



# دامستیک

انجمن علمی - دانشجویی گروه علوم دامی دانشگاه تهران؛ بهار ۱۳۹۹

[https://domesticsj.ut.ac.ir/article\\_76952.html](https://domesticsj.ut.ac.ir/article_76952.html)

## مقاله مروری

### آبر داده‌ها و نقش جدیدترین تکنولوژی‌ها در اصلاح نژاد دام و طیور

فرزاد غفوری<sup>۱\*</sup>، حسن مهربانی یگانه<sup>۲</sup>، صادق محمدیان جشوقانی<sup>۳</sup><sup>۱</sup> دانشجوی دکتری تخصصی ژنتیک و اصلاح نژاد دام و طیور، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران<sup>۲</sup> دانشیار ژنتیک و اصلاح نژاد دام و طیور، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران<sup>۳</sup> کارشناسی ارشد ژنتیک و اصلاح نژاد دام، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران<https://doi.org/10.22059/domesticsj.2020.76952>

## چکیده

دیدگاه کلی ژنتیک و اصلاح نژاد دام و طیور در حال انتقال به عصر دیجیتال با توان عملیاتی بالا است که در آن با استفاده از فناوری‌های جدید سعی می‌شود دقت ثبت اطلاعات و برآورد ارزش‌های اصلاحی افزایش یابد. در ادامه با استفاده از نرم‌افزارهای پیشرفته و کامپیوترهای بزرگ، پردازش حجم بزرگی از داده‌ها انجام می‌شود. ظهور فناوری تعیین ژنوتایپ و شناسایی SNPها همراه با روش‌های آماری جدید برای استفاده از این داده‌ها برای برآورد ارزش اصلاحی، منجر به کاربرد گسترده انتخاب ژنومی در گاوهای شیری و طیور برای انجام انتخاب ژنومیک در دیگر موارد شده است. توسعه الگوریتم‌های داده کاوی مرتبط با آبر داده‌ها در برآورد ارزش‌های اصلاحی نقش قابل توجهی ایفا می‌کنند. مجموعه‌ای از تکنولوژی‌ها همچون هوش مصنوعی، یادگیری ماشین و یادگیری عمیق در عصر جدید فرصت‌های مناسبی را در مقایسه با روش‌های سنتی برای بررسی صفات اقتصادی با معماری پیچیده فراهم ساخته‌اند. این رویکردها تجزیه و تحلیل مجموعه‌های بزرگ داده‌ها و اطلاعات بزرگ ژنومی را در جهت رسیدن به اهداف امکان پذیر کرده است. هدف از این مطالعه بیان توضیحی اجمالی از روش‌ها و فناوری‌های جدید در علوم دامی است که به طور گسترده در رکوردبرداری و ثبت اطلاعات تا برآورد ارزش‌های اصلاحی مورد استفاده قرار می‌گیرند، به گونه‌ای که دیدگاه اصلاح نژادی را در آینده دیجیتال تغییر دهند. بنابراین، افزایش پتانسیل در تجزیه و تحلیل آبر داده‌ها همراه با روش‌های نوین در رکوردبرداری از صفات فنوتیپی و برآورد ارزش‌های اصلاحی، مقدار پیشرفت ژنتیکی دام‌ها را به طور چشمگیری افزایش خواهد داد.

**کلمات کلیدی:** آبر داده، اصلاح نژاد، تکنولوژی، هوش مصنوعی، یادگیری ماشین، یادگیری عمیق

\* نویسنده مسئول: farzad.ghafouri@ut.ac.ir

تاریخ دریافت: ۱۳۹۹/۰۳/۱۸ تاریخ بازنگری: ۱۳۹۹/۰۳/۲۲ تاریخ پذیرش: ۱۳۹۹/۰۳/۳۰ تاریخ انتشار آنلاین: ۱۳۹۹/۰۳/۳۱



## مقدمه

هدف اصلی در این مطالعه ارائه توضیحی اجمالی از روش‌های جدید پیش‌بینی شده است که تاکنون در دامپروری و اصلاح‌نژاد دام و طیور به صورت گسترده چه در رکوردبرداری و ثبت اطلاعات و چه در برآورد ارزش‌های اصلاحی مورد استفاده قرار گرفته است. در ادامه روندهای بالقوه و جهت‌گیری‌های جدید تحقیقاتی همچون تجزیه و تحلیل داده‌های بزرگ (با حجم بالا) با استفاده از تکنیک‌های هوش مصنوعی، یادگیری ماشینی، شبکه عصبی مصنوعی و سایر تکنولوژی‌های روز اشاره می‌شود؛ به گونه‌ای که می‌توانند دیدگاه اصلاح‌نژاد دام و طیور را در آینده دیجیتال، تغییر دهند.

