسب از سامانه‌های پیچیده را دارا می‌باشند. از شبکه‌های عصبی مصنوعی می‌توان برای پیش‌بینی استفاده نمود و برای این کار لازم است که ورودی شناخته شده، خروجی شناخته شده و داده‌های کافی وجود داشته باشد (۲۸). ساختار صفات تولیدی و تولیدمثلی اندازه‌گیری شده در گاوهای شیری پیچیده بوده و به علت وابستگی توزیع احتمالی آن‌ها به یکدیگر، استخراج بسیاری از فراسنجه‌های موردنظر با مشکل روبرو می‌شود (۶، ۱۸، ۲۱). نتایج حاصل از یادگیری شبکه‌ی عصبی مصنوعی در تلقیح‌های مختلف روی فاصله گوساله‌زایی تاکنون بررسی نشده است و انتظار می‌رود نتایج حاصله بتواند مدیریت این صفت را در واحدهای پرورش گاو شیری هموارتر سازد. پژوهش حاضر بر پایه آموزش شبکه عصبی مصنوعی بر روی متغیرهای مؤثر بر فاصله گوساله‌زایی صورت گرفت. هدف از پژوهش حاضر بررسی ارتباط بین فاصله گوساله‌زایی و صفات تولیدی و تولیدمثلی در تلقیح‌های مختلف بود.

مواد و روش‌ها

ویرایش داده‌ها: در این مطالعه، از داده‌های تولیدی و تولیدمثلی شرکت کشت و دامداری فکا وابسته به شرکت تعاونی وحدت گاوداران اصفهان استفاده شد. داده‌های مربوط به فاصله گوساله‌زایی (روز)، طول دوره خشکی (روز)، تعداد تلقیح، شیر تولیدی کل (کیلوگرم)، چربی تولیدی کل (کیلوگرم)، پروتئین تولیدی کل (کیلوگرم)، روزهای شیردهی (روز)، شیر تصحیح شده ۳۰۵ روز دوبار دوشش معادل بلوغ^۱، چربی تصحیح شده ۳۰۵ روز دوبار دوشش معادل بلوغ، پروتئین تصحیح شده ۳۰۵ روز دوبار دوشش معادل بلوغ مورد بررسی قرار گرفت. داده‌ها مربوط به ۱۵۴۶۵ رأس گاو بود که تاریخ زایش آن‌ها در بازه بین سال‌های ۱۳۶۸ تا ۱۳۹۳ قرار داشت. از آنجایی که داده‌های پرت تأثیری نامناسب روی برآوردها خواهند داشت، با استفاده از الگوریتم باکن^۲ نرم‌افزار استاتا^۳ مقادیر پرت برای هر صفت مشخص شدند. در این الگوریتم از معیار چندک تصحیح شده توزیع کای دو^۴ استفاده می‌شود. این الگوریتم تمام مشاهدات با فاصله بیشتر از حد آستانه‌ای را به عنوان مشاهدات پرت در نظر می‌گیرد. همچنین برای درک صحیح از ساختار داده‌های مورد مطالعه، توزیع صفات و نرمال بودن

1- ME-305-2x

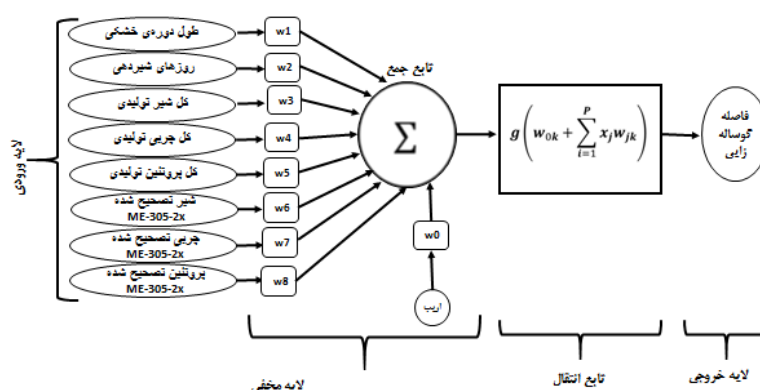
2- Bacon

3- STATA

4- Adjusted quantile of chi-square distribution

نشریه پژوهش در نشخوار کنندگان (۳)، شماره (۴) ۱۳۹۴

داده‌ها با بررسی نمودارهای هیستوگرام و همچنین آزمون شاپیرو-ویلک^۱ مورد بررسی قرار گرفت. بعد از ویرایش داده‌ها و حذف اطلاعات ناقص، تعداد داده‌ها به ۱۴۳۷۱ کاهش پیدا کرد. سپس داده‌ها در بسته نرم‌افزاری کارت^۲ فراخوانده شدند که متکی به محیط نرم‌افزار آماری آر^۳ است. خلاصه آماری صفات مورد بررسی در جدول ۱ نشان داده شده است. داده‌ها از لحاظ تعداد تلقیح، هفت گروه شدند. به طور کلی، از هفت شبکه عصبی که معادل با تعداد گروه تلقیح می‌باشد، استفاده شد. یادگیری شبکه عصبی مصنوعی: برای مدل‌سازی شبکه عصبی از نرم‌افزار وکا^۴ استفاده شد. در مدل‌سازی با استفاده از شبکه عصبی و تعیین معماری آن باید مواردی شامل الگوریتم آموزش شبکه، تابع انتقال و تعداد چرخه‌های مورد استفاده برای آموزش شبکه یا تعداد دوره^۵ را در نظر گرفت. در این تحقیق، معماری سه لایه به کار برده شد که در آن تعداد نوروها در لایه مخفی متفاوت بود. همچنین از تابع انتقال سیگموئیدی^۶ و ۱۰۰ چرخه آموزش شبکه استفاده گردید. تصویر تفهیمی و ساده شبکه عصبی مصنوعی مورد استفاده در شکل ۱ نشان داده شده است.



شکل ۱- تصویر شماتیک شبکه عصبی به کار برده شده در تحلیل داده‌ها.

Figure 1. Schematic representation of neural network used in data analysis.

