مجله علوم آماری، بهار و تابستان ۱۳۹۲
جلد ۷، شماره ۱، ص ۱۲۵-۱۴۳

توابع چندربعی معکوس به عنوان اثرات غیرخطی در مدل‌های رگرسیون لوژستیکی

آرزو مجیری[۱]، سروش علیمرادی[۱]، محمدرضا احمدزاده[۲]
[۱] دانشکده علوم ریاضی، دانشگاه صنعتی اصفهان
[۲] دانشکده برق و علوم کامپیوتر، دانشگاه صنعتی اصفهان

تاریخ دریافت: ۱۳۹۲/۳/۷ تاریخ آخرین بازنگری: ۱۳۹۲/۶/۱۳

چکیده : یک روش آماری رایج برای دسته‌بندی، استفاده از مدل‌های رگرسیون لوژستیک است. این روش با درنظرگرفتن اثرات خطی از ویژگی‌های افراد یا اشیا به مدل‌سازی احتمالات پسین عضویت در هر دسته می‌پردازد. در عمل این گمان وجود دارد که اثرات غیرخطی ویژگی‌ها می‌توانند نقش مؤثری در دسته‌بندی صحیح مشاهدات داشته باشند. اما مسئله‌ای که در پی ورود اثرات غیرخطی به مدل لوژستیک مطرح می‌شود، برآوردیابی پارامترها است. تحقیقات در سال‌های اخیر با فرض اثرات غیرخطی مانند اثرات متقابل و توابع پایه شعاعی گاوسی در مدل، برای پاسخ به مسئلهٔ برآوردیابی، استفادهٔ ترکیبی از ابزارهایی مانند شبکه‌های عصبی تکاملی و روش‌های برآوردیابی ماکسیمم درستنمایی را پیشنهاد کرده‌اند. در این مقاله نوعی از توابع پایه شعاعی با نام توابع چندربعی معکوس به عنوان اثرات غیرخطی در مدل لوژستیک در نظر گرفته می‌شود و با روش ترکیبی، پارامترهای مدل برآورد می‌شوند. آزمایشات تجربی برای مقایسهٔ مدل‌های پیشنهادی در این مقاله، با استفاده از داده‌های پزشکی و داده‌های واقعی مربوط به یک کارخانه تولید

آدرس الکترونیک مسئول مقاله: آرزو مجیری، arezu.mojiri@math.iut.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۲J۱۲

فولاد انجام گرفته است. نتایج نشان می‌دهد که حضور توابع چندربعی معکوس نسبت به توابع گاوسی در مدل، می‌تواند باعث افزایش دقت دسته‌بندی شود.

واژه‌های کلیدی : د سته‌بندی، رگرسیون لوژستیک، توابع چندربعی معکوس، شبکه‌های عصبی تکاملی.

۱ مقدمه

روش رایج آماری برای مدل‌سازی رابطهٔ بین یک یا چند متغیر توضیحی با یک متغیر پاسخ دوحالتی یا چندحالتی، استفاده از مدل رگرسیون لوژستیک است. این مدل‌ها در زمینه‌های بسیاری کاربرد دارند. یکی از این زمینه‌ها تقسیم‌بندی افراد یا اشیا به دسته‌های از پیش تعیین شده بر اساس ویژگی‌های مشاهده شده در آن‌ها است که از آن تحت عنوان دسته‌بندی[۱] یاد می‌شود. در یک مدل لوژستیک دستهٔ مربوط به هر شئ توسط متغیر پاسخ و ویژگی‌های مشاهده شده نیز در قالب متغیرهای توضیحی در مسأله بیان می‌شوند، به طوری که احتمال پسین عضویت یک شئ در هر دسته به وسیلهٔ یک ترکیب خطی از ویژگی‌ها مدل‌سازی و محاسبه می‌شود.

در مدل‌های لوژستیک با درنظر گرفتن اثرات خطی ویژگی‌های افراد سعی می‌شود احتمال عضویت افراد در هر دسته تعیین شود. اما گاهی در عمل ممکن است اثرات غیرخطی ویژگی‌ها (متغیرهای توضیحی) نیز در دسته‌بندی مؤثر باشند. برای حل این مسئله تلاش‌های بسیاری تا به امروز صورت گرفته است. به عنوان مثال، هستی و تیبشیرانی (۱۹۹۰) با در نظر گرفتن تبدیل‌های غیرخطی از متغیرهای توضیحی در مدل‌های لوژستیک، مدل‌های رگرسیون لوژستیک جمعی[۲] را معرفی کردند. مسئلهٔ دیگری که در پی این تلاش‌ها مورد توجه قرار گرفته این است که با وارد نمودن تبدیل‌های غیرخطی به مدل‌های لوژستیک، روش‌های برآوردیابی ماکسیمم درستنمایی در این مدل‌ها لزوما ماکسیمم سراسری را در تابع درستنمایی نتیجه نخواهد داد. از آنجا که تبدیل‌های غیرخطی باعث افزایش تعداد اکسترمم‌های

---

۱ Classification

۲ Additive logistic regression models

موضعی در تابع درستنمایی می‌شوند و از طرف دیگر تعداد این تبدیل‌ها در ابتدا نامعلوم است، لذا نحوهٔ تعیین تعداد و ساختار بهینهٔ تبدیل‌های غیرخطی، مسئله‌ای است که مورد توجه محققین قرار گرفته است. هرواس- مارتینز و مارتینز-استادیلو (۲۰۰۷) با در نظر گرفتن تبدیل‌های غیرخطی به صورت اثرات متقابل بین متغیرهای توضیحی در مدل‌های لوژستیک با پاسخ‌های دوحالتی، پیشنهاد کردند. در پی تلاش آن‌ها هرواس- مارتینز و همکاران (۲۰۰۸) این روش ترکیبی را برای مدل‌های لوژستیک با پاسخ‌های چندحالتی تعمیم دادند. پس از آن گوتییرز و همکاران (۲۰۱۱) نیز با در نظر گرفتن تبدیل‌های غیرخطی دیگری به صورت توابع پایه شعاعی گاوسی [۳] در مدل لوژستیک، روش برآوردیابی ترکیبی مشابه دیگری را مورد بررسی قرار دادند. روش‌های برآوردیابی ترکیبی پیشنهاد شده توسط هرواس-مارتینز و مارتینز-استادیلو (۲۰۰۷)، هرواس-مارتینز و همکاران (۲۰۰۸) و گوتییرز و همکاران (۲۰۱۱) به طور کلی از ترکیب سه نوع ابزار محاسباتی شبکه‌های عصبی مصنوعی [۴]، الگوریتم‌های تکاملی [۵] و روش‌های ماکسیمم درستنمایی استفاده می‌کند.

