

ارائه روشی برای بهینه‌سازی شبکه عصبی برای برآورد عیار با استفاده از اطلاعات سیستم مس پورفیری سوناجیل - اهر

پژمان طهماسبی^۱ و اردشیر هزارخانی^۱

^۱دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر، تهران، ایران.

تاریخ دریافت: ۱۳۸۸/۰۲/۳۰ تاریخ پذیرش: ۱۳۸۸/۱۱/۱۱

چکیده

در این پژوهش، برآورد الگوریتم‌های یادگیری مختلف در شبکه عصبی برای برآورد عیار در سیستم مس پورفیری سوناجیل مقایسه شده است. هدف این پژوهش، بهینه کردن ساختار شبکه مورد استفاده و ارائه روند بهینه‌سازی ساختاری آن برای برآورد عیار مس برای شناسایی بهتر منطقه است. بر این اساس، دوازده الگوریتم یادگیری پس انتشار خطا برای این هدف بررسی شدند. نتایج مطالعه بیانگر آن است که در الگوریتم‌های مورد استفاده دو الگوریتم LM و BFG بهترین کارایی را دارند. دلایل برای نشان دادن کارایی تقریباً مساوی الگوریتم‌های یادگیری دیگر به صورت کمی بیان شده است. متغیرهای ورودی شبکه، موقعیت طول و عرض جغرافیایی و خروجی آن، عیار کانسار در آن مختصات است. همچنین برای به دست آوردن ساختار بهینه شبکه مورد نظر از شبکه‌های با تعداد لایه‌های مختلف استفاده شد که در پایان شبکه با تعداد دوازده نرون مورد استفاده قرار گرفت. برای بررسی تأثیر شکل بهنجار کردن داده‌ها از شکل‌های مختلف داده‌ها استفاده شد که داده‌های بهنجار شده در بازه [0 1] نتایج بهتری داشتند. در پایان برای بهینه‌تر شدن شبکه همچنین از توابع مختلف انتقال در این شبکه استفاده شد که تابع انتقال تانژانت سیگموئیدی با کمترین خطای ممکن همراه بود و این تابع به عنوان تابع بهینه برگزیده شد. با در نظر گرفتن شرایط بهینه مقدار R^2 برای شبکه ۰/۹۴۶ به دست آمد که نویدگر استفاده از شبکه‌های عصبی با ساختار بهینه برای بهبود شرایط برآورد است.

کلیدواژه‌ها: شبکه عصبی، ساختار بهینه، برآورد عیار، الگوی یادگیری، سوناجیل

***نویسنده مسئول:** پژمان طهماسبی

E-mail: pejman@aut.ac.ir

۱- مقدمه

با توجه به قرار گرفتن نواحی اهر و هریس در بخش فلززایی طارم- قره‌داغ و وجود منابع غنی مس در این مناطق، این بخش پس از بررسی‌های مقدماتی زمین‌شناسی به عنوان منطقه، امیدبخش مورد مطالعه قرار گرفت (Iranian company of copper, 2003). در این منطقه توده گرانیتوئیدی نفوذی منبع احتمالی سیال‌های گرمایی برای کانی‌سازی و دگرسانی است (Hezarkhani, 2007). در این پژوهش سعی بر این است که با استفاده از روش‌های موجود، ساختار شبکه عصبی را بهینه و از آن برای برآورد عیار منطقه استفاده کرد. همچنین با استفاده از نتایج حاصل از این عملیات، الگوی جدیدی برای بهینه‌سازی شبکه عصبی و مراحل بهینه‌سازی ساختاری آن پیشنهاد شود. تاکنون در حوزه علوم زمین، شبکه‌های عصبی توانسته‌اند جایگاه ویژه‌ای برای خود بیابند. به گونه‌ای که بیشتر پژوهشگران سعی بر استفاده از آن به عنوان ابزاری قدرتمند دارند. در این زمینه می‌توان به کارهای انجام شده مانند بررسی‌های ژئوشیمیایی به همراه عیار و تناژ آنها با استفاده از شبکه‌های عصبی نوع احتمالی (Singer, 2006)، طبقه‌بندی مواد آلی رسوبی (Weller et al., 2005)، تعیین اندازه‌گیری‌های شوری به دست آمده حاصل از تصاویر ذرات آلی رسوبی با استفاده از شبکه‌های عصبی طبقه‌بندی (Weller et al., 2006; Andrew et al., 2007)، سنگ آتشفشانی و طبقه‌بندی آنها با استفاده از شبکه‌های عصبی مصنوعی (Lacassie et al., 2004 & Lacassie et al., 2006)، تعیین ویژگی‌های سفره آب زیرزمینی با استفاده از شبکه‌های عصبی (Rizzo & Dougherty, 1994)، تقریب برداشت‌های پیوسته و نقشه‌های آن (Funahashi, 1989)، تعیین ویژگی‌های مواد ناخالص در معدن سنگ آهک با استفاده از شبکه‌های عصبی پس‌خور (Koike & Matsuda, 2003)، استفاده از شبکه عصبی برای یافتن موقعیت نهشته‌ها (Singer & Kouda, 1996)، برآورد ذخیره (Wu & Zhou, 1993)، مقایسه نتایج حال از کریچینگ و شبکه عصبی (Samanta et al., 2005; Koike et al., 2001)، و کارهای پیش‌بینی در معدن (Yama & Lineberry, 1999) اشاره کرد. در این منطقه به علت وجود الگوهای ژئوشیمیایی پیچیده از شبکه‌های عصبی استفاده شد (که تفاوت این مطالعه با مطالعه‌های پیشین تلاش برای رسیدن به شبکه با ساختار بهینه به وسیله تمام متغیرهای موجود برای کنترل کارایی