- 1- Shapiro-Wilk Test
- 2- Caret
- 3- R
- 4- WEKA
- 5- Epoch
- 6- Sigmoid Transfer Function

جدول ۱- آمار توصیفی متغیرهای مورد استفاده در این مطالعه در هر تلقیح.

Table 1. Descriptive statistics of the variables used in this study in each insemination.

حداکثر Maximum	حداقل Minimum	ضریب تغییرات Coefficient of Variation(CV)	انحراف معیار Standard Deviation	میانگین Mean	تعداد مشاهدات NO. of Observations	تعداد تلقیح NO. of Insemination	صفت Trait
563	330	0.079868	28.49815	356.8147	3783	1	فاصله‌ی گوساله‌زایی (روز) Calving Interval (day)
717	330	0.115958	44.85663	386.8354	4156	2	
708	333	0.135173	57.54205	425.6926	2690	3	
708	351	0.129338	59.5577	460.4828	1688	4	
714	373	0.124876	61.6029	493.3134	1018	5	
720	407	0.115553	60.37786	522.5101	645	6	
720	442	0.108057	59.45983	550.2634	391	7	
100	35	0.176328	10.30676	58.45229	3783	1	طول دوره‌ی خشکی (روز) Dry off Period (day)
100	35	0.189661	10.92176	57.58566	4156	2	
100	35	0.193294	11.39351	58.94387	2690	3	
100	35	0.191834	11.5654	60.28851	1688	4	
100	35	0.190369	11.6778	61.34283	1018	5	
100	35	0.206145	12.882	62.48992	645	6	
100	36	0.206558	13.22553	64.02813	391	7	
521	233	0.097637	29.12839	298.3349	3783	1	روزهای شیردهی (روز) Days in milk (day)
656	237	0.133178	43.84702	329.2351	4156	2	
655	264	0.153045	56.12369	366.713	2690	3	
643	297	0.146146	58.48765	400.1996	1688	4	
652	313	0.139713	60.35295	431.9774	1018	5	
671	350	0.130725	60.13728	460.0279	645	6	
659	375	0.119971	58.33528	486.243	391	7	
18922.84	4028.79	0.229586	2303.736	10034.32	3783	1	کل شیر تولیدی (کیلوگرم) Total Milk Production (kg)
23331.8	4443.64	0.225684	2521.87	11174.33	4156	2	
25794.91	4856.15	0.233954	2836.111	12122.51	2690	3	
27771.73	4201.99	0.219012	2896.393	13224.8	1688	4	
27316.59	6787.67	0.219825	3077.818	14001.22	1018	5	
25302.31	6435.04	0.21168	3120.972	14743.83	645	6	
24614.61	7479.47	0.199351	3064.219	15370.98	391	7	
579.26	101.99	0.265861	75.58757	284.3125	3783	1	کل چربی تولیدی (کیلوگرم) Total Fat (kg) Production
708.35	103.08	0.265947	85.02688	319.7135	4156	2	
842.42	111.69	0.276335	96.81964	350.3711	2690	3	
861.88	131.53	0.268433	102.1181	380.4234	1688	4	
843.03	150.1	0.261997	105.7196	403.5149	1018	5	
847.4	164.29	0.264281	113.6399	429.9958	645	6	
824.67	172.43	0.233548	104.4772	447.3487	391	7	
538.77	90.17	0.305446	78.91108	258.3467	3783	1	کل پروتئین تولیدی (کیلوگرم) Total Protein Production (kg)
727.11	90.43	0.310792	89.13846	286.8106	4156	2	
699.29	90.63	0.320732	99.79946	311.1614	2690	3	
776.5	92.55	0.306783	103.9294	338.7716	1688	4	
793.25	104.32	0.29157	104.7391	359.225	1018	5	
722.62	92.47	0.306888	114.8682	374.2997	645	6	
727.8	100.76	0.267164	105.4927	394.8617	391	7	
16627.5	3616.25	0.213697	2001.235	9364.823	3783	1	شیر تصحیح شده (کیلوگرم) Adjusted Milk (ME-305-2x) (kg)
17240.85	4072.92	0.193992	1898.501	9786.498	4156	2	
16025.31	4353.05	0.194306	1911.22	9836.11	2690	3	
15608.61	3728.84	0.177503	1790.004	10084.34	1688	4	
15926.2	5352.89	0.179528	1811.05	10087.86	1018	5	
15917.75	5017.74	0.17038	1733.66	10175.24	645	6	
15686.58	5722.64	0.164053	1687.378	10285.57	391	7	
493.09	101.71	0.251489	66.61082	264.8661	3783	1	چربی تصحیح شده (کیلوگرم) Adjusted Fat (ME-305-2x) (kg)
507.53	101.3	0.241976	67.21158	277.7617	4156	2	
562.11	100.73	0.251043	70.35158	280.2372	2690	3	
500.75	101.76	0.245177	69.54588	283.656	1688	4	
491.88	101.41	0.245214	69.67012	284.1192	1018	5	
506.36	112.43	0.250018	71.96567	287.8417	645	6	
464.38	100.75	0.235831	68.16248	289.0307	391	7	
484.81	80.94	0.294052	69.35771	235.869	3783	1	پروتئین تصحیح شده (کیلوگرم) Adjusted Protein (ME-305-2x) (kg)
726.82	80.04	0.290358	72.07396	248.2248	4156	2	
488.22	80.71	0.296695	74.30437	250.4399	2690	3	
459.99	80.52	0.281682	71.52309	253.9144	1688	4	
459.78	80.94	0.268364	68.09905	253.7561	1018	5	
422.89	82	0.282462	71.36414	252.6501	645	6	
411.2	81.27	0.251798	64.07755	254.4802	391	7	

میزان اریب^۱ معمولاً در شبکه‌های عصبی اعمال می‌گردد که مهم‌ترین نقش آن این است که باعث تغییر تابع انتقال به چپ یا راست می‌شود. معمولاً در شبکه‌های عصبی ترکیب خطی از متغیرهای مستقل، به یک تابع غیرخطی مانند تابع سیگموئیدی $g(.)$ ، تبدیل می‌شوند. تابع سیگموئیدی مورد استفاده در پژوهش حاضر به صورت زیر است.