شبکه‌های عصبی مصنوعی از جمله روش‌های نوین و قدرتمند محاسباتی به شمار می‌روند که با ایده‌برداری از عملکرد نرون‌های موجود در سیستم اعصاب بدن انسان به‌وجود آمده‌اند (هیکین، ۱۹۹۹). شبکه‌های عصبی مصنوعی لایه‌هایی از نرون‌های مصنوعی را تشکیل می‌دهند که غالبا دارای سه نوع لایه‌ی ورودی، پنهان و خروجی هستند. در هر لایه، پردازش بر روی داده‌ها به کمک توابعی به نام توابع تحریک انجام می‌شود. این شبکه‌ها بر اساس نوع تابع تحریک [۶] مورد استفاده و همچنین بر اساس تعداد لایه‌های پنهان، می‌توانند نام‌گذاری شوند. (هیکین، ۱۹۹۹). شبکه‌های عصبی مورد استفاده در روش ترکیبی‌هرواس- مارتینز و مارتینز-استادیلو (۲۰۰۷) و هرواس- مارتینز و همکاران (۲۰۰۸) از نوع شبکه‌های عصبی واحد

---

۳ Gaussian radial basis functions

۴ Artificial neural networks

۵ Evolutionary algorithms

۶ Activation function

ضربی[۷] و شبکه‌های عصبی مورد استفاده در روش ترکیبی گوتییرز و همکاران (۲۰۱۱) از نوع شبکه‌های عصبی تابع پایه شعاعی[۸] هستند (انگلبریچت، ۲۰۰۷).

الگوریتم‌های تکاملی نیز از جمله روش‌های نوین محاسباتی در زمینهٔ مسائل بهینه‌سازی به شمار می روند که با ایده‌برداری از فرایند تکامل در طبیعت به وجود آمده‌اند. یک ویژگی بارز این الگوریتم‌ها نسبت به روش‌های بهینه‌ساز کلاسیک، جمعیت محور بودن آن‌ها است، که این الگوریتم‌ها را قادر می‌سازد به طور همزمان جمعیتی از جواب‌های ممکن را در مسئله مورد بررسی قرار دهند و باعث بالا بردن توانایی آن‌ها در حل مسائلی می‌شود که دارای فضای جواب گسترده و پیچیده‌ای هستند. این الگوریتم‌ها نیز دارای انواع متفاوتی هستند (انگلبریچت، ۲۰۰۷). الگوریتم تکاملی مورد استفاده در روش ترکیبی گوتییرز و همکاران (۲۰۱۱) تحت عنوان الگوریتم برنامه‌نویسی تکاملی[۹] است که از تحقیقات فوگل (۱۹۶۲) در زمینهٔ شبیه‌سازی تکامل سرچشمه می‌گیرد.

در این مقاله با در نظر گرفتن نوع دیگری از توابع پایه شعاعی به نام توابع چندربعی معکوس[۱۰] به عنوان تبدیل‌های غیرخطی از متغیرهای توضیحی، نوع دیگری از مدل‌های لوژستیک معرفی شده و با به کار بردن یک روش ترکیبی به برآوردیابی پارامترها در مدل پیشنهادی پرداخته می‌شود. در مرحلهٔ اول از روش ترکیبی، ساختار و تعداد بهینهٔ توابع چندربعی معکوس در مدل لوژستیک به کمک تکامل یک شبکهٔ عصبی تابع پایه شعاعی به‌دست می‌آید. در مرحلهٔ دوم، توابع چندربعی معکوس حاصل، به مجموعهٔ متغیرهای توضیحی اولیه افزوده می‌شود. در مرحلهٔ سوم با استفاده از دو الگوریتم مجزا ضرایب باقی‌مانده در مدل برآورد می شوند، که هریک از این دو الگوریتم منجر به مدل‌های متفاوتی می‌شوند. الگوریتم اول تحت عنوان الگوریتم MLogistic[۱۱] به دلیل وجود وابستگی میان متغیرهای توضیحی و تبدیل‌های غیرخطی از آن‌ها یعنی توابع چندربعی معکوس، به

---

۷ Product unit neural networks

۸ Radial basis functions neural networks

۹ Evolutionary programming algorithm

۱۰ Inverse multiquadratic functions

۱۱ MultiLogistic

منظور پیش‌گیری از بروز مسئلهٔ بیش‌برازش در مدل، برآوردهای ستیغی (کیسی و هویلینگن، ۱۹۹۲) را مورد محاسبه قرار می‌دهد. الگوریتم دوم نیز تحت عنوان الگوریتم SLogistic[۱۲] با استفاده از الگوریتم LogitBoost (فریدمن و همکاران، ۲۰۰۰) عمل می‌کند. این الگوریتم نیز با در نظر گرفتن یک پارامتر تنظیم به کنترل پیچیدگی مدل پرداخته و از بروز مسئلهٔ بیش‌برازش در مدل جلوگیری می‌کند.

به منظور نشان دادن عملکرد مناسب‌تر توابع پایه شعاعی چندربعی معکوس ابتدا براساس سه مجموعه داده به مقایسهٔ مدل‌های لوژستیک شامل توابع چندربعی معکوس با مدل‌های لوژستیک شامل توابع پایه شعاعی گاوسی پرداخته می‌شود، به طوری که به ازای هر مجموعه داده، بهترین مدل به گونه‌ای انتخاب می‌شود که دارای بیش‌ترین دقت در امر دسته‌بندی باشد. در نهایت پس از تعیین مدل‌های لوژستیک با بالاترین دقت در مجموعه داده‌های واقعی مربوط به فرایند تولید ورقه‌های فولاد، به بررسی ساختار آن مدل‌ها پرداخته می‌شود.

در بخش ۲ مدل‌های لوژستیک شامل توابع پایه شعاعی چندربعی معکوس معرفی می‌شود. در بخش ۳ با معرفی روش ترکیبی پیشنهاد شده توسط گوتییرز و همکاران (۲۰۱۱)، به برآوردیابی پارامترهای این مدل‌ها پرداخته می‌شود. در بخش ۴ با استفاده از سه مجموعه داده، دقت دسته‌بندی مدل‌های پیشنهادی با دقت دسته‌بندی مدل‌های لوژستیک شامل توابع گاوسی مورد مقایسه قرار می‌گیرد. در بخش ۵ نیز ساختار بهترین مدل‌های لوژستیک در مجموعه داده‌های واقعی تشریح می‌شود. در بخش ۶ مدت زمان مورد نیاز برای اجرای روش ترکیبی بیان می‌شود و در بخش پایانی، بحث و نتیجه‌گیری ارائه می‌شود.

۲ معرفی مدل LIIM

در مسئلهٔ دسته‌بندی مجموعهٔ مشاهدات $D = \{(x_n, y_n), n = ۱, \ldots, N\}$، بردار $x_n = (x_{n۱}, \ldots, x_{nk})$ نشان دهندهٔ مقادیر k ویژگی مشاهده شده مربوط به فرد یا شئ n ام و بردار $y_n = (y_n^{(۱)}, \ldots, y_n^{(J)})$ نشان دهندهٔ دسته‌ای است که فرد یا

[۱۲] SimpleLogistic

شئ n ام در بین J دسته، به آن تعلق دارد، به طوری که $y_n^{(\ell)} = ۱$ اگر فرد n ام به دسته‌ی ℓ تعلق داشته باشد و در غیر این صورت $y_n^{(\ell)} = ۰$. مدل‌های لوژستیک در مسائل دسته‌بندی، با فرض وجود J دسته و با در نظر گرفتن یک ترکیب خطی از ویژگی‌های مشاهده شده، رابطهٔ بین ویژگی‌ها و احتمال عضویت فرد با بردار ویژگی x به دسته‌ی ℓ ام را به‌صورت

$$\ln \frac{P(y^{(\ell)} = ۱|x)}{P(y^{(J)} = ۱|x)} = \alpha_۰^\ell + \alpha_۱^\ell x_۱ + \ldots + \alpha_k^\ell x_k \ , \ \ell = ۱, \ldots, J \ . \qquad (۱)$$