## آینده دیجیتال و عصر آبر داده‌ها

دیدگاه کلی اصلاح‌نژاد دام و طیور در حال گذر از دوره کلاسیک به سمت دوره دیجیتال با توان بالا است که در آن مقدار زیادی از اطلاعات به صورت لحظه‌ای با استفاده از نرم‌افزارهای پیشرفته و کامپیوترهای بزرگ، ذخیره و تجزیه و تحلیل می‌شوند و تصمیم‌گیری‌های عملی در استراتژی‌های اصلاح‌نژادی براساس نتایج خروجی صورت می‌گیرد (Nayeri et al., 2019).

در حال حاضر اصلاح‌نژاد دام و طیور از قاعده توسعه روش‌های محاسباتی کارآمدتر و سریع‌تر مستثنا نیست و اکنون با انتقال روش‌ها و برنامه‌های جدید و مفهومی با یک دوره محاوره‌ای به نام "عصر بزرگ داده" روبرو است. در حالی که منابع اطلاعاتی سنتی برای اصلاح‌نژاد در حیوانات شامل اطلاعات فنوتیپی و شجره‌ای است، امروزه هجوم وسیعی از داده‌های ژنومی متشکل از چندشکلی‌های تک نوکلئوتیدی (SNP)، حاشیه‌نویسی ژن، مسیرهای متابولیکی، شبکه‌های متعامل پروتئینی، بیان ژن و اطلاعات مربوط به ساختار پروتئین‌ها وجود دارد. استفاده از داده‌های SNP برای پی بردن به معماری ژنوم تا فنوم صفات تنها می‌تواند بخشی از واریانس فنوتیپی را توجیه نماید؛ به همین دلیل فراهم آوری تمام اطلاعات امیکس که به طور خاص نیازمند شبکه‌های عصبی عمیق برای تحلیل ارتباطات بین لایه‌های مختلف است امری ضروری به نظر می‌رسد.

مطالعات بسیار زیادی در رابطه با روش‌های آماری جدید وجود دارد که بر اساس آن‌ها قابلیت تجزیه و تحلیل مدل‌های پیچیده‌تر با کارایی بیشتر به ویژه برای آبر داده‌ها وجود خواهد داشت، به گونه‌ای مدت زمان لازم برای محاسبات به طور قابل توجهی کاهش یابد. امروزه بیشتر محققان و پژوهشگران علم اصلاح‌نژاد دام و طیور به دنبال روش‌های محاسباتی کارآمدتری هستند که بتوانند با استفاده از ظرفیت محاسباتی کمتر، پردازش سریع‌تر نمونه داده‌ها را انجام دهند.

به منظور بهبود و پیشرفت در اصلاح‌نژاد، اصلاح‌گران باید برآورد پارامترهای ژنتیکی مانند میزان تغییرات ژنتیکی در برخی خصوصیات را رصد نمایند و اطلاعات فنوتیپی و ژنوتیپی با دقت و صحت بالایی ثبت گردند. با در نظر گرفتن تغییرات موجود در منابع مختلف، می‌توان از مدل‌های مختلط متناسب با یک استراتژی راهبردی استفاده کرد و برای هر منبع تغییر باید یک پارامتر آماری به نام واریانس تخمین زده شود (CM Dekkers, 2012).