$$h_k(x) = g\left(w_{0k} + \sum_{i=1}^P x_i w_{ik}\right)$$

در این رابطه، ضرایب وزن w همانند ضرایب رگرسیون هستند که در واقع w_{jk} اثر k امین متغیر مستقل روی k امین واحد مخفی شبکه می‌باشد. در این مطالعه داده‌های هر تلقیح به دو قسمت آموزش (۸۰ درصد) و آزمون (۲۰ درصد) تقسیم‌بندی شدند. برای جلوگیری از اثر همبستگی خطی بین متغیرهای لایه ورودی شبکه عصبی از دو رویکرد حذف همبستگی^۲ بالا و تحلیل مؤلفه اصلی^۳ استفاده شد. در رویکرد اول در هر مجموعه از داده‌های متعلق به هر گروه تلقیح، ابتدا همبستگی خطی بین متغیرها محاسبه گردید. سپس متغیرهایی که همبستگی بین آن‌ها بیش از ۷۵ درصد بود، شناسایی شدند و تنها یکی از آن دو متغیر در مدل‌سازی شبکه‌های عصبی مورد استفاده قرار گرفت. در رویکرد دوم، ابتدا متغیرهای ورودی به روش تحلیل مؤلفه اصلی مورد بررسی و تحلیل قرار گرفتند و سپس برازش شبکه‌های عصبی روی متغیرهای جدید انجام شد. انجام این روش باعث شد که تعداد متغیرهای ورودی مورد استفاده در تلقیح‌های مختلف، یکسان نباشند. برای این که بهترین نتایج حاصل گردد، شبکه‌های عصبی مختلف با تعداد لایه‌های مخفی و وزن‌های اختصاص داده شده مختلف مورد بررسی قرار گرفت. نتایج شبکه‌های عصبی مختلف بر اساس دو معیار ضریب تعیین و جذر میانگین مربعات خطا مورد ارزیابی قرار گرفتند. شبکه‌ای که ضریب تعیین بیشتر و جذر میانگین مربعات خطای کمتری داشت به عنوان بهترین شبکه عصبی انتخاب شد. این معیارها با استفاده از فرمول‌های زیر محاسبه شدند.

$$RMSE = \frac{\sqrt{\sum_{i=1}^n (y_i - y_{id})^2}}{n} \quad \text{معادله (۱)}$$

1- Bias

2- Pearson Correlation

3- Principle Component Analysis

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_{id})^2}{\sum_{i=1}^n (y_{id} - y_m)^2} \quad \text{معادله (۲)}$$

که در آن y_i مقدار برآورد شده فاصله گوساله‌زایی با استفاده از شبکه عصبی است، y_{id} مقدار مشاهده شده فاصله گوساله‌زایی است و y_m میانگین مقدار مشاهده شده فاصله گوساله‌زایی است.

جلوگیری از بیش‌برازش شبکه‌های عصبی: شبکه‌های عصبی به دلیل تعداد زیاد ضرایب وزن (w_{jk}) ، تمایل به ایجاد بیش‌برازشی از ارتباط بین متغیرهای پیشگو و پاسخ دارند. در این پژوهش از روش وزن‌های جریمه‌ای^۱ برای از بین بردن این مشکل استفاده شد (۱۹). در این روش، جریمه‌ای برای ضرایب شبکه عصبی در نظر گرفته می‌شود. بنابراین مقدار وزن‌های جریمه‌ای باید اثر معنی‌داری روی خطاهای مدل جهت هموارکردن آن داشته باشد. فرایند بهینه‌سازی به دنبال کمینه کردن، نسخه دیگری از مجموع مربعات خطاها برای مقدار مشخص λ به دست می‌آورد که به صورت زیر می‌باشد:

$$\sum_{i=1}^n (y_i - f_i(x))^2 + \lambda \sum_{k=1}^H \sum_{j=0}^P w_{jk}^2 + \lambda \sum_{k=0}^H y_k^2.$$

در این حالت شبکه برازش داده شده از نظر ریاضی هموارتر^۲ می‌شود و احتمال بیش‌برازش کاهش می‌یابد. انتخاب وزن‌های جریمه‌ای و تعداد لایه‌های پنهان، نیازمند نمونه‌گیری‌های مجدد از داده‌هاست. به همین منظور از مدل‌های شبکه عصبی میانگین شده استفاده شد (۲۴، ۳۳). سه مقدار $\lambda=0/00$ ، $\lambda=0/01$ و $\lambda=0/10$ برای وزن‌های جریمه‌ای کاهنده به همراه یک لایه پنهان با اندازه بین ۱ تا ۱۳ واحد پنهان در نظر گرفته شد. همچنین میزان رابطه خطی و غیرخطی بین متغیرهای ورودی با فاصله گوساله‌زایی نیز با معیار ضریب اطلاعات بیشین^۳ بررسی گردید.

نتایج و بحث

مدل‌سازی به کار رفته در این پژوهش برای مدیریت بهتر فرایند تولیدمثل در گاو شیری انجام گرفت. فرض اساسی پژوهش حاضر این بود که به کارگیری مدل‌های مبتنی بر شبکه عصبی مصنوعی می‌تواند توانایی مدیریت‌های تولیدمثلی را در گاوداری‌ها افزایش دهد. گاوها به نحوی گروه‌بندی

1- Penalty weights

2- Smoothness

3- Maximal Information Coefficient

نشریه پژوهش در نشخوارکنندگان (۳)، شماره (۴) ۱۳۹۴

شدند که امکان محاسبه فاصله گوساله‌زایی در آن‌ها وجود داشته باشد. با در نظر گرفتن کل داده‌ها (بدون گروه‌بندی آن در تلقیح‌های مختلف)، میانگین فاصله گوساله‌زایی ۴۱۲/۹۳ روز بود. گزارش شده است که به ازای هر روز افزایش در فاصله گوساله‌زایی، نزدیک به ۱۴۰ هزار ریال هزینه ایجاد می‌شود (۳۰). ضعف در تشخیص فحلی با افزایش فاصله زایش و متعاقب آن به علت کاهش تعداد تلیسه جایگزین و کاهش شدت انتخاب در تلیسه‌ها، سبب کاهش روند ژنتیکی تولید شیر در گله‌ها خواهد شد. فاصله بین گوساله‌زایی تا اولین تلقیح و دوره تلقیح و همچنین تعداد تلقیح به ازای آبستنی مهم‌ترین فاکتورهای مؤثر بر فاصله گوساله‌زایی می‌باشند (۳۰).