مدل‌سازی می‌کنند، که در این روش فرد با بردار ویژگی x به دستهٔ j اختصاص می‌یابد اگر (وب، ۲۰۰۲):

$$\ln P(y^{(j)} = ۱|x) = \max_\ell P(y^{(\ell)} = ۱|x) \ , \ \ell = ۱, \ldots, J \ .$$

گاهی در عمل ممکن است علاوه بر اثرات خطی، اثرات غیرخطی ویژگی‌ها نیز در مسئلهٔ دسته‌بندی و در نتیجه در مدل رگرسیون لوژستیک نقش مؤثری داشته باشند. گوتییرز و همکاران (۲۰۱۱) اثرات غیرخطی را در مدل‌های لوژستیک به‌صورت توابع پایه شعاعی گاوسی در نظر گرفتند. توابع پایه شعاعی، دارای مقادیری متقارن حول یک مرکز، با یک شعاع مشخص هستند. مقدار این توابع در هر نقطه تنها به فاصلهٔ آن تا مرکز بستگی دارد. یک نوع از این توابع، با عنوان توابع پایه شعاعی گاوسی به‌صورت

$$f(\|x - c\|) = \exp(-\frac{\|x - c\|^۲}{r^۲})$$

تعریف می‌شود و نوع دیگر آن‌ها تابع چندربعی معکوس به‌صورت

$$f(\|x - c\|) = \frac{۱}{\sqrt{r^۲ + \|x - c\|^۲}}$$

است، که در آن‌ها c مرکز و r شعاع تابع است و در هر دو تابع با افزایش فاصلهٔ x از مرکز c، کاهش می‌یابند. با وارد نمودن m تبدیل غیرخطی به‌صورت توابع چندربعی معکوس به رابطهٔ (۱) مدل لوژستیک پیشنهادی به‌صورت

$$\ln \frac{P(y^{(\ell)} = ۱|x)}{P(y^{(J)} = ۱|x)} = \alpha_۰^\ell + \sum_{i=۱}^{k} \alpha_i^\ell x_i + \sum_{j=۱}^{m} \frac{\beta_j^\ell}{\sqrt{r_j^۲ + \|x - c_j\|^۲}}, \ \ell = ۱, \ldots, J \quad (۲)$$

حاصل می‌شود که تحت عنوان مدل LIIM[۱۳] نام گذاری می‌شود.

دلیل استفاده از توابع گاوسی و چندربعی معکوس در مدل‌های لوژستیک به عملکرد این توابع در مسائل دسته‌بندی مربوط می‌شود. مقدار این توابع به ازای یک بردار ویژگی x و یک شعاع مشخص r، با افزایش فاصلهٔ x از مرکز ثابت c، کاهش می‌یابد. این کاهش می‌تواند در روش‌های دسته‌بندی به گونه‌ای استفاده شود که منجر به کاهش احتمال عضویت فرد در دسته‌ای شود که به آن تعلق ندارد که این امر، خود مشروط بر انتخاب مناسب مرکز و شعاع در تابع است. بنابراین توابع پایه شعاعی گاوسی و چندربعی معکوس می توانند در دسته‌بندی صحیح افراد و در نتیجه در مدل‌های لوژستیک نقش مؤثری داشته باشند.

۳ روش برآوردیابی ترکیبی در مدل LIIM

در مدل‌های لوژستیک معمولاً با روش‌های تکرار شونده‌ای[۱۴] مانند نیوتون–رافسون[۱۵]، به محاسبهٔ برآوردهای ماکسیمم درستنمایی پارامترها پرداخته می‌شود (اگرستی، ۲۰۰۲). در مدل (۲)، منفی لگاریتم تابع درستنمایی به‌صورت

$$\begin{aligned} L(\theta) &= -\sum_{n=1}^{N} \ln P(y_n = 1|x_n) \\ &= -\sum_{n=1}^{N}[\sum_{\ell=1}^{J-1} y_n^{(\ell)} f_\ell(x_n, \theta_\ell) - \ln \sum_{\ell=1}^{J-1} \exp\{f_\ell(x_n, \theta_\ell)\}] \end{aligned} \qquad (۳)$$

است، که در آن

$$f_\ell(x_n, \theta_\ell) = \alpha_0^\ell + \sum_{i=1}^{k} \alpha_i^\ell x_{in} + \sum_{j=1}^{m} \frac{\beta_j^\ell}{\sqrt{r_j^2 + \|x_n - c_j\|^2}}, \ \ell = 1, \ldots, J-1 \quad (۴)$$

و $\theta_\ell = (\alpha^\ell, \beta^\ell, W)$ ماتریس پارامترهای تابع f_ℓ است. $\alpha^\ell = (\alpha_0^\ell, \ldots, \alpha_k^\ell)$ و $\beta^\ell = (\beta_0^\ell, \ldots, \beta_m^\ell)$ ضرایب در مدل و $W = (w_1, \ldots, w_m)$ ماتریسی با درایه‌های $w_j = (c_j', r_j)$ است به‌طوری که $c_j = (c_{j1}, \ldots, c_{jk})$ مرکز و r_j شعاع در j امین

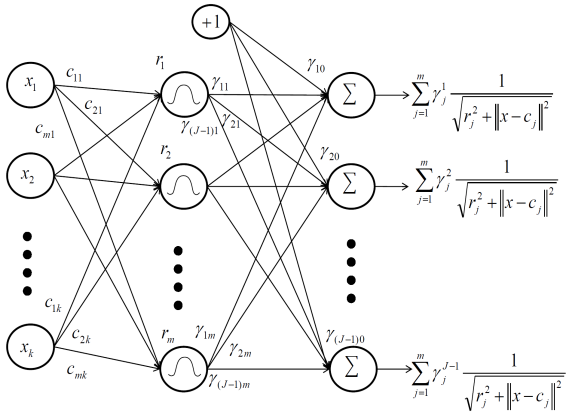
[۱۳] Logistic Initial Inverse Multiquadratic

[۱۴] Iterative methods

[۱۵] Newton-Raphson

تابع چندربعی معکوس هستند و برای هر i و j، $c_{ij} \in R$ و $r_j \in R^+$. وجود قسمت غیرخطی $\sum_{j=1}^{m} \beta_j^{\ell} \frac{1}{\sqrt{r_j^2 + \|x_n - c_j\|^2}}$ در تابع (۴) از یک طرف باعث افزایش تعداد اکسترمم‌های موضعی در تابع درستنمایی شده و از طرف دیگر نامعلوم بودن تعداد بهینهٔ توابع چندربعی معکوس باعث ناتوانی روش‌های معمول برآوردیابی ماکسیمم درستنمایی در محاسبهٔ مینیمم سراسری تابع (۳) شده و برآوردیابی پارامترها را با مشکل مواجه می‌سازد. این دلایل می‌توانند انگیزه‌های مناسبی برای استفاده از یک روش ترکیبی به منظور برآوردیابی پارامترها در چنین مدلی به شمار روند.