روش‌های مورد استفاده برای ارزیابی پارامترهای واریانس و مدل‌هایی که ویژگی‌ها و برتری‌های حیوانات را نشان می‌دهند، "هسته پژوهش" را تشکیل می‌دهند. با این وجود، با توجه به نمونه‌های بزرگ داده‌ها و مدل‌های پیچیده‌ای که اغلب در ارزیابی‌های دامپروری مورد استفاده قرار می‌گیرند، بازده محاسباتی یکی از ویژگی‌های اصلی روش‌های تخمین است (Pérez-Enciso, 2017). راه‌حل‌های سریع‌تر برای مدل‌های پیچیده استفاده از نرم‌افزارهای پر قدرت در دسترس برای ارزیابی منحنی‌های مختلف هر پارامتر است. به گونه‌ای که نرم‌افزارها و برنامه‌های جدید هم از نظر حافظه رایانه و هم از نظر مدت زمان محاسبه باید کارآمد باشند. بدان معنی که پارامترهای واریانس ژنتیکی مورد نیاز در برنامه‌های انتخاب حیوانات می‌توانند در آینده سریع‌تر و با صحت بالاتری محاسبه شوند (Chu and Corey, 2012).

شناسایی و یافتن حیوانات (به عنوان مثال حیوانات فحل در مزرعه) استفاده کرد (Britt et al., 2018).

امروزه عصر دیجیتال با حجم بسیار زیادی از داده‌ها همراه است. به گونه‌ای که با استفاده از داده‌های با حجم بسیار بالا، بخش پرورش و اصلاح نژاد دام کارآمدتر می‌شود. اینترنت اشیا (IoT: Internet of Things) همراه با هوش مصنوعی و یادگیر ماشین و پردازش تصویر فناوری‌های جدیدی هستند که با هم گره خورده‌اند تا پرورش، مدیریت و اصلاح نژاد دام با سهولت بیشتری انجام گیرد. اینترنت اشیا دنیای فیزیکی (اشیا و تجهیزات محیط پیرامون) را با دنیای دیجیتال پیوند می‌دهد و قابلیت‌هایی را در بخش اصلاح نژاد ارائه می‌دهد. با این وجود در پیش‌بینی وقایع آینده، پشتیبانی از تصمیم‌گیری و بهبود سیستم‌های یاری‌رسان خواهد بود (Perera et al., 2015). اینترنت اشیا با اطلاع‌رسانی سریع، هشدار می‌دهد که چگونه و چه زمانی در تغذیه، تولیدمثل، بهداشت و مراقبت از حیوانات مداخله کنند و جهت رفع مشکل قدم بردارند. همچنین از برنامه‌های تلفن همراه متصل به پیشرفته‌ترین نرم‌افزارهای مدیریتی، برای ردیابی سلامت حیوانات و افزایش بهره‌وری استفاده می‌شود.

فناوری‌های هوش مصنوعی (AI: Artificial Intelligence) و رباتیک (Robotic) جدیدترین فناوری‌هایی هستند که در عصر جدید در اصلاح نژاد دام و طیور از آن‌ها استفاده می‌شود. با سرمایه‌گذاری صنایع مختلف در هوش مصنوعی ارزش و فناوری آن همچنان در حال افزایش است؛ به گونه‌ای که در سال ۲۰۱۶ درآمد سالانه از سیستم‌های فعال شده با هوش مصنوعی تا سال ۲۰۲۵ از ۱/۴ میلیارد دلار به ۵۹/۸ میلیارد دلار افزایش می‌یابد. از یادگیری ماشینی به طور فزاینده‌ای در بسیاری از فناوری‌های جدید بهره گرفته می‌شود و الگوریتم‌های رایانه‌ای مانند الگوریتم ژنتیک بینش عمیقی از معیارهای تجاری و بهبود تصمیم‌گیری مبتنی بر داده‌ها را ارائه می‌دهند. استفاده از ربات‌های ثبت‌کننده رکوردها و اطلاعات مانند ربات‌های شیردوش، نقش قابل توجهی در

الگوریتم‌های نوین یادگیری عمیق تمام اطلاعات امیکس افراد را طبق لایه‌بندی‌های مختلف دریافت نموده و با استفاده از عملگرهای ریاضی اقدام به بررسی ارتباطات بین این لایه‌ها می‌کنند (Yang Ya-Ian et al., 2017).