نتایج شبکه‌های عصبی در رویکرد حذف: نتایج بررسی نرمال بودن متغیرها با استفاده از آزمون شاپیرو-ویلک نشان داد که در تمام آن‌ها مقدار ارزش پی^۱ کمتر از ۰/۰۵ به دست آمد. لذا می‌توان گفت که امکان استفاده از روش‌های رایج رگرسیونی برای مدل‌سازی این داده‌ها و با توجه به مفروضات این مدل‌ها به سختی وجود دارد. در بررسی اولیه صفات مختلف توسط ضریب همبستگی مشخص گردید که متغیرهای کل شیر تولیدی با شیر تصحیح شده، کل چربی تولیدی با چربی تصحیح شده و کل پروتئین تولیدی با پروتئین تصحیح شده دارای ضریب همبستگی پیرسون بالای ۰/۷۵ بودند. از این رو تنها یکی از آن‌ها (رکوردهای تصحیح شده) در شبکه عصبی نهایی مورد استفاده قرار گرفت. در جدول ۲ نتایج برازش هفت مدل شبکه عصبی مورد استفاده به تفکیک تعداد تلقیح نشان داده شده است.

جدول ۲- نتایج شبکه‌های عصبی با روش حذف در تلقیح‌های مختلف.

Table 2. Results of neural networks using elimination method in different inseminations.

Table 2: Results of neural networks using elimination method in different inseminations.									
آزمون ۲۰ درصد (20% testing)			آموزش ۸۰ درصد (80% training)			تعداد	تعداد واحد	تعداد تلقیح No. of insemination	
RMSE	R ²	تعداد No.	RMSE	R ²	تعداد No.	وزن No. of weights	پنهان No. of hidden units		
0.4294	0.9998	755	0.3814	0.9998	3028	3783	50	7	1
1.0544	0.9994	830	0.4708	0.9998	6332	6415	21	4	2
0.4851	0.9999	535	0.4753	0.9999	2155	2690	57	8	3
0.5463	0.9999	337	0.4714	0.9999	1135	8168	71	10	4
0.5282	0.9999	203	0.4783	0.9999	815	1018	71	10	5
0.7123	0.9998	128	0.4753	0.9999	517	654	71	10	6
0.6293	0.9999	77	0.5374	0.9999	314	391	50	10	7

1- P-value

در گروه تعداد تلقیح یک، بهترین ساختار شبکه عصبی دارای ۷ نورون پنهان با اختصاص ۵۰ وزن بود. در مجموع ۳۷۸۳ مشاهده در این دسته قرار داشتند که با تعداد ۳۰۲۸ مشاهده شبکه عصبی آموزش داده شد و ضریب تعیین و جذر میانگین مربعات خطا به ترتیب برابر ۰/۹۹۹۸ و ۰/۳۸۱۴ به دست آمد. سپس این شبکه با استفاده از ۷۵۵ مشاهده باقی‌مانده برازش داده شد که ضریب تعیین آن ثابت ماند ولی جذر میانگین مربعات خطای آن اندکی افزایش پیدا کرد و به مقدار ۰/۴۲۹۴ رسید. در سایر گروه‌های تلقیح نیز تفسیر جدول به همین صورت انجام می‌گیرد. تعداد وزن‌ها، در ستون سوم جدول ۲ نشان داده شده است. در شبکه‌های عصبی مورد استفاده، تنها یک لایه پنهان وجود دارد که خود شامل چندین واحد پنهان می‌باشد. تعداد وزن‌های این شبکه از فرمول $H(P + 1) + H + 1$ به دست می‌آید. در این فرمول H تعداد واحدهای پنهان و P تعداد متغیرهای ورودی شبکه است. در گروه تعداد تلقیح یک، ۷ واحد پنهان و ۵ متغیر ورودی وجود دارد که ایجاد ۵۰ وزن می‌کند. در گروه تعداد تلقیح دو، ۴ واحد پنهان و ۵ متغیر ورودی وجود دارد که ایجاد ۲۹ وزن می‌کند. این در حالی است که در جدول ۲، تعداد وزن‌ها برابر با ۲۱ می‌باشد که نشان‌دهنده این است که نیاز به برآورد ۸ وزن در مدل نبوده است و مدل مورد استفاده ساده‌تر است.

نتایج شبکه‌های عصبی در رویکرد تحلیل مؤلفه اصلی: تحلیل مؤلفه اصلی روی متغیرهای ورودی انجام شد. جدول ۳ نتایج برازش شبکه عصبی روی مؤلفه‌های اصلی استخراج شده از هر مجموعه داده را نشان می‌دهد. در هر مجموعه، تعداد مؤلفه‌های استخراجی برای به دست آوردن ۹۹ درصد از کل واریانس نشان داده شده است. در جدول ۳ مشاهده می‌شود که در گروه تلقیح یک، ۵ مؤلفه اصلی ۹۹ درصد واریانس موجود در داده‌ها را بیان می‌کردند. شبکه عصبی برازش یافته با ۴ نورون پنهان و بر روی این ۵ مؤلفه دارای ضریب تعیین ۰/۹۹۴۱ و ۰/۹۹۴۲ و جذر میانگین مربعات ۲/۲۰۲۶ و ۲/۱۳۹۵ به ترتیب در مجموعه آموزش و آزمون می‌باشد. در بقیه گروه‌های تلقیح نیز تفسیر نتایج جدول به همین صورت می‌باشد. همان‌طور که در جدول ۳ مشاهده می‌شود، تعداد مؤلفه‌های اصلی در تلقیح‌های مختلف متفاوت بوده که این شاید ساختار خاص داده‌ها را در ارتباط با فاصله گوساله‌زایی نشان دهد.