در روش ترکیبی برای برآوردیابی پارامترهای مدل LIIM، در مرحلهٔ اول تعداد و ساختار بهینهٔ توابع چندربعی معکوس، به کمک تکامل یک شبکهٔ عصبی تابع پایه شعاعی به‌دست می‌آیند. به این منظور ابتدا، قسمت غیرخطی مدل پیشنهادی (۲)، تعداد نرون‌های لایهٔ پنهان و c_jها وزن‌ها در اتصالات لایهٔ ورودی به لایهٔ پنهان در نظر گرفته می‌شوند. همچنین نرون‌ها در لایهٔ پنهان به کمک توابع تحریک چندربعی معکوس، وظیفهٔ پردازش داده‌ها را بر عهده دارند. در لایهٔ خروجی نیز نرون‌ها با دریافت خروجی‌های مربوط به لایهٔ پنهان، به عنوان ورودی‌های خود و همچنین با دریافت وزن‌هایی، حاصل‌جمع وزنی ورودی‌ها را به عنوان خروجی‌های شبکه محاسبه می‌کنند. شکل ۱ این شبکه را نمایش می‌دهد. هدف از نمایش قسمت



شکل ۱: شبکهٔ عصبی تابع پایه شعاعی

غیرخطی مدل (۲) توسط شبکه، تعیین تعداد و ساختار بهینهٔ تبدیل‌های غیرخطی به کمک تکامل شبکه است. به بیان دیگر تعیین وزن‌ها و همچنین تعداد نرون‌های لایهٔ پنهان یعنی m در این شبکه طی فرایند تکامل شبکه انجام می‌پذیرد. فرایند تکامل شبکه با استفاده از الگوریتم برنامه‌نویسی تکاملی انجام می‌شود. این الگوریتم مانند دیگر الگوریتم‌های تکاملی، جواب‌های ممکن در یک مسأله را به عنوان افراد در یک جمعیت در نظر می‌گیرد. سپس براساس یک جمعیت اولیه از این افراد و با به کار بردن عملگرهای تکاملی در هر مرحله، باعث بهبود افراد (جواب‌های) موجود در این جمعیت می‌شود. این مراحل در الگوریتم‌های تکاملی تا برقراری یک یا چند شرط خاتمه تکرار شده و در پایان، بهترین فرد در جمعیت نهایی به عنوان جواب بهینه در مسئله انتخاب می‌شود (انگلبریچت، ۲۰۰۷).

الگوریتم برنامه‌نویسی تکاملی مورد استفاده در روش ترکیبی با هدف تعیین بهترین شبکهٔ عصبی تابع پایه شعاعی، برمبنای یک جمعیت تصادفی اولیه از این نوع شبکه‌ها عمل می‌کند و با در نظر گرفتن منفی لگاریتم تابع درستنمایی به عنوان معیار خطا، سعی می‌کند در هر مرحله، ساختار و پارامترهای شبکه‌های عصبی در این جمعیت را در جهت مینیمم‌سازی این معیار، بهبود بخشد. در هر مرحله از این الگوریتم، عملگر جهش پارامتری برای تغییر و بهبود وزن‌ها و جهش ساختاری به منظور تغییر تعداد نرون‌های لایه‌ی پنهان (m)، استفاده می‌شوند و یک جمعیت جدید از شبکه‌های عصبی تابع پایه شعاعی را تولید می‌کنند. این مراحل تا رسیدن به یک یا چند شرط خاتمه تکرار می‌شود (گوتییرز و همکاران،۲۰۱۱). پس از برقراری شروط خاتمه، شبکهٔ عصبی تابع پایه شعاعی که در جمعیت نهایی، دارای کم‌ترین مقدار خطا باشد به عنوان بهترین شبکه انتخاب می‌شود. به این ترتیب تعداد و ساختار بهینهٔ توابع چندربعی معکوس در فرایند تکامل شبکهٔ عصبی تابع پایه شعاعی تعیین می‌شود. پس از تعیین پارامترهای مرکز و شعاع و تعداد توابع چندربعی معکوس، در مرحلهٔ دوم، توابع چندربعی معکوس به‌دست آمده، به‌صورت متغیرهای توضیحی جدید (z_jها) در مدل در نظر گرفته می‌شوند و مدل حاصل

به‌صورت

$$\ln \frac{P(y^{(\ell)} = ۱|x)}{P(y^{(J)} = ۱|x)} = \alpha_{\circ}^{\ell} + \sum_{i=۱}^{k} \alpha_i^{\ell} x_i + \sum_{j=۱}^{m} \beta_j^{\ell} z_j \ , \ \ell = ۱, \ldots, J \qquad (۵)$$

بازنویسی می‌شود. در مرحلهٔ نهایی، برای برآورد بقیه‌ی ضرایب، یعنی α_iها و β_jها، از دو الگوریتم مجزا استفاده می‌شود. در الگوریتم اول تحت عنوان الگوریتم MLogistic بر اساس روش پیشنهاد شده توسط کیسی و هویلینگن (۱۹۹۲)، با افزودن یک عبارت جریمه‌ای به رابطه‌ی (۳) به‌صورت $L^{\lambda}(\gamma) = L(\gamma) - \lambda\|\gamma\|^{۲}$ برآوردگرهای ستیغی[۱۶] ضرایب محاسبه می‌شوند، که در آن $\gamma' = (\alpha', \beta')$ و λ پارامتر ستیغی یا پارامتر تنظیم نامیده شده و باعث تنظیم پایداری برآوردهای حاصل می‌شود به گونه‌ای که به‌ازای $\lambda = ۰$ برآوردگرهای ضرایب همان برآوردگرهای ماکسیمم درستنمایی معمولی هستند در حالی که وقتی $\lambda \to \infty$ برآوردهای تمام ضرایب به سمت صفر میل می‌کنند.

الگوریتم دوم تحت عنوان SLogistic بر مبنای الگوریتم LogitBoost در برازش مدل‌های لوژستیک جمعی معرفی شده توسط فریدمن (۲۰۰۰) عمل می‌نماید. این الگوریتم مانند یک روش گام به گام پیش‌رونده[۱۷] به انتخاب بین متغیرهای توضیحی اولیه و توابع چندربعی معکوسی که بیش‌ترین ارتباط را با مقادیر پاسخ دارند، می‌پردازد. انتخاب در این الگوریتم بر مبنای معیار کم‌ترین توان‌های دوم خطا (SE) از ورود متغیرهای توضیحی که باعث ایجاد مسئلهٔ بیش‌برازش در مدل شده و ناپایداری را در بین برآوردها در پی دارند، جلوگیری به عمل می‌آورد:

$$SE^{\lambda} = SE - \lambda\|\gamma\|$$

جزئیات الگوریتم‌های MLogistic و SLogistic توسط گوتییرز و همکاران (۲۰۱۱) ارائه شده است. محاسبهٔ پارامتر تنظیم λ در هر دو الگوریتم با استفاده از مجموعهٔ مقادیر $\left\{ ۱۰^{۲}, ۱۰^{-۱.۵}, \ldots, ۱۰^{-۲} \right\}$ و روش اعتبار سنجی متقابل ۱۰-بخشی[۱۸]

[۱۶] Ridge estimators

[۱۷] Forward stepwise

[۱۸] 10-fold cross validation

انجام می‌پذیرد به طوری که هربار با انتخاب یک مقدار در این مجموعه و اجرای روش اعتبار سنجی به‌ازای آن مقدار، میانگین خطای دسته‌بندی محاسبه می‌شود. در نهایت مقداری از این مجموعه که به ازای آن کم‌ترین میانگین خطا به‌دست آمده، به‌عنوان مقدار مناسب برای پارامتر λ انتخاب می‌شود. در اجرای روش اعتبار سنجی متقابل ۱۰-بخشی ابتدا مجموعه داده به ۱۰ زیر مجموعهٔ مجزا و تقریبا برابر تقسیم می‌شود، سپس هر بار یک زیر مجموعه به‌عنوان مجموعهٔ آزمایشی برای محاسبهٔ خطا و مجموعهٔ دیگر به منظور مدل‌سازی و برآورد پارامتر در نظر گرفته می‌شوند. این فرایند تا جایی ادامه می‌یابد که هر بخش از داده‌ها یک بار به‌عنوان مجموعهٔ آزمایشی مورد استفاده قرار گرفته باشد.