صنعت دام نیز در دهه گذشته تحت تأثیر تغییرات اساسی در انتخاب ژنتیکی، مدیریت گله و از همه مهم‌تر جمع‌آوری داده‌های مبتنی بر فناوری‌های نوین بر پایه پردازش سیگنال قرار گرفته است. فناوری سنسور با نظارت بر فعالیت دام‌ها و گله‌های بزرگ و ارسال هشدار برای سلامتی، باروری و یا جمع‌آوری فنوتیپ‌های با اندازه‌گیری سخت‌گیرانه، پتانسیل بزرگی برای افزایش بهره‌وری اصلاح نژاد دارد. این مجموعه داده‌های بزرگ به همراه سایر منابع ژنومی فرصتی برای ایجاد مدل‌های پیش‌بینی سلامت، باروری و سایر رویدادها به وجود می‌آورد (Nayeri et al., 2019).

### فناوری‌های جدید در افزایش دقت ثبت اطلاعات

اصلاح‌گران دام و طیور برای تأمین تقاضای آینده جمعیت ۹/۸ میلیارد نفری انسان تا سال ۲۰۵۰، امروزه از فناوری و نوآوری‌های جدید زیادی استفاده می‌کنند. به گونه‌ای که این فناوری‌ها سریع‌تر، کارآمدتر و در دسترس‌تر از گذشته می‌باشند. پیش‌بینی می‌شود که تولید گوشت و فرآورده‌های دامی و طیور در سال‌های آینده با توجه به تغییر رژیم غذایی و افزایش سطح درآمد چند برابر شود، که این افزایش میزان تولید، لازمه آن استفاده از روش‌های مدرن اصلاح نژادی است. استفاده از فناوری‌های جدید در کشاورزی و دامپروری مانند تکنولوژی RFID، حسگرهای بیولوژیک و GPS امکان نظارت بر حیوانات مزرعه‌ای را فراهم می‌کند. این فناوری‌های نوین حجم عظیمی از داده‌ها را به دنبال خواهند داشت. همچنین سایر فناوری‌های جدید در دامداری‌ها شامل استفاده از هواپیماهای بدون سرنشین برای نظارت بر دام‌ها در کشاورزی توسعه یافته و به صورت گسترده مورد استفاده قرار می‌گیرد که در آینده می‌توان با استفاده از آن‌ها تکنولوژی‌های همراه مانند GPS را برای

الگوریتم‌های داده کاوی (تجزیه و تحلیل داده‌ها) مرتبط با هر نوع داده برای مجموعه‌های بزرگ و حجم عظیمی از داده‌ها ممکن است به شناسایی روابط غیرمنتظره و ناشناخته در داده‌ها کمک شایانی داشته باشد. در حالی که انتخاب ژنومی با استفاده از داده‌های با حجم بسیار زیاد به قدرت محاسباتی بالایی نیاز دارد، به خصوص زمانی که بخش عظیمی از جمعیت تعیین ژنوتایپ می‌شوند. پیشرفت تئوری وارون ماتریس‌های با روابط عددی بزرگ، امکان محاسبات و حل سیستم‌های بزرگ مطالعات را مجاز ساخته است و الگوریتم‌های با عملکرد بالا برای برآورد مؤلفه‌های واریانس متناسب با ساختار داده‌ها طراحی شده‌اند (Cole *et al.*, 2012).

اخیراً نشان داده شده است که رویکردهای تک مرحله‌ای ترکیب BLUP با ماتریس روابط ژنومی (G) الزامات محاسباتی مشابه با BLUP سنتی را دارند و عامل محدودکننده‌ای جهت ساخت و وارون ماتریس G برای بسیاری از ژنوتایپ‌ها می‌باشد. یک الگوریتم ساده جهت ایجاد ماتریس G برای ۱۴۰۰۰ فرد تقریباً ۲۴ ساعت زمان نیاز دارد، در صورتی که الگوریتم‌های پیچیده‌تر و پیشرفته‌تر مدت زمانی مورد نظر را به حداقل مقدار ممکن کاهش داده‌اند. البته در این میان باید عنوان کرد که مجموعه داده‌های بزرگ چالش‌هایی را برای ارائه ارزیابی‌های ژنتیکی ایجاد می‌کنند که باید به نحوی برطرف شوند. به گونه‌ای که افزایش دقت در ارزیابی ژنومی و تسریع در پیشرفت ژنتیکی را به دنبال داشته باشند (Cole *et al.*, 2012).