شبکه است) تا با کاربرد آن بتوان ابهامات موجود در منطقه را برطرف کرد. در پایان با اطمینان از کارکرد شبکه می‌توان برای مناطقی که به اندازه کافی از آنها اطلاعات در دسترس نیست پیش‌بینی‌های اکتشافی نمود و ویژگی‌های منطقه را بهتر شناسایی کرد.

۲- زمین‌شناسی منطقه

منطقه مورد مطالعه در فاصله ۱۷ کیلومتری خاور هریس و ۴۵ کیلومتری جنوب خاور شهرستان اهر واقع در استان آذربایجان شرقی قرار گرفته است. این منطقه براساس تقسیمات زمین‌شناسی ایران در بخش البرز باختری- آذربایجان قرار گرفته و بخشی از نوار آتشفشانی البرز- آذربایجان به سن تشریری است. موقعیت سیستم پورفیری سوناجیل در شکل ۱ آورده شده است. واحدهای زمین‌شناسی موجود در این محدوده به شرح زیر هستند:

۲-۱. واحد آندزیتی ائوسن (E_{an})

این واحد آتشفشانی به طور گسترده در بخش‌هایی از جنوب و شمال باختر منطقه حضور دارد و به نظر می‌رسد که کهن‌ترین واحد سنگ‌شناختی در این ناحیه باشد. در بخش‌های جنوبی منطقه (شمال روستای آتمیان)، حالت لایه‌بندی حاصل از جریان‌های گدازه‌ای در این سنگ‌ها دیده می‌شود. دایک‌هایی با ترکیب دیوریتی-گرانودیوریتی و دیابازی به درون این توده نفوذ کرده‌اند. برونزدهایی از توف برشی شده و ایگنمریت با دانه‌بندی نامنظم دارای قطعات زاویه‌دار با زمینه آندزیتی در این واحد دیده می‌شود (Iranian company of copper, 2003).

۲-۲. واحد آندزیت پورفیری (E_{ap})

این واحد گسترده‌ترین واحد سنگ‌شناسی منطقه مورد مطالعه است و در بخش‌های گسترده‌ای از مرکز محدوده رخنمون دارد. این سنگ‌ها از دید ماکروسکوپی بافت ریزدانه تا پورفیری و اندیس رنگ ملاتوکرات تا مزوکرات دارند. شکل کانی‌سازی در این محدوده بیشتر رگه‌ای- رگچه‌ای و پراکنده و در گمانه‌های حفر شده در این واحد نیز شایع‌ترین دگرسانی، پتاسیک- فلیک بوده و با افزایش ژرفا مجموعه‌های پتاسیک چیره‌تر می‌شود. دایک‌ها و آپوفیزهای فراوانی از توده‌های نفوذی گرانیتوئیدی به درون این توده نفوذ کرده‌اند که می‌توان این عوامل را به عنوان منبع

الگوریتم بر پایه قانون تصحیح خطا بنا شده است که می‌توان آن را تعمیم الگوریتم مرسوم به کمترین میانگین مربعات دانست. یادگیری با این روش (پس انتشار) دو مرحله دارد: مرحله پیشروی و مرحله بازگشت. در مرحله پیشروی، ورودی‌ها به صورت لایه به لایه در شبکه پیش می‌روند و در پایان یک سری خروجی به عنوان جواب حقیقی شبکه به دست می‌آید، در این مرحله توزان اتصال ثابت است. در مرحله بازگشت، اوزان اتصال بر اساس قانون تصحیح خطا، تغییر می‌کند. تفاضل پاسخ حقیقی شبکه و پاسخ مورد انتظار که خطا نامیده می‌شود در سوی مخالف اتصالات در شبکه منتشر می‌شود و اوزان به گونه‌ای تغییر می‌یابند که پاسخ حقیقی شبکه به پاسخ مطلوب نزدیک تر شود.

مراحل الگوریتم پس انتشار را می‌توان به صورت زیر بیان کرد (Hornik et al., 1989):
 ۱- شبکه یک مثال آموزشی را دریافت می‌کند و با استفاده از اوزان موجود در شبکه که در ابتدا به صورت تصادفی مقداردهی می‌شوند، خروجی را محاسبه می‌کند.
 ۲- خطا (اختلاف میان نتیجه محاسبه شده (خروجی) و مقدار مورد انتظار) محاسبه می‌شود.