پس از اجرای هریک از دو الگوریتم MLogistic و SLogistic مدل‌های لوژستیک متفاوتی حاصل می‌شوند به طوری که با در نظر گرفتن همهٔ متغیرهای توضیحی اولیه و توابع چندربعی معکوس به دست آمده در مرحله‌ی اول و اعمال دو الگوریتم MLogistic و SLogistic به‌صورت مجزا، مدل‌های حاصل به ترتیب MLIIM[۱۹] و SLIIM[۲۰] نام‌گذاری می‌شوند. ساختار کلی این دو مدل به‌صورت

$$\ln \frac{P(y^{(\ell)}=1|x)}{P(y^{(J)}=1|x)} = \alpha_0^{\ell} + \sum_{i=1}^{k} \alpha_i^{\ell} x_i + \sum_{j=1}^{m} \beta_j^{\ell} \frac{1}{\sqrt{r_j^2 + \|x-c_j\|^2}},\ \ell = 1,\ldots,J$$

است، با این تفاوت که ضرایب α_i و β_j در این دو مدل به‌ترتیب با الگوریتم‌های MLogistic و SLogistic برآورد می‌شوند.

با در نظر گرفتن تنها توابع چندربعی معکوس (z_jها) و اعمال نمودن دو الگوریتم MLogistic و SLogistic به‌صورت مجزا برای برآورد ضرایب، مدل‌های حاصل به‌ترتیب MLIM و SLIM نام‌گذاری می‌شوند و دارای ساختار کلی

$$\ln \frac{P(y^{(\ell)}=1|x)}{P(y^{(J)}=1|x)} = \alpha_0^{\ell} + \sum_{j=1}^{m} \beta_j^{\ell} \frac{1}{\sqrt{r_j^2 + \|x-c_j\|^2}},\ \ell = 1,\ldots,J$$

هستند، که تنها تفاوت آن‌ها نیز در برآورد ضرایب α_i و β_j است.

[۱۹] MultiLogistic Initial inverse multiquadratic

[۲۰] SimpleLogistic Initial Inverse Multiquadratic

۴ آزمایشات تجربی

در این بخش مدل‌های لوژستیک پیشنهادی به سه مجموعه داده برازش داده می‌شود. سپس به مقایسه و ارزیابی این مدل‌ها با مدل‌های لوژستیک شامل توابع پایه شعاعی گاوسی پرداخته می‌شود. در مجموعه دادهٔ اول، تحت عنوان مجموعه دادهٔ تشخیص بیماری قلبی[۲۱]، به بررسی عوامل مؤثر بر وجود یا عدم بیماری قلبی در افراد پرداخته می‌شود. این مجموعه شامل ۳۰۲ مشاهده است که متغیر پاسخ بیان‌کنندهٔ وجود یا عدم بیماری قلبی در افراد و متغیرهای توضیحی بیان‌کنندهٔ میزان کلسترول، قندخون و سایر نتایج آزمایشگاهی مربوط به هر فرد است. در مجموعه داده‌های دوم و سوم با استفاده از یک نمونه به حجم ۱۰۰۰ از تولیدات داخلی ورقه‌های فولادی در یک کارخانه به بررسی معایب در این ورقه‌ها پرداخته می‌شود. به این منظور خواص مکانیکی و آزمایشگاهی ورقه‌های تولیدی به‌عنوان متغیرهای توضیحی منظور می‌شوند. خواص مکانیکی ورقه‌ها مانند استحکام کشش نهایی، ضخامت آن‌ها و خواص آزمایشگاهی ورقه‌ها مانند میزان کربن، سیلیکون، منیزیم و مواد به کار رفته‌ی دیگر در تولید ورقه‌ها در نظر گرفته می‌شوند. مجموعه دادهٔ دوم تحت عنوان مجموعهٔ نورد، مهم‌ترین نوع عیب‌ها را در بین ورقه‌های معیوب و سالم مورد بررسی قرار می‌دهد، یعنی متغیر پاسخ چندحالتی است که نوع عیب در ورقه‌ها را بیان می‌کند. در مجموعه دادهٔ سوم تحت عنوان مجموعه نورد دودوئی، تنها معیوب و سالم بودن ورقه‌ها مورد توجه قرار می‌گیرند، یعنی متغیر پاسخ دوحالتی است.

ارزیابی مدل‌ها با روش اعتبارسنجی متقابل ۱۰-بخشی انجام می‌شود که به منظور افزایش دقت در نتایج، در هر بخش، طراحی و آزمایش مدل‌ها، ۳ بار تکرار می‌شود. نرخ دسته بندی صحیح[۲۲] در هر بار آزمایش به عنوان شاخص دقت در مدل اندازه‌گیری می‌شود. این شاخص برابر با نسبت مواردی است که توسط مدل،

---

[۲۱] http://www.uco.es/ayrna/index.php?lang=en

[۲۲] Correct classification rate

به‌درستی دسته‌بندی شده است و به‌صورت

$$CCR = \frac{1}{N}\sum_{i=1}^{N} I(C(x_n) = y_n)$$

محاسبه می‌شود، که در آن $C(x_n)$ نشان دهندهٔ دسته‌ای است که توسط مدل به فرد یا شئ n ام اختصاص می‌یابد و $I(\cdot)$ نشان‌دهندهٔ تابع نشان‌گری است که اگر $C(x_n) = y_n$ آن‌گاه مقدار این تابع برابر با یک و در غیر این صورت برابر با صفر است. پس از اجرای همهٔ ۳۰ آزمایش، میانگین و انحراف معیار نرخ های دسته‌بندی صحیح برحسب درصد محاسبه می‌شود.

جدول ۱ میانگین (انحراف معیار) شاخص دقت مورد نظر را برای مدل‌های لوژستیک پیشنهادی و مدل‌های لوژستیک شامل توابع گاوسی در ۳۰ آزمایش انجام شده را نمایش می‌دهد. مدل‌های MLIRBF، SLIRBF، MLRBF و SLRBF مدل‌های لوژستیکی هستند که تبدیل‌های غیرخطی در آن‌ها، از نوع توابع گاوسی است. برآوردیابی پارامترها در این مدل‌ها به‌ترتیب متناظر با مدل‌های پیشنهادی MLIIM، SLIIM، MLIM و SLIM است (گوتییرز و همکاران، ۲۰۱۱).