در جدیدترین بررسی‌ها GWAS در تجزیه و تحلیل "trans-omic" برای بازسازی شبکه‌های بیوشیمیایی جهانی در لایه‌های مختلف omic با استفاده از اندازه‌گیری‌های چند omic و ادغام داده‌های محاسباتی استفاده می‌شود (Zhang *et al.*, 2012). در این راستا فناوری‌هایی برای اتصال داده‌های چندتایی مبتنی بر دانش قبلی در مورد فعل و انفعالات بیوشیمیایی معرفی و یک شبکه ترانس اومیک بیوشیمیایی با مفاهیمی از ماهیت

ارزیابی‌های ژنتیکی به دلیل دقت و صحت بالا در ثبت اطلاعات را به دنبال خواهند داشت که در واحدهای دامداری کشورهای توسعه یافته به طور گسترده مورد استفاده قرار می‌گیرند. استفاده از این تکنولوژی‌های مدرن افزایش بهداشت، افزایش عملکرد، کاهش هزینه‌ها و ضایعات کمتر محیطی واحد دامپروری را به دنبال خواهند داشت. پیشرفت در زمینه هوش مصنوعی در واحدهای دامپروری تا حدی افزایش داشته است که اکنون می‌تواند مواردی مانند سن، کیفیت و ترکیبات شیر و یا وضعیت سلامتی حیوان را ردیابی کنند (Murase, 2000).

### داده‌های با حجم بالا و ارزیابی‌های ژنتیکی

مجموعه داده‌های جدید خروجی از واحدهای دامپروری امروزه بزرگ و بزرگ‌تر می‌شوند و حجم بالایی از اطلاعات خروجی را به خود اختصاص می‌دهند. بخش عظیمی از این اطلاعات به دنبال ظهور ژنتیک مولکولی و فناوری‌های آمیکس به دلیل دسترسی به آرایه‌های SNP با حجم بالا و فناوری‌های نوین توالی‌یابی با هزینه بسیار پایین شکل گرفتند. این در حالی است که روش‌های محاسباتی با کارایی بالا برای ذخیره‌سازی و تجزیه و تحلیل داده‌های کارآمد در دست تهیه است (Kumar *et al.*, 2012). به دنبال افزایش چشمگیر استفاده از نسل جدید توالی‌یابی‌ها و تعیین ژنوتایپ گونه‌های مختلف، نیاز به ذخیره‌سازی فایل‌های داده‌های ژنوتیپی است، اگر چه داده‌های مربوط به توالی‌یابی کامل ژنوم ارگانیزم‌ها به ظرفیت بیشتری برای ذخیره‌سازی نیاز دارند (Cole *et al.*, 2012).

بیشترین میزان دقت در انتخاب ژنومیک برای صفات با وراثت‌پذیری کم محقق شده است و علاقه به ویژگی‌های جدید بهداشتی و مدیریتی (در کل محیطی) افزایش یافته است. به گونه‌ای که جمع‌آوری اطلاعات فنوتیپی کافی برای انجام ارزیابی‌های دقیق ممکن است سال‌ها طول بکشد و در ادامه اثبات قابلیت اطمینان بالا برای گاوهای مسن‌تر برای برآورد اثرات نشانگر لازم می‌باشد (Cole *et al.*, 2012).

تاریخچه حیوانات به سوی اطلاعات ژنومی برای بهبود پیش‌بینی و برآورد فنوتیپ آینده حیوانات استفاده کرد. امروزه با پیشرفت فناوری‌ها و سیستم‌های مدیریتی مدرن در علوم دامی، مقادیر زیادی از داده‌ها در دسترس خواهند بود که به طور خاص برای تجزیه و تحلیل ژنتیکی طراحی نشده‌اند؛ اما با این وجود ممکن است حاوی اطلاعات مفیدی باشند.