۳- خطا درون شبکه منتشر می‌شود و اوزان برای کمتر کردن خطا، از نو تنظیم می‌شوند. مهم‌ترین بخش، تنظیم اوزان است که پس از محاسبه خطای پیشگویی برای نمونه اول ورودی به سیستم، وزن‌ها از آخرین لایه به سوی نخستین لایه طوری تغییر می‌کنند که خطای پیشگویی کمتر می‌شود. در واقع BP (پس انتشار خطا) سرشکن کردن خطا بر روی سلول‌های (گره‌های) یک لایه و نیز لایه بعدی است. پس از این اطلاعات، نمونه دوم به شبکه داده می‌شود (Gallagher, 1999). از آنجا که با همان وزن‌ها نمونه جدید دوباره خطا خواهد داشت، روش پراکندگی وارون دوباره دست به کار می‌شود و وزن‌ها را به گونه‌ای تغییر می‌دهد که کمترین خطا را (هم برای این نمونه و هم برای نمونه پیشین) ایجاد کند. به این ترتیب پس از خواندن تعداد نمونه کافی به ورودی شبکه، اصطلاحاً شبکه Converge یا همگرا می‌شود (یعنی میزان خطا به کمترین مقدار خود می‌رسد). این به معنای موفقیت در مرحله یادگیری است و شبکه آماده است تا برای مرحله پیشگویی به کار رود.

توجه به این نکته اهمیت دارد که اگر تعداد نرون‌ها و لایه پنهان مورد استفاده، بیش از حد معمول باشد، سیستم به جای تجزیه و تحلیل داده‌ها، آنها را حفظ می‌کند و اصطلاحاً دچار بیش‌برازشی (Over Fitting) می‌شود. در این حالت مدل به دست آمده قادر خواهد بود که داده‌های مشابهی که در مرحله یادگیری مورد استفاده قرار گرفته است را به طور دقیق پیش‌بینی کند. اما اگر داده‌های جدیدی که در مرحله آموزش از آنها استفاده شده است به کار گرفته شوند، سیستم، عملکردی بسیار بدی خواهد داشت و خطای پیش‌بینی زیاد خواهد بود (Gallagher, 1999). برای جلوگیری از این پدیده، از روش اعتبارسنجی متقاطع استفاده می‌شود. در این روش، مجموعه داده‌های اولیه به سه دسته آموزش، آزمون و اعتبار تقسیم‌بندی می‌شوند. اعتبار شبکه همزمان با آموزش در هر دور سنجیده می‌شود و درست وقتی که خطای روی داده‌های اعتبار شروع به بالا رفتن می‌کند، آموزش شبکه قطع می‌شود.

۳-۳. تفاوت روش محاسباتی متداول با روش محاسباتی شبکه‌های عصبی

در روش‌های معمولی، گام‌های محاسباتی از پیش تعیین شده‌اند و توالی منطقی دارند، در مقابل، ANN‌ها نه توالی دارند و نه الزاماً از پیش تعیین شده هستند. در این حالت پردازشگرهای پیچیده مرکزی وجود ندارند، بلکه تعداد زیادی پردازشگر ساده وجود دارد که کاری جز گرفتن جمع وزنی ورودی‌هایشان از دیگر پردازشگرها ندارند. مدل‌سازی کلاسیک از نخستین گام، خطای بزرگی را مرتکب می‌شود که فقط در سیستم‌های ساده (خطی یا نزدیک به خطی) قابل چشم‌پوشی است. نخستین گام در روش کلاسیک برای بررسی داده‌ها، بررسی شاخص‌هایی مثل میانگین، انحراف معیار و ... است. از این مرحله به بعد در روش کلاسیک، کاری با تک‌تک نمونه‌ها نداریم و اهمیت فردی آنها از بین می‌رود. در واقع روش کلاسیک

احتمالی سیال‌های گرمایی برای کانی‌سازی و دگرسانی فرض کرد. حجم عظیمی از واریزه‌های توده بازالتی کواترن بر سطح این واحد لغزیده و مانع از پیدایش و بروز کانی‌سازی و دگرسانی در سطح شده است. حتی در مطالعات ژئوشیمیایی نیز، این واحد با یک افت عیار عناصر مشخص می‌شود (Iranian company of copper, 2003).

۲-۳. واحد گرانیتویدی الیگوسن (O_۲)

این واحد در گستره وسیعی از شمال منطقه مورد مطالعه و به صورت توده‌های کوچکی در جنوب و باختر محدوده رخنمون دارد. این توده، یک استوک کوارتز-مونزونیتی تا کوارتز مونزودپوریتی است که در راستای N-S و به طول تقریبی ۱۰ کیلومتر و عرض میانگین ۲ کیلومتر گسترش دارد و عرض رخنمون در بخش جنوب بیشتر و حدود ۴ کیلومتر است. این توده در بخش جنوبی و پهن‌تر خود، سبب تشکیل بخش‌های دگرسانی خیلی گسترده شده است، اما در بخش شمالی این حالت کمتر دیده می‌شود. به نظر می‌رسد که آپوفیرهایی از این توده در بخش‌های دگرسانی گرمایی (آندزیتی) تزریق شده‌اند که ابعاد آنها حداکثر ۱ کیلومتر مربع بوده و برخی از آنها نهشته‌های آتشفشانی مجموعه‌های پایینی اتوسن را قطع کرده‌اند.