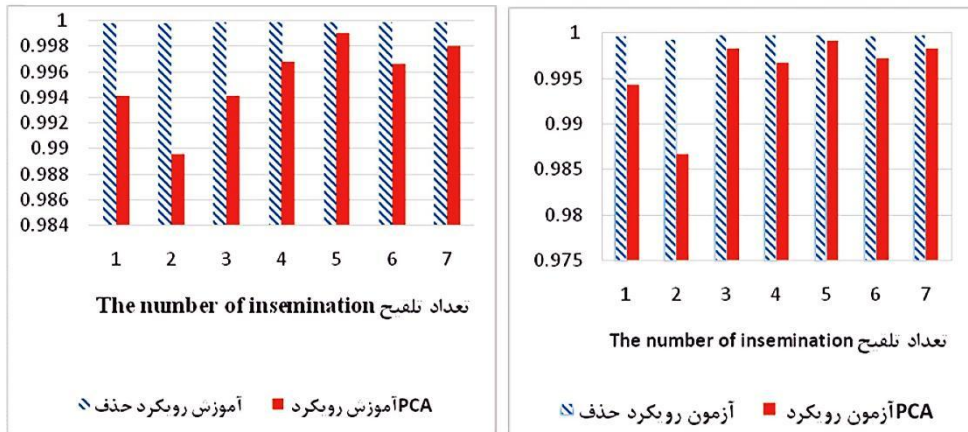
نشریه پژوهش در نشخوارکنندگان (۳)، شماره (۴) ۱۳۹۴

جدول ۳- نتایج شبکه‌های عصبی با روش تحلیل مؤلفه‌های اصلی در تلقیح‌های مختلف.

Table 3. Results of neural networks using principle component analysis method in different inseminations.

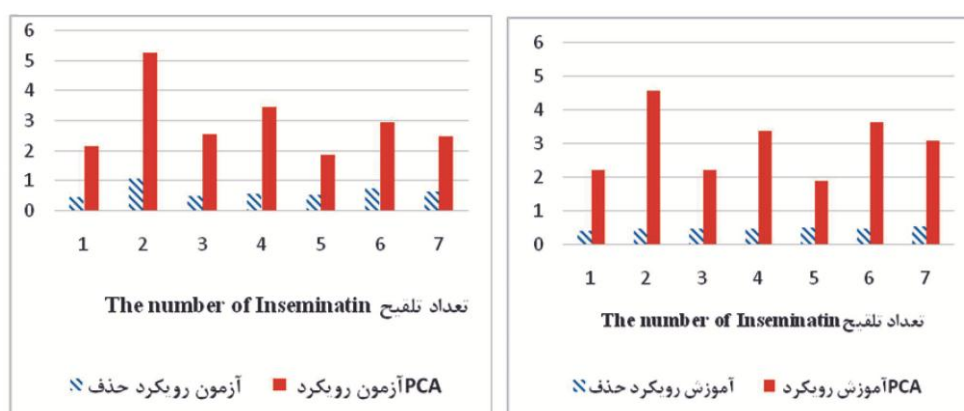
آزمون ۲۰ درصد (20% testing)			آموزش ۸۰ درصد (80% training)			درصد واریانس تبیین شده	تعداد واحد پنهان	تعداد مؤلفه‌های اصلی	تعداد تلقیح
RMSE	R ²	تعداد NO	RMSE	R ²	تعداد NO	The percentage of explained variance	No. of hidden units	No. of principal components	No. of insemi nation
2.1395	0.9942	755	2.2026	0.9941	3028	99	4	5	1
5.2509	0.9866	830	4.5756	0.9895	6332	99	5	5	2
2.5366	0.9983	535	2.2026	0.9941	2155	99	7	6	3
3.4629	0.9966	337	3.3865	0.9968	1135	99	7	6	4
1.8594	0.9991	203	1.8831	0.9990	815	99	3	6	5
2.9540	0.9971	128	3.6337	0.9965	517	99	7	6	6
2.4775	0.9982	77	3.0711	0.9980	314	99	6	6	7

مقایسه ضریب تعیین دو رویکرد حذف و تحلیل مؤلفه اصلی در مجموعه آموزش (سمت راست) و مجموعه آزمون (سمت چپ) در شکل ۲ نشان داده شده است. مقادیر ضریب تعیین در مجموعه آموزش و آزمون در هر دو رویکرد بالای ۰/۹۸ است و رویکرد حذف در تمام تلقیح‌ها دارای ضریب تعیین بالاتری است. پس هر دو رویکرد از نظر ضریب تعیین مناسب هستند ولی مدل با رویکرد حذف از دقت بالاتری برخوردار است. در استخراج شبکه چه بر اساس رویکرد حذف و چه بر اساس تحلیل مؤلفه اصلی، با افزایش تعداد تلقیح، ضریب تعیین افزایش و دقت برآورد شبکه دقیق‌تر شده است. البته همان‌طور که مشخص است اختلاف دقت در حد کمتر از ۰/۰۱ است.



شکل ۲- مقایسه ضریب تعیین (R^2) دو رویکرد در مجموعه آموزش (سمت راست) و مجموعه آزمون (سمت چپ).
Figure 2. Comparison the coefficient of determination (R^2) of two approaches in the training set (left) and test set (right).

مقایسه معیار جذر میانگین مربعات خطا در دو رویکرد حذف و تحلیل مؤلفه اصلی در مجموعه آموزش (سمت چپ) و مجموعه آزمون (سمت راست) در شکل ۳ نشان داده شده است. مقادیر جذر میانگین مربعات خطا مربوط به مجموعه آموزش و آزمون در رویکرد حذف کمتر از یک بوده ولی در رویکرد تحلیل مؤلفه اصلی همگی بالای ۱/۵ هستند و در تلفیق دوم این معیار در هر دو رویکرد بیشینه می‌شود. این نتایج نشان می‌دهد که رویکرد حذف مناسب‌تر می‌باشد.