یادگیری ماشینی (ML: Machine Learning) طیف گسترده‌ای از تکنیک‌ها را برای شناسایی الگوها یا ویژگی‌ها در مجموعه داده‌های بزرگ ارائه می‌دهد. از یادگیری ماشینی امروزه در جامعه و صنعت استفاده گسترده‌ای می‌شود. در اصلاح‌نژاد دام و طیور از الگوریتم‌های یادگیری ماشینی برای ارزیابی‌های ژنومی استفاده می‌گردد. در عصر جدید در صنعت دامپروری از مجموعه داده‌های بزرگ و تکنیک‌های آماری استفاده می‌شود که در محدوده یادگیری ماشینی قرار می‌گیرند. مجموعه داده‌های ساده مانند داده‌های فنوتیپی و شجره‌ای و امروزه نشانگرها به سادگی در بانک‌های اطلاعاتی مناسب قرار می‌گیرند، در حالی که اطلاعات بیولوژیکی مدرن عمدتاً بدون ساختار می‌باشند. به عنوان مثال داده‌های ژنومی به صورت مجموعه‌ای از داده‌های متفاوت که غالباً حجم بالایی را به خود اختصاص می‌دهند؛ شامل اطلاعاتی همچون حاشیه‌نویسی SNPها و ژن‌ها، مسیرهای متابولیکی، مجموعه داده‌های بیان ژن، متیلاسیون و پروتئین‌ها هستند. روش‌های اصلاح‌نژادی سنتی برای استفاده کامل از این نوع اطلاعات ناتوان هستند، در حالی که ابزارهای ML مانند شبکه عصبی برای این کار مناسب هستند و اکنون تحقیقات زیادی در این زمینه در حال انجام می‌باشد. از ابزار ML می‌توان به سادگی برای اختصاص اولویت‌ها به هر SNP استفاده کرد و سپس پیش‌بینی‌های ژنومی را انجام داد. محدودیت این روش ترجیحاً این است که برخی از داده‌های ژنومی یا متابولیکی برای جمع‌آوری پرهزینه خواهند بود و تنها در تعداد کمی از افراد اندازه‌گیری می‌شوند (Pérez-Enciso, 2017).

ایستا و پویا بازسازی می‌گردد. یک مطالعه ارتباطی trans-Ome-Wide (trans-OWAS) اتصال فنوتیپ‌ها با شبکه‌های ترانس اومیک نشان‌دهنده این است که هم عوامل ژنتیکی و هم محیطی را منعکس می‌کنند. همچنین روش‌های گسترده‌ای بر اساس GWAS مانند مطالعات گستره کل فنوم (PheWAS) و مطالعات گستره کل مسیره‌ها (PWAS) و ... در حال گسترش است. امروزه از GWAS در مطالعات بسیاری همانند تجزیه و تحلیل جایگاه‌های صفات کمی به صورت مشترک استفاده می‌گردد (Zhang et al., 2012).

از مزایای مختلف این روش‌ها می‌توان مواردی چون گردش کار ساده، طراحی ساده طرح‌های تحقیقاتی و دستور کار مرتبط با آن، سرعت بالای خروجی داده‌ها با سهولت، گزینه‌ها و پلتفرم‌های مختلف در دسترس، برنامه‌های بیوانفورماتیک متعدد موجود و غیره را ذکر نمود. همچنین به عنوان یک روش تازه معرفی شده با معایب چندی از قبیل فایل‌های خروجی با حجم بسیار بالا، نیاز به حافظه قوی و بالا برای ذخیره فایل‌ها و داده‌ها، اطلاعات و داده‌های حجیم با نیاز به تجهیزات و ابزار قوی و با قدرت بالا همچون ابر رایانه‌ها، واحدهای محاسبه‌گر، اتصال به اینترنت با سرعت بسیار بالا برای تجزیه و تحلیل داده‌ها، در دسترس نبودن یک سیستم پردازش و دستورکار معمول و مشترک و گوناگونی در پردازش داده‌ها وابسته به ابزار محاسبه‌ای، پلتفرم و موجود زنده مورد بررسی نیز مواجه است که انتظار می‌رود به زودی راه‌حلی برای آن‌ها ارائه شوند (Chu and Corey, 2012).