به نظر می‌رسد دو پالس مجزای ماگمایی و به دنبال آن نفوذ مواد سیلیسی در این منطقه رخ داده است. در مرحله اول، واحد نفوذی با اندیس رنگ تیره‌تر، سپس واحد نفوذی روشن و در پایان رگه سیلیسی نفوذ کرده است. دایک‌های بسیار بزرگ و گسترده سیلیسی در درون توده نفوذی دیده می‌شود که در پیرامون این دایک‌ها مقدار ناچیزی کانی‌سازی مس به صورت ملاکیت نیز صورت گرفته است (Iranian company of copper, 2003).

۲-۴. واحد بازالتی کواترن (Q_۲)

این واحد شامل گدازه‌هایی با ترکیب آلکالی‌الیون‌بازالت تا تراکی‌بازالت بوده و آشکارترین نشانه‌های آتشفشانی منطقه را به نام کوه اکوزداغی با ارتفاعی حدود ۲۶۰۰ متر شکل داده است.

این واحد، جوان‌ترین واحد سنگی منطقه بوده و به صورت سقفی بر بالای واحدهای آتشفشانی اتوسن قرار گرفته است. واریزه‌هایی از این واحد به صورت توده‌های حجیم بر روی واحد آندزیت پورفیری لغزیده و موجب حذف بی‌هنجاری ژئوشیمیایی از این بخش‌ها شده‌اند. این روانه‌های بازالتی یک جبهه مشخص در بخش‌های جنوبی دارند قطعات بزرگ بازالتی حالت بالشی و گرد شده از خود نشان می‌دهند (Iranian company of copper, 2003).

۳- شبکه‌های عصبی (Neural Network)

شبکه‌های عصبی مصنوعی در واقع از ساختارهای بسیار پیچیده مغز انسان الهام گرفته شده است که در آن میلیون‌ها سلول عصبی با ارتباطی که با هم دارند به حل مسائل یا ذخیره‌سازی اطلاعات می‌پردازند. وظیفه شبکه عصبی یادگیری است. این فرایند ابتدا از راه آموزش یا همان مرحله کسب تجربه که به کمک یک سری داده‌های ورودی و خروجی مطلوب انجام می‌پذیرد، اجرا می‌شود. به این صورت که مجموعه‌ای از ورودی‌ها و خروجی‌های درست به شبکه داده می‌شود و شبکه عصبی با استفاده از این ورودی‌ها (مثال‌ها) مدل ریاضی پیچیده‌ای می‌سازد که در صورت دادن ورودی‌های جدید، پاسخ درستی را تولید کند (Krose & Smagt, 1996).

۳-۱. ساختار شبکه عصبی

همان‌گونه که یاد شد شبکه‌های عصبی مصنوعی از یک سری واحدهای ساختمانی اولیه تشکیل می‌شوند که با هم ترکیب شده و پس از انجام عملیات پردازش، یک خروجی به دست می‌دهند. این واحدهای اولیه مانند شکل ۲ به هم متصل هستند به گونه‌ای که خروجی هر واحد به عنوان ورودی واحدهای دیگر مورد استفاده قرار می‌گیرد.

۳-۲. معرفی روش پس انتشار (Backpropagation)

رایج‌ترین تکنیک آموزش نظارتی، الگوریتم پس انتشار خطا است. یادگیری این

۵-۲. انتخاب داده‌ها

انتخاب داده‌ها فرایندی است شامل گزینش داده‌ها برای ارائه به ANN و بیشتر شامل دیدگاهی قابل توجه است. کاملاً روشن است که محدودیت‌های سیستم اندازه‌گیری می‌تواند مجموعه ورودی‌های ممکن را برای یک کاربرد فرض شده، محدود کند. برخی از محدودیت‌های اصولی اثبات شده نیز می‌تواند بر انتخاب داده‌ها تأثیر بگذارد (به این معنی که مثلاً کمترین و بیشترین مقادیر ورودی و خروجی باید در مرحله یادگیری یا آموزش شبکه مورد استفاده قرار گیرد).

در مجموعه داده‌ها، داده‌هایی مفیدترند که اطلاعات بیشتری در مورد سیستم مورد نظر داشته باشند. از سویی معمولاً داده‌های حاوی اطلاعات در مورد یک سیستم این است که از تمامی نواحی سیستم مورد نظر، مثال وجود داشته باشد و تنوع در مورد مجموعه داده‌ها زیاد باشد. از سویی از آنجا که در شبکه عصبی مدل کاملاً به کیفیت و دقت داده‌های به کار رفته وابسته است، بنابراین باید بیشترین تلاش برای ارائه داده‌هایی با دقت بیشتر و عدم قطعیت کمتر به شبکه صورت گیرد.

مجموعه داده‌ها باید به اندازه کافی بزرگ باشد. عوامل مؤثر بر اندازه مورد نیاز سری داده‌ها در حل یک مسئله از راه روش شبکه‌های عصبی مصنوعی عبارتند از: ۱- نوع مسئله ۲- میزان پیچیدگی مسئله ۳- نوع روابط احتمالی و تقریبی میان متغیرهای مسئله ۴- تجربه کاربر ۵- دقت مورد نیاز در حل مسئله ۶- نوع شبکه عصبی مورد استفاده ۷- زمان حل و اعمال تجزیه بر نتایج مسئله.