شکل ۳- مقایسه جذر میانگین مربعات خطای دو رویکرد در مجموعه آموزش (سمت چپ) و مجموعه آزمون (سمت راست).
Figure 3. Comparison between RMSE of the two approaches in the training set (left) and test set (right).

اندازه‌گیری رابطه خطی و غیر خطی متغیرهای ورودی با متغیر خروجی: در محاسبه معیار ضریب اطلاعات بیشین از اطلاعات متقابل^۱ بین دو متغیر استفاده می‌شود. در جدول ۴ میزان تأثیر متغیرهای ورودی بر متغیر فاصله گوساله‌زایی توسط معیار ضریب اطلاعات بیشین تعیین شده است. همان‌طور که در جدول ۴ نشان داده شده است، متغیرهای روزهای شیردهی و کل شیری تولیدی، کل چربی تولیدی و کل پروتئین تولیدی، بیشترین میزان ارتباط آماری را با فاصله گوساله‌زایی نشان می‌دهند. این روند ارتباطی تقریباً با افزایش تعداد تلقیح‌ها افزایش یافت. همان‌طور که ملاحظه می‌شود، ارتباط بین متغیرهای ورودی و خروجی با استفاده از شبکه عصبی و در تلفیق‌های مختلف تعیین گردید.

1- Mutual information

نشریه پژوهش در نشخوار کنندگان (۳)، شماره (۴) ۱۳۹۴

جدول ۴- میزان رابطه خطی و غیرخطی با استفاده از معیار ضریب اطلاع بیشین.

Table 4. Measurement of linear and nonlinear relationship with maximal information coefficient.

متغیرها							
Variables							
طول دوره خشکی Length of dry off period	کل شیر تولیدی Total Milk Production	کل چربی تولیدی Total Fat Production	کل پروتئین تولیدی Total Protein Production	روزهای شیردهی Days in Milk	شیر تصحیح شده Adjusted Milk (ME-305-2x)	چربی تصحیح شده Adjusted Fat (ME-305-2x)	پروتئین تصحیح شده Adjusted Protein (ME-305-2x)
تعداد تلقیح No. of Insemination							
1	0.0436	0.1179	0.1024	0.1012	0.5875	0.1041	0.0972
2	0.0761	0.1477	0.1318	0.1115	0.6987	0.0876	0.0905
3	0.0914	0.1904	0.1753	0.1490	0.7888	0.1034	0.1049
4	0.1028	0.1955	0.1927	0.1725	0.8017	0.1078	0.1190
5	0.1278	0.2403	0.2036	0.1739	0.8300	0.1367	0.1432
6	0.1584	0.2254	0.2716	0.1904	0.8809	0.1671	0.1455
7	0.1804	0.3109	0.2360	0.2555	0.8539	0.1783	0.1530

انجوبی و همکاران (۲۰۰۹) توانستند رکوردهای تولید شیر اولین دوره شیردهی گاوها را با استفاده از شبکه عصبی پیش‌بینی کنند (۲۱). بهترین توپولوژی شبکه عصبی مورد استفاده این محققین شامل یک لایه پنهان به همراه ۸ نورون و تابع انتقال از نوع تانزانت سیگموئیدی^۱ بود. آن‌ها نشان دادند که با وارد کردن اطلاعات والد نر به شبکه، معیار جذر میانگین مربعات خطا کاهش یافته است. برای بهینه‌سازی شبکه عصبی باید تعداد لایه‌های پنهان و تعداد نورون‌های موجود در آن تعیین گردد. انتظار می‌رود که هرچه تعداد لایه‌های پنهان بالا رود، شبکه بهینه‌تر گردد. اما افزایش تعداد لایه‌های پنهان همراه با افزایش هزینه محاسباتی است (۱۴). در مطالعه کومار و شارما (۲۰۱۴) برای بررسی میزان تولید شیر و طول عمر، شبکه عصبی با لایه‌های مختلف آموزش داده شد. در این پژوهش از دو تابع انتقال متفاوت تانسینگ^۲ و کاملاً خطی^۳ به ترتیب در لایه‌های اول و دوم مخفی استفاده گردید (۱۸). از شبکه عصبی برای گروه‌بندی متغیرها نیز استفاده می‌شود. در مطالعه‌ای که روی میزان سخت‌زایی در گاو شیری صورت گرفت، تعداد ۱۵ متغیر با ماهیت‌های متفاوت هورمونی، مدیریتی و تولیدی به عنوان متغیرهای ورودی شبکه عصبی به کار برده شدند. هدف این بود که مشخص گردد چه تعداد از متغیرها باعث بروز سخت‌زایی می‌گردند. نشان داده شد که متغیرهای فصل گوساله‌زایی، ژنوتیپ ژن

1- Tangent sigmoid

2- Tansig

3- Pure line

CYP19-PvuII و طول آبستنی بیشترین ارتباط را با بروز سخت‌زایی داشتند. در این مطالعه تابع شعاعی پایه^۱ در دو توپولوژی مختلف با یک و دو لایه مخفی به کار برده شد (۲۰). همچنین گزارش شده است که در گروه‌بندی متغیرهای مؤثر در عملکرد تلقیح مصنوعی، استفاده از رگرسیون سازش‌پذیر اسپلاین چند متغیره^۲ و شبکه عصبی بهتر از رگرسیون لجستیک و توابع گروه‌بندی^۳ عمل می‌کند (۷). با توجه به نتایج حاصل، پیشنهاد می‌شود که این‌گونه آموزش‌های شبکه عصبی روی داده‌های حاصل از واحدهای پرورش گاو شیری اعمال گردد. همچنین نیاز به استفاده از متغیرهای ورودی که در این مطالعه به کار نرفتند نیز احساس می‌شود. مثلاً ثابت شده است که سن اولین زایش نیز یکی از عوامل مؤثر بر عملکرد تولیدمثلی می‌باشد که تأثیر آن بر عملکرد تولیدمثلی مشابه با دوره شیردهی می‌باشد (۱۷).