### آبر داده‌ها و جدیدترین تکنولوژی‌ها

حجم داده‌هایی که در دوره پرورش دام و طیور جمع‌آوری می‌شوند، بسیار سریع در حال گسترش هستند، این در حالی است که فناوری تجزیه و تحلیل این حجم عظیم از داده‌ها عقب افتاده است. برای حل این موضوع پیشنهاد این است که از کاربرد روش‌های پیشرفته یادگیری ماشینی برای بررسی بینش تغییرات در الگوهای فنوتیپی با گذشت زمان و تغییر استفاده از داده‌های مربوط به

که با افزایش حجم داده‌های ژنومی و فنوتیپی تأثیر قابل توجهی در اصلاح‌نژاد دام و طیور خواهند داشت. از جمله محدودیت‌ها، دشواری در تفسیر الگوریتم‌ها از نظر ژنتیکی و حجم وسیعی از داده‌های مورد نیاز برای یادگیری بیشتر است. توسعه تکنیک‌ها و رویکردهای پیشرفته در عصر جدید فرصت‌های بسیاری را برای محققان و پژوهشگران فراهم کرده است تا روش‌های تحلیلی جدید را در زمینه‌های مختلف علوم به ویژه علوم دامی ایجاد کنند. این رویکردها رایانه‌های پیشرفته را قادر می‌سازند تا به انسان‌ها در تجزیه و تحلیل دانش مجموعه‌های بزرگ و پیچیده مانند داده‌های ژنومی کمک کنند.

### منابع

- Britt, J. H., Cushman, R. A., Dechow, C. D., Dobson, H., Humblot, P., Hutjens, M. F., Jones, G. A., Ruegg, P. S., Sheldon, I. M. and Stevenson, J. S. (2018). Invited review: Learning from the future—A vision for dairy farms and cows in 2067. *Journal of Dairy Science*, 101(5), 3722-3741.
- Chu, Y., and Corey, D. R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Therapeutics*, 22(4), 271-274.
- CM Dekkers, J. (2012). Application of genomics tools to animal breeding. *Current Genomics*, 13(3), 207-212.
- Cole, J. B., Newman, S., Foertter, F., Aguilar, I., and Coffey, M. (2012). Breeding and genetics symposium: Really big data: Processing and analysis of very large data sets. *Journal of Animal Science*, 90(3), 723-733.
- Kumar, S., Banks, T. W., and Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics*, 2012, 1-15.
- Murase, H. (2000). Artificial intelligence in agriculture. *Computers and Electronics in Agriculture*, 29(1/2).
- Nayeri, S., Sargolzaei, M., and Tulpan, D. (2019). A review of traditional and machine learning methods applied to animal breeding. *Animal Health Research Reviews*, 20(1), 31-46.
- Perera, C., Liu, C. H., and Jayawardena, S. (2015). The emerging internet of things marketplace from an industrial perspective: A survey. *IEEE Transactions on Emerging Topics in Computing*, 3(4), 585-598.
- Pérez-Enciso, M. (2017). Animal breeding learning from machine learning. *Journal of Animal Breeding and Genetics*, 134(2), 85-86.
- YANG, Y. L., Rong, Z., and Kui, L. (2017). Future livestock breeding: Precision breeding based on multi-omics information and population personalization. *Journal of Integrative Agriculture*, 16(12), 2784-2791.
- Zhang, H., Wang, Z., Wang, S., and Li, H. (2012). Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology*, 3(1), 26.

یادگیری ماشینی با هدف یافتن کارآمدترین الگوریتم یا روش انجام می‌شود و به مفهوم "مدل" اهمیت می‌یابد. محدودیت جدی این است که ML به طور معمول برای تفسیر از نظر اثرات مستقیم SNP یا اجزای واریانس دشوار است. ML می‌تواند برای توسعه پیش‌بینی/برآورد شایستگی‌های ژنتیکی، کشف روابط ناشناخته بین فنوتیپ‌ها و یا حتی بین فنوتیپ‌ها و متغیرهای محیطی مورد استفاده قرار گیرد (Pérez-Enciso, 2017).