۵-۳. تعیین نسبت تعداد داده‌ها در سری‌های داده‌ای متفاوت

در این راستا باور بیشتر متخصصان بر این است که ۷۰ درصد داده‌ها برای آموزش شبکه، ۱۵ درصد برای صحت و ۱۵ درصد دیگر برای آزمایش شبکه مورد استفاده قرار بگیرد (Gallagher, 1999). در مطالعه حاضر نیز روند تقسیم‌بندی داده‌ها به همین شیوه بوده است. در مطالعه حاضر تعداد داده‌های موجود برابر با ۱۷۰ تجزیه XRF بوده است که مورد استفاده قرار گرفتند.

انتخاب تصادفی داده‌های آموزش و بقیه باید به گونه‌ای صورت گیرد که در حالت آزمایش کارآیی شبکه، داده مورد آزمایش به جای عمل درونیابی در خارج از محدوده آموزشی، مورد برون‌یابی قرار نگیرد. به عبارت دیگر، بهتر است داده‌های آموزشی به گونه‌ای انتخاب شوند که تمام نقاط مرزی (و نیز نقاطی که در آن خاصیت مورد بررسی دچار تغییرات شدیدی می‌شود) را در بر بگیرد.

دو نکته بسیار مهم و قابل توجه نیز در مورد انتخاب داده‌های آزمایشی که باید رعایت شود این است که اولاً این بخش از داده‌ها نباید پیش‌تر در مرحله آموزش مدل پیشنهادی به کار رفته باشند و ثانیاً محدوده تغییرات این سری از داده‌ها باید در بازه آموزش شبکه قرار گیرد. به این معنی که تمام داده‌های مجموعه آزمایش باید میان مقادیر بیشینه و کمینه داده‌های آموزش قرار بگیرند.

با توجه به اهمیت شکل داده‌ها در آموزش شبکه عصبی در این مطالعه تأثیر انواع داده‌های ورودی بر روی شبکه، مورد بررسی قرار گرفت:

۱- داده‌های بهنجار شده در بازه [0 1] ۲- داده‌های خام ۳- داده‌های بهنجار شده در بازه [-1 1] با به کارگیری بیشینه و کمینه داده‌ها ۴- داده‌های بهنجار شده با استفاده از میانگین و انحراف معیار داده‌ها.

با توجه به شکل ۳ همان‌گونه که می‌بینید بهترین نتیجه از به کار بردن داده‌های بهنجار شده، در بازه [0 1] حاصل شده است.

در این مرحله برای بررسی تأثیر توابع مختلف آموزشی بر نتایج حاصل از شبکه عصبی، دو نوع تابع انتقال سیگموئیدی (تانژانتی و لگاریتمی) بررسی شده است که براساس مطالعات صورت گرفته این توابع تأثیر بسزایی در کارآیی شبکه دارند (Hartman et al., 1990; Cybenko, 1989). نتایج این بررسی راد نمودار شکل ۴ می‌بینید. مدل‌های شبکه عصبی و تعیین ساختار بهینه با در نظر گرفتن تغییر در نوع الگوریتم یادگیری و ساختار شبکه انجام گرفت. در این راستا، انواع الگوریتم‌های آموزشی

با عملی شبیه همگن کردن داده‌ها، پیچیدگی روابط آنها را محو می‌کند و به همین دلیل از کشف این پیچیدگی‌ها باز می‌ماند. به همین دلیل ترتیب سیستم کلاسیک در استخراج معنی از داده‌ها ضعیف و با بازده پایین عمل می‌کند و در بسیاری از موارد از کشف روابط میان داده‌ها ناکام می‌ماند. اگر می‌شد سیستمی باشد که با اهمیت دادن به فرد فرد داده‌ها آنها را تجزیه و تحلیل و نیز بدون پیش‌داوری در مورد شکل تابع هر متغیر (خطی یا غیر خطی بدون تابع) آن را ذخیره و ارزیابی کند، چنین سیستمی می‌توانست نتایج بیشتری را از ژرفای داده‌ها بیرون بکشد. شبکه‌های عصبی مصنوعی این قابلیت را دارند و به همین خاطر بسیار مورد توجه قرار گرفته‌اند (Hassanipak & Sharafeddin, 2005).

۴- مراحل عملی در ایجاد، توسعه و تست مدل شبکه عصبی

به طور کلی سه مرحله در این مطالعه انجام گرفته است. هر چند در اینجا کل مطالعه یاد شده به سه بخش تقسیم‌بندی می‌شود اما این سه بخش خود شامل مراحل و گام‌هایی جزئی هستند.

۴-۱. مرحله اول: ایجاد مدل خام

در آغاز هر مرحله اجرایی برای کشف اصلاحات مورد نیاز در مدل، پیش از هر کاری یک ساختار اولیه و خام برای شبکه در نظر گرفته و در هر مرحله تا تعیین مدل بهینه اصلاحاتی در مدل اولیه ایجاد می‌شود. تعیین این مدل بهینه پایانی به طور معمول عبارت است از تعیین نوع الگوریتم یادگیری، تعداد لایه‌ها، تعداد نرون‌ها در هر لایه، شکل مناسب داده‌های ورودی، نوع توابع تحریک و روش ارتباط میان نرون‌ها و متغیرهای آزاد مدل که هر کدام از این مسائل به نوبه خود بر روند کار تأثیر می‌گذارند.