جدول ۱: میانگین و (انحراف معیار) شاخص‌های دقت

مجموعه داده	مدل							
	MLIM*	MLRBF	MLIIM*	MLIRBF	SLIM*	SLRBF	SLIIM*	SLIRBF
بیماری قلبی	۷۵/۸۰ (۵/۸۹)	۸۴/۱۶ (۶/۸۳)	۸۴/۷۶ (۷/۱۰)	۸۳/۷۵ (۶/۸۲)	۷۶/۵۷ (۵/۳۶)	۸۲/۸۷ (۶/۹۶)	۸۴/۳۰ (۶/۴۱)	۸۴/۱۹ (۶/۲۸)
نورد	۶۱/۹۳ (۲/۹۲)	۶۲/۱۰۳ (۲/۵۴)۱	۶۳/۳۳ (۲/۹۴)	۶۳/۹ (۲/۳۵)	۶۱/۹۳ (۲/۸۵)	۶۱/۸۷ (۲/۵۶)	۶۳/۲ (۲/۳۳)	۶۲/۸۷ (۲/۸۵)
نورد دودوئی	۷۸/۱۳ (۴/۷۰)	۷۵/۱۷۸ (۶/۷۲)	۷۸/۲ (۴/۴۴)	۷۵/۵۶ (۵/۸۳)	۷۸/۲۳ (۴/۵۱)	۷۵/۶۷ (۶/۶۶)	۷۸/۸۳ (۴/۴۰)	۷۶/۸۹ (۶/۰۶)

* مدل‌های پیشنهادی

در مقایسه‌ی میانگین و انحراف معیار شاخص‌های دقت مندرج در جدول ۱ برای مدل‌های لوژستیک پیشنهادی با مدل‌های لوژستیک دربردارندهٔ توابع گاوسی، نتایج زیر به تفکیک مجموعه‌های داده حاصل می‌شوند:

- در مجموعه دادهٔ بیماری قلبی، مدل‌های MLIIM و SLIIM (به ترتیب با مقادیر ۸۴/۷۶ و ۸۴/۳۰) دارای میانگین نرخ دسته‌بندی صحیح بالاتری نسبت

به مدل‌های لوژستیک شامل توابع گاوسی متناظر یعنی مدل‌های MLIRBF و SLIRBF (به ترتیب با مقادیر۸۳/۷۵ و ۸۴/۱۹) هستند.

- در مجموعه دادهٔ نورد، مدل‌های SLIM و SLIIM (به ترتیب با مقادیر۶۱/۹۳ و۶۳/۲) دارای میانگین نرخ دسته‌بندی صحیح بالاتری نسبت به مدل‌های لوژستیک شامل توابع گاوسی متناظر یعنی مدل‌های SLRBF و SLIRBF (به ترتیب با مقادیر ۶۱/۸۷ و ۶۲/۸۷) هستند.
- در مجموعه دادهٔ نورد دودوئی، مدل‌های MLIM، MLIIM، SLIIM و SLIIM (به ترتیب با مقادیر۷۸/۱۳، ۷۸/۲، ۷۸/۲۳ و ۷۸/۸۳) دارای میانگین نرخ دسته‌بندی صحیح بالاتری نسبت به مدل‌های لوژستیک شامل توابع گاوسی متناظر یعنی مدل‌های MLRBF، MLIRBF، SLRBF و SLIRBF (به ترتیب با مقادیر ۷۵/۷۸، ۷۵/۵۶، ۷۵/۶۷ و ۷۶/۸۹) هستند.

بار دیگر با مشاهده جدول ۱ می‌توان دریافت که در بین ۸ مدل برازش شده:

- در مجموعه دادهٔ بیماری قلبی، مدل MLIIM (با مقدار ۸۴/۷۶) دارای بالاترین شاخص دقت است. به بیان دیگر این مدل در ۸۴/۷۶ درصد موارد توانسته است وجود یا عدم بیماری قلبی را در افراد مورد آزمایش، به درستی تشخیص دهد.
- در مجموعه دادهٔ نورد، مدل MLIRBF (با مقدار ۶۳/۹) دارای بالاترین شاخص دقت است. به بیان دیگر این مدل در ۶۳/۹ درصد موارد توانسته است نوع عیب را در ورقه‌های فولادی، به درستی تشخیص دهد. نکتهٔ جالب توجه این که مدل پیشنهادی MLIIM نیز متناظر با اختلاف اندکی در رتبهٔ دوم قرار دارد.
- در مجموعه دادهٔ نورد دودوئی، مدل SLIIM (با مقدار ۷۸/۸۳) با اختلاف نسبتا زیادی دارای بالاترین شاخص دقت است. به بیان دیگر این مدل در ۷۸/۸۳ درصد موارد توانسته است معیوب یا سالم بودن ورقه‌های فولادی را به درستی تشخیص دهد.

۵ ساختار مدل‌های لوژستیک مربوط به داده‌های واقعی

همان‌طور که در بخش قبل گفته شد، مدل های MLIRBF و SLIIM به ترتیب در مجموعه داده‌های واقعی نورد و نورد دودوئی دارای بالاترین میانگین نرخ دسته‌بندی صحیح در بین ۸ مدل معرفی شده هستند. در این بخش به بیان ساختار هریک از این دو مدل پرداخته می‌شود . مدل‌های برازش‌یافته در این بخش با اجرای روش ترکیبی و با استفاده از کل داده‌ها در هر مجموعه داده، به‌دست آمده‌اند.

- مدل MLIRBF در مجموعه دادهٔ نورد:

هدف استفاده از ویژگی‌های مشاهده شده مربوط به ورقه‌های تولید شده مانند خواص آزمایشگاهی و مکانیکی آن‌ها به منظور تشخیص نوع عیب در این ورقه‌ها است. در این مجموعه داده، مهم‌ترین نوع معایب با کدهای $\ell = ۱,\ldots,۶$ و سالم بودن ورقه با کد $\ell = ۷$ نمایش داده شده است. پس از اجرای روش ترکیبی بر روی کل داده‌ها در این مجموعه و برازش مدل MLIRBF نرخ دسته‌بندی صحیح برابر با ۶۳/۹ درصد به‌دست می‌آید. ساختار کلی مدل برازش یافته به‌صورت

$$\ln \frac{P(y^{(\ell)} = ۱|x)}{P(y^{(۷)} = ۱|x)} = \alpha_۰^{\ell} + \sum_{i=۱}^{۲۴} \alpha_i^{\ell} x_i + \sum_{j=۱}^{۴} \beta_j^{\ell} z_j \ , \ \ell = ۱,\ldots,۶ \qquad (۶)$$

است، که در آن z_jها مقادیر مربوط به توابع گاوسی حاصل از روش ترکیبی هستند و به عنوان متغیرهای توضیحی جدید (z_jها) در نظر گرفته می‌شوند. برآورد ضرایب α_iهاو β_jها در جدول ۲ آورده شده است.