ادغام داده‌های بزرگ ژنومی مانند نشانگرهای SNP در اصلاح‌نژاد دام معمولاً با محدودیت‌هایی روبرو بوده است. این امر نیاز به استفاده از مدل‌های غیرپارامتری از یادگیری ماشینی دارد. در دسترس بودن اطلاعات در حال افزایش است و افزایش همزمان در قدرت پردازش محاسباتی منجر به توسعه رویکردهای پیشرفته‌تر یادگیری ماشینی و کشف مجدد انواع خاصی از ابزارها به ویژه شبکه عصبی مصنوعی (ANN: Artificial Neural Network) می‌شود. روش‌هایی که یک کلاس ویژه به نام "یادگیری عمیق (DL: Deep Learning)" را تشکیل می‌دهند. تکنیک‌های یادگیری ماشینی مانند یادگیری عمیق می‌توانند تا حد زیادی به استخراج الگو و روابط کمک کنند، زمانی که مدل‌های سنتی نتوانند داده‌های بزرگ را با ساختارهای پیچیده مدیریت کنند (Nayeri et al., 2019). در ادامه روش‌های پیشرفته هوش مصنوعی (AI: Artificial Intelligence) می‌توانند با ارائه شواهد محکم منجر به تصمیم‌گیری آگاهانه با استفاده از جمع‌آوری داده‌های واقعی با دقت بالا با استفاده از انواع حسگرها روند مدیریت گله را به شدت تحت تأثیر قرار دهند.

### نتیجه‌گیری کلی

در این مطالعه به طور خلاصه به استراتژی‌های پیش‌بینی سنتی مانند BLUP، GBLUP، MAS و GWAS تا رویکردهای پیشرفته‌تر مانند یادگیری ماشینی، هوش مصنوعی، شبکه عصبی مصنوعی و یادگیری عمیق اشاره شد. این رویکردها مجموعه‌ای از فناوری‌های متنوع هستند





## Review Article

# Big data and the role of high-throughput technologies in livestock and poultry breeding

Farzad Ghafouri<sup>1\*</sup>, Hassan Mehrabani Yeganeh<sup>2</sup>, Sadegh Mohamadian Jeshvaghani<sup>3</sup>

<sup>1</sup> Ph.D. Student of Animal and Poultry Breeding & Genetics, Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

<sup>2</sup> Associate Professor of Animal Breeding and Genetics, Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

<sup>3</sup> M.Sc. of Animal Breeding and Genetics, Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

 <https://doi.org/10.22059/domesticj.2020.76952>

## Abstract

The general perspective of livestock and poultry breeding is being transferred to the digital era with high operational capacity, in which high-throughput technologies are utilized to boost the accuracy of phenotypic records collection and estimation of breeding values. Then, using advanced software and large computers, high amount of data is processed. The advent of NGS and the identification of SNPs along with new statistical methods for using this data to estimate the breeding value has led to the widespread use of genomic selection in dairy cattles and poultry. The development of data mining algorithms related to big data plays a significant role in estimating breeding values. A range of novel technologies, such as artificial intelligence, machine learning and deep learning, provide proper opportunities compared to traditional methods for examining economic traits with complex architecture. These approaches have made it possible to analyze large data sets and large genomic information in order to achieve desirable results. The purpose of this study is to provide a brief explanation of the new methods and novel technologies in animal sciences which are widely used in phenotype data collection and data registration in order to estimate accurate breeding values, in such a way as to lead to a digital future. Therefore, increasing the potential of big data analysis, along with new methods for recording phenotypic traits and estimating the breeding values, will dramatically augment genetic improvement.

**Keyword(s):** Big data, Breeding, Technology, Artificial Intelligence, Machine Learning, Deep Learning

\*Corresponding Author E-mail: Farzad.ghafouri@ut.ac.ir

Received: 07 June 2020

Revised: 11 June 2020

Accepted: 19 June 2020

Published online: 20 June 2020



**Citation:** Ghafouri, F., Mehrabani Yeganeh, H., Mohamadian Jeshvaghani, S. Big data and the role of high-throughput technologies in livestock and poultry breeding. *Professional Journal of Domestic*, 2020; 20(1): 34-40.