۴-۲. مرحله دوم: اصلاح ساختار مدل اولیه و ایجاد مدل نهایی (آموزش شبکه)
فرایند آموزش شبکه را می‌توان به عنوان جستجویی برای تعیین کمینه کلی یک تابع هزینه (که به طور معمول مجموع مربعات خطا یا میانگین مربعات خطا است) در نظر گرفت. سطح خطا در فضای وزن ارتباطی موجود بوده و بعد این فضا برابر است با تعداد اوزان موجود در شبکه. بنابراین در طول فاز آموزش، جستجوی مجموعه مختصاتی از فضای اوزان که مربوط به کمینه مقدار عمومی سطح خطای حاصل از انطباق مدل با داده‌های واقعی، صورت می‌گیرد.

مراحلی که در این مطالعه برای یافتن یک مدل قابل قبول در شبکه‌ها در نظر گرفته شده‌اند شامل موارد زیر هستند:

۱- انتخاب یک ساختار و الگوریتم مناسب برای آموزش شبکه ۲- تعیین شکل بهینه داده‌های ارائه شده به شبکه ۳- انتخاب تابع محرک (انتقال) مناسب. ۴- معرفی مدل بهینه.

۴-۳. مرحله سوم: آزمایش و کنترل کارآیی و دقت کلی مدل شبکه عصبی

پیش از این مرحله متغیرهای مدل بهینه شبکه عصبی (که در مراحل مختلف یاد شده در مرحله آموزش تعیین شده‌اند) ذخیره می‌شوند و اکنون این شبکه، آموزش دیده نامیده می‌شود. در این مرحله ورودی‌هایی که برای آزمایش مدل مورد نظر هستند به شبکه وارد می‌شوند و خروجی‌های حاصل از مدل برای محاسبه مدل ایجاد شده مورد بررسی و تجزیه و تحلیل قرار می‌گیرند.

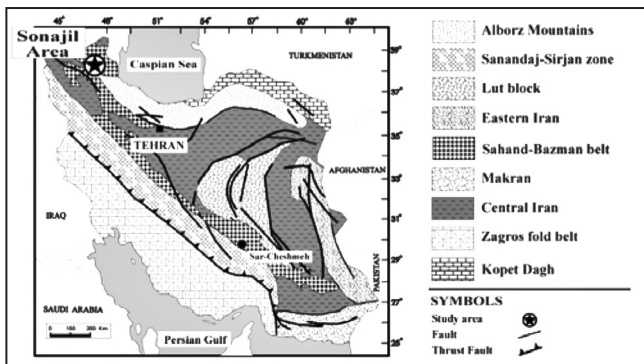
۵- بحث و بررسی نتایج

۵-۱. جمع‌آوری، انتخاب و تقسیم بانک داده‌ها به سری داده‌های مورد نیاز

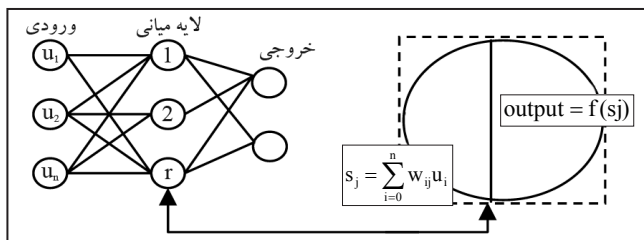
جمع‌آوری داده‌ها برای شبکه عصبی مصنوعی بی‌شک یکی از حساس‌ترین مراحل این کار بوده است. داده‌های مورد استفاده در این مطالعه به صورت حقیقی مورد استفاده قرار گرفتند.

کارآیی معتبر یک شبکه عصبی مصنوعی در جزء و هر مدل تجربی در کل، هنگامی صورت می‌گیرد که مجموعه داده‌های ارائه شده به مدل کاملاً نشان دهنده ویژگی‌های منطقه مورد مطالعه باشند.

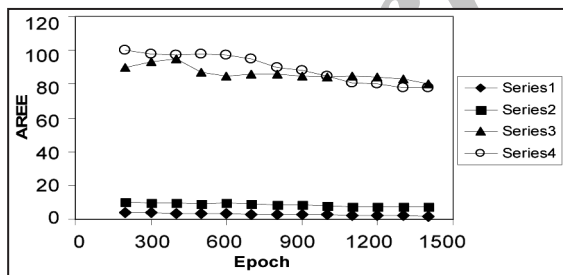
با تعداد نرون‌های مختلف وابسته است، به گونه‌ای که در این مطالعه پس از امتحان تعداد نرون‌های مختلف در لایه میانی، شبکه با دوازده نرون در لایه میانی بهترین نتیجه را در برداشت. همچنین با تأثیر دیده شده دیگر از عوامل بر روی کارکرد شبکه مانند شکل داده‌های ورودی، نوع تابع انتقال و دیگر متغیرهای بررسی شده برای دستیابی به یک برآورد مناسب، این متغیرها نیز باید در طراحی شبکه مورد توجه قرار گیرد و انواع مناسب آنها که منجر به ساختار بهینه شبکه می‌شوند، استفاده می‌شود.