نتیجه‌گیری کلی

به‌طور کلی نتایج این پژوهش نشان داد که متغیرهای به‌کار رفته در این مطالعه توانسته است پیش‌بینی خوبی از فاصله گوساله‌زایی ارائه دهد. از طرفی مشخص شد که این معیارها در تلقیح‌های مختلف تقریباً یکسان است و از این‌رو نباید انتظار داشت که مدل‌سازی تلقیح‌های مختلف، جواب متفاوتی ارائه کند. در این مطالعه مشخص شد که با افزایش تعداد تلقیح‌ها، روند ارتباطی متغیرها با فاصله گوساله‌زایی تغییر نمی‌کند. بعضی از این متغیرها مثل روزهای باز ارتباط ساختاری زیادی با فاصله گوساله‌زایی دارند. لذا پیشنهاد می‌شود با استفاده از روش‌های بیزی^۴ و لاسو^۵ ابتدا انتخاب متغیر صورت گیرد و سپس آموزش شبکه عصبی انجام پذیرد. همچنین با توجه به این که شبکه‌های عصبی مصنوعی نیاز به فرضیه‌ها و مدل‌های از پیش تعیین شده ندارند، می‌توانند به‌عنوان جایگزینی برای مدل‌های رگرسیون مطرح شوند.

سپاسگزاری

نویسندگان از مدیریت محترم مجتمع کشت و دامداری فکا، آقای مهندس جلیل‌نژاد که داده‌های این تحقیق را در اختیار ما گذاشتند، تشکر و قدردانی می‌نمایند.

- 1- Radius Base Function
- 2- Multivariate Adaptive Regression Spline
- 3- Classification Functions
- 4- Bayesian
- 5- Lasso

منابع

1. Ansari-Lari, M., Rezagholi, M., and Reiszadeh, M. 2009. Trends in calving age and calving intervals for Iranian Holstein in Fars province, Southern Iran. *Trop. Anim. Health Prod.* 41: 1283-1288.
2. Borecki, M., Korwin-Pawłowski, M.L., Bełłowska, M., Szmidt, M., Urbańska, K., Kalenik, J., Chudzian, L., Szczepański, Z., Kopczyński, K., Jakubowski, A., and Szmidt, A.J. 2010. Capillary microfluidic sensor for determining the most fertile period in cows. *Acta Physica Polonica A.* 118(6): 1093-1099.
3. Cavero, D., Tölle, K.H., Henze, C., Buxadé, C., and Krieter, J. 2008. Mastitis detection in dairy cows by application of neural networks. *Livest. Sci.* 114: 280-286.
4. Chen, L.J., Cui, L.Y., Xing, L., and Han, L.J. 2008. Prediction of the nutrient content in dairy manure using artificial neural network modeling. *J. Dairy Sci.* 91: 4822-4829.
5. Chen, L.J., Xing, L., and Han, L.J. 2009. Quantitative determination of nutrient content in poultry manure by near infrared spectroscopy based on artificial neural networks. *Poult. Sci.* 88: 2496-2503.
6. Curtis, C.R., Erb, H.N., Sniffen, C.J., Smith, R.D., and Kronfeld, D.S. 1985. Path analysis of dry period nutrition, postpartum metabolic and reproductive disorders, and mastitis in Holstein cows. *J. Dairy Sci.* 68: 2347-2360.
7. De Vries, A. 2006. Economic value of pregnancy in dairy cattle. *J. Dairy Sci.* 89: 3876-3885.
8. Finn G.D., Lister, R., Szabo, R., Simonetta, D., Mulder, H., and Young, R. 1996. Neural Networks applied to a large biological database to analyse dairy industry pattern. *Neural Comput. Appl.* 4: 237-253.
9. Ghiasi, H., Pakdel, A., Nejati-Javaremi, A., Mehrabani-Yeganeh, H., Honarvar, M., Gonzalez- Recio, O., Jesus Carabano, M., and Alenda, R. 2011. Genetic variance components for female fertility in Iranian Holstein cows. *Livest. Sci.* 139(3): 277-280.
10. Gorgulu, O. 2012. Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *S. Afr. J. Anim. Sci.* 42(3): 280-287.
11. Grzesiak, W., Lacroix, R., Wójcik, J., and Blaszczyk, P. 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Can. J. Anim. Sci.* 83: 307-310.
12. Grzesiak, W., Zaborski, D., Sablik, P., Żukiewicz, A., Dybus, A., and Szatkowska, I. 2010. Detection of cows with insemination problems using selected classification models. *Comput. Electron. Agr.* 74(2): 265-273.
13. Hassan, K.J., Samarasinghe, S., and Lopez-Benavides, M.G. 2009. Use of neural networks to detect minor and major pathogens that cause bovine mastitis. *J. Dairy Sci.* 92: 1493-1499.

14. Haykin, S. 1999. Neural Networks, A comprehensive Foundation. Prentice Hall, Upper saddle River, NJ. Pp: 156-158.
15. Hosseinnia, P., Edrisi, M., Edriss, M.A., and Nilforooshan, M.A. 2007. Prediction of second parity milk yield and fat percentage of dairy cows based on first parity information using neural networks system. J. Appl. Sci. 7: 3274-3279.
16. Krieter, J., Stamer, E., and Junge, W. 2006. Control charts and neural networks for oestrus detection in dairy cows. Lecture Notes in Informatics. 133-136.
17. Kuhn, M. 2008. Building predictive models in R using the caret package. J Stat Softw. 28(5): 1-26.
18. Kumar, S., and Sharma, S.K. 2014. Anticipating milk yield using artificial neural network. Int. J. Appl. Sci. Engeer. Res. 3(3): 690-695.
19. Krogh, A., and Hertz, J.A. 1992. A simple weight decay can improve generalization. In Advances in neural information processing systems. 4: 950-957.
20. Nielen, M., Spigt, M.H., Schukken, Y.H., Deluyker, H.A., Maatje, K., and Brand, A. 1995. Application of a neural network to analyse on-line milking parlour data for the detection of clinical mastitis in dairy cows. Prev. Vet. Med. 22: 15-28.
21. Njubi, D.M., Wakhungu, J., and Badamana, M.S. 2009. Milk yield prediction in Kenyan Holstein-Friesian cattle using computer neural networks system. Livest. Res. Rural Dev. 21(4): 46-51.
22. Pahlavan Afshar, K., Honarvar, M., and Lavvaf, A. 2012. Estimated economic value of estrous detection using bio-economic model in Iranian Holeystein herds. J. Anim. Sci. Res. 12: 37-50. (In Persian)
23. Pastell, M.E., and Kujala, M. 2007. A probabilistic neural network model for lameness detection. J. Dairy Sci. 90: 2283- 2292.
24. Ripley, B.D. 1995. Statistical ideas for selecting network architectures. In Neural networks: Artificial intelligence and industrial applications. Springer. Pp: 183-190.
25. Salehi, F., Lacroix, R., and Wade, K.M. 1998. Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. Comput. Electron. Agr. 20: 199-213.
26. Salehi, F., Lacroix, R., Yang, X.Z., and Wade, K.M. 1997. Effects of data preprocessing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. Trans. ASAE, 40(3): 839-846.
27. Sanzogni, L., and Kerr, D. 2001. Milk production estimates using feed forward artificial neural networks. Comput. Electron. Agric. 32: 21-30.
28. Shahabifar, M. 2005. The ore reserve estimate by artificial neural network technique. Iran mining Engineering Conference, Tarbiat modares university, Tehran, Iran. (In Persian)