- مدل SLIIM در مجموعه دادهٔ نورد دودوئی:

هدف، استفاده از ویژگی‌های مربوط به ورقه‌های تولید شده مانند خواص آزمایشگاهی و مکانیکی آن‌ها به منظور تعیین معیوب یا سالم بودن ورقه‌ها است. در این مجموعه داده، ورقه‌های معیوب با کد صفر و ورقه‌های سالم با کد یک مشخص شده‌اند. پس از اجرای روش ترکیبی بر روی کل داده‌ها در این مجموعه، مدل SLIIM برازش یافته دارای نرخ دسته‌بندی صحیح برابر با

جدول ۲: برآورد ضرایب در مدل (۶)

متغیر توضیحی	ضریب	j ۱	۲	۳	۴	۵	٦
x_1 (ضخامت ورق)	α_1^j	۰/۱۴۸	−۰/۲۱۸	۰/۱۱	۰/۱۷	۰/۱۷۳	−۰/۰۷۹
x_2 (تیم کاری A)	α_2^j	−۰/۱۲۶	−۰/۰۹۱	−۰/۰۰۹	−۰/۰۶۹	−۰/۰۱۳	−۰/۰۰۸
x_3 (تیم کاری B)	α_3^j	۰/۰۶۳	۰/۰۴۴	۰/۰۳۳	۰/۰۴۴	۰/۰۰۸	−۰/۰۸۱
x_4 (تیم کاری C)	α_4^j	۰/۱۲۷	−۰/۰۳۹	۰/۰۳۹	۰/۰۳۲	−۰/۰۰۱	−۰/۰۵
x_5 (تیم کاری D)	α_5^j	−۰/۰۵۶	۰/۰۸۴	−۰/۰۶	−۰/۰۰۴	۰/۰۰۵	۰/۱۳۴
x_6 (شیفت کاری ۱)	α_6^j	−۰/۰۴۸	−۰/۰۰۱	۰/۰۳۹	۰/۰۱۸	−۰/۰۰۷	۰/۰۱۴
x_7 (شیفت اری ۲)	α_7^j	۰/۰۱۴	−۰/۰۱۷	−۰/۰۰۶	−۰/۰۲۱	−۰/۰۱۴	−۰/۰۱۱
x_8 (شیفت کاری ۳)	α_8^j	۰/۰۲۹	۰/۰۱۸	−۰/۰۲۹	۰/۰۰۵	۰/۰۲	−۰/۰۰۱
x_9 (روغن امولسیون)	α_9^j	۰/۰۲۱	۰/۱۱۶	۰/۰۳۶	۰/۰۹۷	−۰/۰۷۵	−۰/۱۱۴
x_{10} (ناهمواری سطح)	α_{10}^j	−۰/۲۱۶	۰/۱۱۵	۰/۰۹۸	۰/۱۱	۰/۱۰۴	−۰/۰۸۹
x_{11} (وزن کل ورق)	α_{11}^j	−۰/۲۶۴	۰/۰۵۳	۰/۹۹۸	۰/۰۱۱	۰/۰۴۵	۰/۲۸۶
x_{12} (نقطه‌ی تسلیم)	α_{12}^j	۰/۵۸۸	−۰/۰۲۲	−۰/۱۸۴	−۰/۰۰۱	−۰/۲۹۲	−۰/۰۶۱
x_{13} (استحکام کشش نهایی)	α_{13}^j	۰/۲۴۳	−۰/۰۶۳	−۰/۰۵۷	۰/۰۳۸	−۰/۰۴۷	−۰/۰۰۸
x_{14} (میزان کشش ورق)	α_{14}^j	۰/۲۸۹	۰/۳۵۷	−۰/۰۴۲	۰/۰۱۸	−۰/۰۷۷	−۰/۲۱۵
x_{15} (سختی مرکز ورق)	α_{15}^j	۰/۱	−۰/۰۱۳	۰/۰۹۱	۰/۲۶۳	−۰/۰۳۱	−۰/۰۵۷
x_{16} (میزان کربن)	α_{16}^j	−۰/۱۰۶	۰/۰۳۱	۰/۱۳	−۰/۲۵۸	−۰/۳۶۴	۰/۰۶۳
x_{17} (میزان سیلیکون)	α_{17}^j	−۰/۰۸	۰/۱۳۹	−۰/۰۴۱	−۰/۰۷	−۰/۰۹۳	۰/۰۱۴
x_{18} (میزان منیزیم)	α_{18}^j	۰/۲۶۱	−۰/۰۵	۰/۱۴۹	−۰/۱۶۹	−۰/۰۴۷	−۰/۲۹
x_{19} (میزان آلومینیوم)	α_{19}^j	−۰/۰۴	۰/۰۳۸	−۰/۰۹۲	−۰/۰۸۲	−۰/۱۶	−۰/۰۱۵
x_{20} (میزان نیتروژن)	α_{20}^j	۰/۰۷۲	۰/۰۷۶	۰/۰۱۴	−۰/۱۲۸	−۰/۰۰۳	−۰/۰۳۲
x_{21} (میزان وانادیوم)	α_{21}^j	−۰/۰۰۵	۰/۱۸۷	۰/۱۱۷	−۰/۰۹۷	−۰/۰۴۴	−۰/۰۶۳
x_{22} (میزان تیتانیوم)	α_{22}^j	−۰/۱۰۴	۰/۱۵۲	۰/۰۱	۰/۰۴۲	−۰/۰۳۷	۰/۰۶۹
x_{23} (میزان سولفور)	α_{23}^j	۰/۰۱۴	−۰/۰۵۸	۰/۲۱۲	−۰/۱۰۷	۰/۰۷۹	−۰/۰۸۷
x_{24} (میزان فسفر)	α_{24}^j	−۰/۲۱۶	۰/۲۲	۰/۰۲۷	−۰/۱۸۶	۰/۳۷۱	−۰/۰۲۷
z_1	β_1^j	۱/۶۴۷	−۱/۲۲۶	−۰/۷۴۳	−۰/۹۷۳	−۱/۰۸۸	۰/۱۸۸
z_2	β_2^j	۲/۴۳۰	−۱/۵۵۱	۰/۱۲۴	۰/۲۷۱	۰/۰۸۱	۰/۴۳۹
z_3	β_3^j	۰/۵۹۵	−۱/۱۲۱	−۱/۰۵۹	−۱/۳۸۳	−۰/۹۸۲	۱/۵۱۷
z_4	β_4^j	۳/۹۴۹	۰/۸۳۲	۲/۶۵۹	۳/۰۴۳	۳/۰۴۰	−۳/۲۸۸
عرض از مبدأ	α_0^j	۶/۱۷۷	۱/۸۶۷	−۰/۵۵۴	−۱/۵۲۲	−۱/۰۰۳	۱/۶۲۱

۷۹ درصد است. ساختار کلی مدل برازش یافته به‌صورت

$$\ln \frac{P(y^{(1)} = 1|x)}{P(y^{(0)} = 1|x)} = \alpha_0^1 + \sum_{i=1}^{24} \alpha_i^1 x_i + \sum_{j=1}^{6} \beta_j^1 z_j \ . \qquad (۷)$$

است، که در آن z_jها مقادیر مربوط به توابع چندربعی معکوس حاصل از روش ترکیبی هستند و به عنوان متغیرهای توضیحی جدید در نظر گرفته می‌شوند. برآورد ضرایب α_iها و β_jها در جدول ۳ آورده شده است.