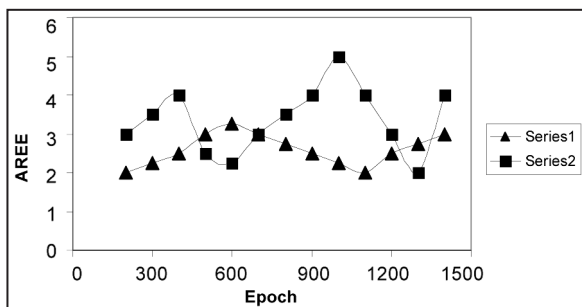
شکل ۱- نقشه زمین‌شناسی ایران و موقعیت نهشته سوناچیل (Hezarkhani & Ghayouri, 2007).



شکل ۲- ساختار شبکه عصبی.



شکل ۳- نتایج حاصل از تغییرات AREE با به کار بردن چهار شکل داده مورد استفاده.



شکل ۴- نمودار تغییرات AREE حاصل از برآورد عیار مس توسط یک شبکه عصبی نمونه با به کارگیری دو نوع تابع انتقال مورد بررسی 1- logsig 2- tansig.

زیر مجموعه الگوریتم BP به همراه شبکه‌های با تعداد نرون‌های مخفی متفاوت (جدول ۵) برای تعیین بهترین الگوریتم‌های آموزشی و ساختار بهینه مدل شبکه به کار رفته و عملکرد مدل‌های حاصل، بررسی و تجزیه شد. در پایان این مرحله، شبکه مورد نظر با بهترین کارایی و با تعداد لایه‌های مخفی بهینه مورد استفاده قرار گرفت.

۶- تعیین و ارائه ساختار پایانی و مقادیر بهینه متغیرهای شبکه عصبی مورد استفاده

با توجه به اعمال حالت‌های مختلف برای تعیین وضعیت بهینه، شبکه عصبی مورد نظر از نوع الگوریتم BR (Bayesian Regularization) مناسب است. در این پژوهش الگوریتم‌های مختلفی مورد بررسی قرار گرفت تا از آن، زیر شاخه الگوریتم BP که منجر به کمترین خطا در آموزش و آزمایش داده‌ها می‌شود به عنوان شبکه دلخواه، مورد توجه قرار بگیرد. ویژگی‌های کامل و ساختار شبکه بهینه پایانی شبکه عصبی مورد نظر در برآورد عیار مس به ترتیب در جدول ۴ و شکل ۵ ارائه شده است. با توجه به جدول ۲ و عملیات‌های انجام شده برای یافتن شبکه بهینه مورد استفاده، می‌توان ویژگی‌های آن را در جدول ۴ دید. برای بررسی کارایی شبکه ایجاد شده از نمودارهای زیادی استفاده شد که در زیر به شرح آنها پرداخته می‌شود.

۶-۱. نمودارهای اختلاف در تخمین

بررسی اختلاف‌های موجود میان مقادیر واقعی و مقادیر برآوردی می‌تواند درک مستقیمی را نسبت به میزان توانایی مدل‌های متفاوت در پیش‌بینی عیار مس ارائه دهد. در این نوع نمودارها محاسبه اختلاف‌های موجود از راه کاستن مقادیر برآوردی از مقادیر واقعی صورت گرفته است (شکل ۶).

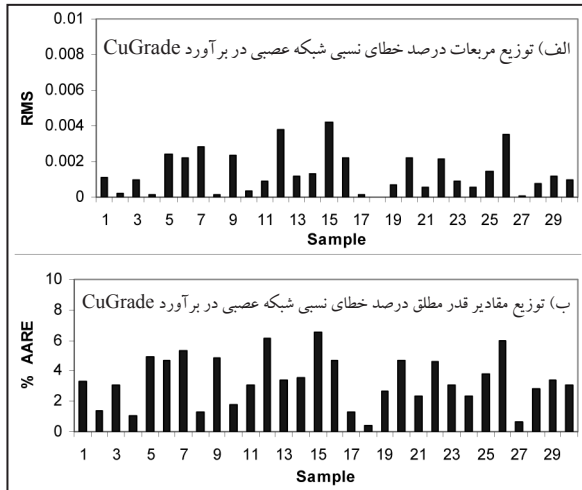
۶-۲. نمودار رسم متقابل یا پراکندگی

در این روش که با نام نمودارهای رسم متقابل یا متقاطع (CrossPlot) یا نمودارهای پراکندگی (Scatterplot) نامیده می‌شوند استفاده شده است که از رسم داده‌های برآوردی در برابر مقدار واقعی آنها ایجاد می‌شود (شکل ۷). در نهایت نمودارهای پراکندگی مربعات درصد خطای نسبی شبکه (RMS) و گسترش مقادیر قدر مطلق درصد خطای نسبی (AARE) رسم می‌شود (شکل ۸).