29. Shahinfar, S., Mehrabani-Yeganeh, H., Lucas, C., Kalhor, A., Kazemian, M., and Weigel, K.A. 2012. Prediction of breeding values for dairy cattle using artificial neural networks and neuro-fuzzy systems. *Comput. Math. Methods Med.* 2012: 127130.
30. Slama, H., Wells, M.E., Adams, G.D., and Morrison, R.D. 1996. Factors affecting calving interval in dairy herds. *Anim. Sci.* 59: 1334-1339.
31. Sun, Z. 2008. Application of artificial neural networks in early detection of mastitis from improved data collected on-line by robotic milking stations. Ph.D. Dissertation, Lincoln University, New Zealand.
32. Toghdiani Pozveh, S., Shadparvar, A.A., Moradi Shahrabak, M., and Dadpasand Taromsari, M. 2009. Genetic analysis of reproduction traits and their relationship with conformation traits in Holstein cows. *Livest. Prod. Sci.* 125: 84-87.
33. Tumer, K., and Ghosh, J. 1996. Analysis of decision boundaries in linearly combined neural classifiers. *Pattern Recog.* 29(2): 341-348.
34. Wang, E., and Samarasinghe, S. 2005. On-Line Detection of Mastitis in Dairy Herds Using Artificial Neural Networks. 273-278.
35. Yang, X.Z., Lacroix, R., and Wade, K.M. 1999. Neural detection of mastitis from dairy herd improvement records. *Trans. ASAE.* 42: 1063-1071.
36. Yang, X.Z., Lacroix, R., and Wade, K.M. 2000. Investigation into the production and conformation traits associated with clinical mastitis using artificial neural networks. *Can. J. Anim. Sci.* 80: 415-426.
37. Zaborski, D., and Grzesiak, W. 2011. Detection of difficult calving in dairy cows using neural classifier. *Archiv Tierzucht*, 54(5): 477-489.



Gorgan University of Agricultural
Sciences and Natural Resources
J. of Ruminant Research, Vol. 3(4), 2016
<http://ejrr.gau.ac.ir>

Association between calving interval and productive traits in dairy cattle over different inseminations using artificial neural network

***M. Ghaderi Zefrehei¹, M.R. Bahreini Behzadi¹, M. Fayaz² and S. Sharifi³**

¹Assistant Prof., Dept. of Animal Sciences, Faculty of Agriculture, Yasouj University,

²Ph.D. Student, Dept. of Biology, Faculty of Basic Sciences, Shahid Beheshti University,

³Ph.D. Student, Dept. of Animal Sciences, Faculty of Agriculture, Isfahan University of Technology

Received: 21/12/2015; Accepted: 17/03/2016

Abstract

Background and objectives: Calving interval is sought to be substantially affected by management and environmental effects. Using intellectual algorithms of machine learning methods to investigate complex systems are growing and these algorithms could be assumed as right approach to analysis dairy cattle industry data. Artificial neural networks which are part of artificial intelligence, gain from these algorithms. The objective of this study was to find the association between productive and reproductive traits with calving interval over different course of inseminations. It was assumed that extracting this association may improve the management of this trait.

Materials and methods: In this research, the productive and reproductive traits from FOKA, an agriculture and animal husbandry, associated to Isfahan Vahdat Cooperative was used. The records were calving interval (day), the length of dry off period (day), number of insemination, total milk production (kg), total fat production (kg), total protein production (kg), adjusted milk production (kg), adjusted fat production (kg) and adjusted protein production (kg). The data were due to 15465 cows in which their parturition date spanned between 1368 and 1393. The data dimensions were reduced using elimination of high level of correlation among variables and principal component analysis before undertaking artificial neural network modeling. For each insemination, we learned a neural network. We used coefficient of determination and root mean square error to investigate the

*Corresponding author: mosmos741@yahoo.com

efficiency of neural network. The linear and nonlinear relationship among input variables with calving interval was measured using maximal information criterion.

Results: The similar values of coefficient of determination and root mean square error due to neural network in different insemination obtained which would indicated that there is almost identical associations between calving interval input variables. Higher values of root mean square error obtained in neural network which learned based upon principal component analysis of independent data than elimination one. Therefore, it was concluded that elimination approach is suitable choice in this context. The maximal information criterion showed almost an identical association between calving interval and input variables over different inseminations. In this way, days in milk and milk production showed high amount of correlation with calving interval.

Conclusion: The results indicated that coefficient of determination and root mean square error of different inseminations was similar. Therefore, it is hardly expected that the association between calving interval and independent variables to be different over different inseminations. Based upon these results, it was elucidated that independent variables used in this study could fairly adequately predict the calving interval. Also, we could say that management – environmental conditions over different inseminations had minuscule effect on the associations of calving interval with independent variables.

Keywords: Principal component analysis, artificial neural network, Maximal information coefficient, Calving interval