جدول ۳: برآورد ضرایب در مدل (۷)

متغیر توضیحی	ضریب	$j = ۱$
$x_۱$ (ضخامت ورق)	$\alpha_۱^j$	۰/۱۲
$x_۲$ (تیم کاری A)	$\alpha_۲^j$	−۰/۱۰۱
$x_۳$ (تیم کاری B)	$\alpha_۳^j$	۰/۰۹۰
$x_۴$ (تیم کاری C)	$\alpha_۴^j$	۰/۲۰۲
$x_۵$ (تیم کاری D)	$\alpha_۵^j$	−۰/۱۷۷
$x_۶$ (شیفت کاری ۱)	$\alpha_۶^j$	۰/۰۳۵
$x_۷$ (شیفت اری ۲)	$\alpha_۷^j$	۰/۰۴۴
$x_۸$ (شیفت کاری ۳)	$\alpha_۸^j$	−۰/۱۲۵
$x_۹$ (روغن امولسیون)	$\alpha_۹^j$	۰/۰۰۵
$x_{۱۰}$ (ناهمواری سطح)	$\alpha_{۱۰}^j$	−۰/۱۹۲
$x_{۱۱}$ (وزن کل ورق)	$\alpha_{۱۱}^j$	−۰/۵۲۴
$x_{۱۲}$ (نقطه‌ی تسلیم)	$\alpha_{۱۲}^j$	۱/۸۷۲
$x_{۱۳}$ (استحکام کشش نهایی)	$\alpha_{۱۳}^j$	−۰/۳۷۶
$x_{۱۴}$ (میزان کشش ورق)	$\alpha_{۱۴}^j$	۱/۰۸۰
$x_{۱۵}$ (سختی مرکز ورق)	$\alpha_{۱۵}^j$	۰/۳۲۱
$x_{۱۶}$ (میزان کربن)	$\alpha_{۱۶}^j$	−۰/۱۸۵
$x_{۱۷}$ (میزان سیلیکون)	$\alpha_{۱۷}^j$	−۰/۰۳۳
$x_{۱۸}$ (میزان منیزیم)	$\alpha_{۱۸}^j$	۰/۵۵۹
$x_{۱۹}$ (میزان آلومینیوم)	$\alpha_{۱۹}^j$	۰/۰۳۱
$x_{۲۰}$ (میزان نیتروژن)	$\alpha_{۲۰}^j$	۰/۰۰۱
$x_{۲۱}$ (میزان وانادیوم)	$\alpha_{۲۱}^j$	۰/۰۰۳
$x_{۲۲}$ (میزان تیتانیوم)	$\alpha_{۲۲}^j$	−۰/۲۱۷
$x_{۲۳}$ (میزان سولفور)	$\alpha_{۲۳}^j$	۰/۱۲۱
$x_{۲۴}$ (میزان فسفر)	$\alpha_{۲۴}^j$	−۰/۶۹۷
$z_۱$	$\beta_۱^j$	−۹/۹۳۲
$z_۲$	$\beta_۲^j$	−۱۳/۸۱۳
$z_۳$	$\beta_۳^j$	۲۷/۱۳۹
$z_۴$	$\beta_۴^j$	−۲/۶۱۴
$z_۵$	$\beta_۵^j$	۲/۸۸۸
$z_۶$	$\beta_۶^j$	۲۳/۰۳۸
عرض از مبدأ	$\alpha_۰^j$	−۰/۵۴

۶ تخمین مدت زمان محاسباتی

جدول ۴ به طور تقریبی مدت زمان‌های مورد نیاز (برحسب ساعت) برای اجرای کامل آزمایشات مربوط به بخش ۴ را به تفکیک مدل‌ها نشان می‌دهد. روش ترکیبی با استفاده از دو نرم‌افزار MatlabR2012a و WEKA اجرا شده است به طوری که مرحلهٔ اول این روش، یعنی اجرای الگوریتم برنامه‌نویسی تکاملی در محیط برنامه‌نویسی Matlab و اجرای دو الگوریتم MLogistic و SLogistic در محیط WEKA انجام پذیرفته است. نتایج ارائه شده در جدول ۲ براساس سیستم 8 Core

3.00-GHz Intel Xeon(R) X5450 CPUs با حافظه ۱۲ گیگا بایت ۶۴ بیتی است.

جدول ۴: تخمین مدت زمان محاسبات

مجموعه داده	مدل							
	MLIM*	MLRBF	MLIIM*	MLIRBF	SLIM*	SLRBF	SLIIM*	SLIRBF
بیماری قلبی	۲۴	۱۶۸	۲۴	۱۶۸	۲۴	۱۶۹	۲۴	۱۶۹
نورد	۹۶	۲۴۱	۹۶	۲۴	۹۶	۲۴	۹۶	۲۴
نورد دودوئی	۴۸	۱۶۰	۴۸	۱۶۰	۴۸	۱۶۱	۴۸	۱۶۱

* مدل‌های پیشنهادی

بحث و نتیجه‌گیری

در این مقاله به معرفی یک روش دسته‌بندی با استفاده از مدل‌های لوژستیکی پرداخته شد که این مدل‌ها شامل تبدیل‌های غیرخطی از متغیرهای توضیحی، به‌صورت توابع چندربعی معکوس هستند. به منظور برآوردیابی پارامترهای مدل‌های پیشنهادی، روشی ترکیبی مورد استفاده قرار گرفت و مدل‌های برازش یافته با مدل‌های لوژستیکی که شامل توابع گاوسی هستند، بر اساس سه مجموعه داده، مورد ارزیابی و مقایسه قرار گرفتند. نتایج آزمایشات حاکی از مؤثر بودن نقش توابع چندربعی معکوس نسبت به توابع گاوسی، در بسیاری از موارد بوده است.

تقدیر و تشکر

نویسندگان از اعضای محترم هیئت تحریریه و داوران گرامی مجله تشکر و قدردانی می‌نمایند.

مراجع

Agresti, A. (2002), *Categorical Data Analysis*, 2nd Ed., John Wiley and Sons, New York.

Cessie, S. and Houwelingen, J. (1992), Ridge Estimators in Logistic Regression, *Applied Statistics*, **41**, 191-201.

Engelbrecht, A. P. (2007), *Computational Intelligence*, John Wiley and Sons, New York.

Friedman, J., Hastie, T. and Tibshirani, R. (2000), Additive Logistic Regression: A Statistical View of Boosting, *Annals of Statistics*, **38**, 337-374.

Fogel, L. J. (1962), Autonomous Automata, *Industrial Research*, **4**, 14-19.

Gutierrez, P. A., Hervas-Martinez, C. and Martinez-Estudillo, F. J. (2011), Logistic Regression by Means of Evolutionary Radial Basis Function Neural Networks, *IEEE Transactions on Neural Networks*, **22**, 246-263.

Hastie, T. J. and Tibshirani, R. J. (1990), *Generalized Additive Models*, Chapman & Hall.

Haykin, S. (1999), *Neural Networks: A Comprehensive Foundation*, 2nd Ed., Pearson Education.

Hervas-Martinez, C. and Martinez-Estudillo, F. J. (2007), Logistic Regression Using Covariates Obtained by Product Unit Neural Networks, *Pattern Recognition*, **40**, 52-64.

Hervas-Martinez, C., Martinez-Estudillo, F. J. and Carbonero-Ruz, M. (2008), MultiLogistic Regression by Means of Evolutionary Product Unit Neural Networks, *Neural Networks*, **21**, 951-961.

Webb, A. R. (2002), *Statistical Pattern Recognition*, John Wiley and Sons, New York.