روش مورد استفاده در این مطالعه روشی مبتنی بر سعی و خطا است که تا حد زیادی به تجربه کاربر و پراکندگی داده‌ها و دیگر متغیرهایی که قبلاً بحث شد، دارد. به تازگی روش‌های دیگری به موازات شبکه عصبی مانند الگوریتم ژنتیک استفاده می‌شود که می‌تواند کنترل‌کننده روند انتخاب متغیرها باشد (Samanta et al., 2004). هدف این روش‌ها جلوگیری از قرار گرفتن شبکه در کمینه محلی و رسیدن به کمینه کلی است که به این دلیل چنین روش‌هایی را روش‌های کلی (Global) می‌نامند (Shang & Wah, 1996).

۷- نتیجه‌گیری

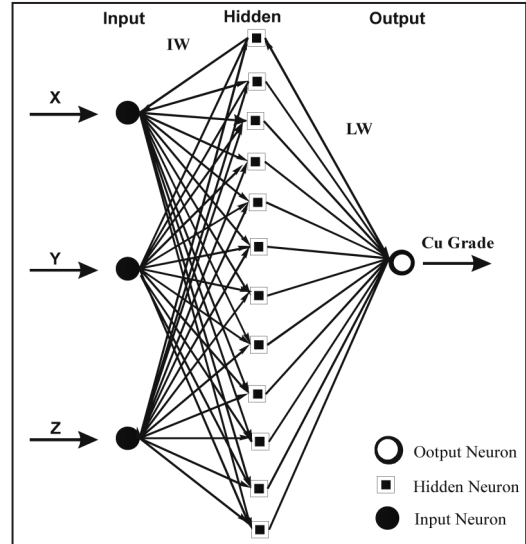
شبکه‌های عصبی مصنوعی در واقع سیستم‌های موازی هستند که می‌توان از آنها برای تشخیص الگوهای پیچیده در میان داده‌ها استفاده کرد. این ساختار موازی شبکه‌های عصبی آنها را قادر به انجام محاسبات پیچیده می‌کند. این روش نیازی به یک مدل پیچیده ریاضی و فرض خطی بودن تعداد زیادی از متغیرها ندارد. همان‌گونه که بیان شد شبکه‌های عصبی در تشخیص الگوها و نیز زمانی که اطلاعات در دسترس برای تفسیر کافی نیستند، می‌توانند ابزاری سودمند باشند. از آن جایی که تعداد نرون‌های لایه ورودی تابعی از عوامل مؤثر بر خروجی مورد نظر هستند، بنابراین کاربر باید در گزینش متغیرهای ورودی به گونه‌ای عمل کند که بیشتر عوامل مؤثر بر خروجی (مثلاً عیار) پوشش داده شوند. انتخاب تعداد نرون‌های لازم در لایه‌های ورودی کاملاً وابسته به خروجی شبکه است و تعداد بیشتر یا کمتر آنها در دقت شبکه چندان تأثیر ندارد. اما تعیین تعداد نرون‌های موجود در لایه میانی، کاملاً به تجربه و تکرار شبکه



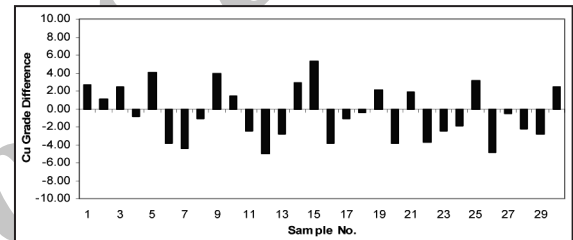
شکل ۸- نمودارهای الف) توزیع مربعات درصد خطای نسبی شبکه (RMS) و ب) توزیع مقادیر قدر مطلق درصد خطای نسبی (AARE).

جدول ۱- ویژگی‌های شبکه مورد استفاده برای برآورد عیار مس.

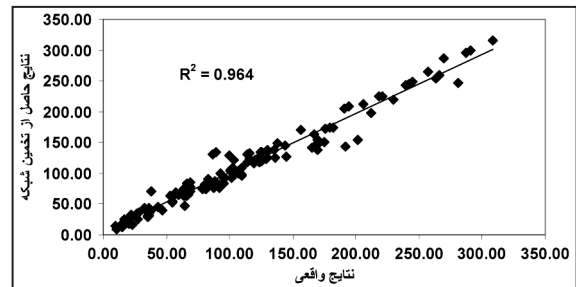
مقدار/ نوع	متغیرها
پیشخور	نوع شبکه
۳	تعداد ورودی
۱	تعداد خروجی
۱	تعداد لایه پنهان
۱۲	تعداد نرون لایه پنهان
خطی	تابع انتقال لایه ورودی
خطی	تابع انتقال لایه خروجی
تانژانت سیگموئیدی	تابع انتقال لایه پنهان
مجموع وزنی ورودی باضافه بایاس	ترکیب ورودی‌ها
گروهی	شکل ارائه داده‌ها
نرمال شده در بازه [0 1]	وضعیت عددی داده‌ها
MSE	تابع عملکرد شبکه
10^{-5}	مقدار قرار داده شده تابع عملکرد
پس انتشار خطا- LM	نوع الگوریتم آموزشی



شکل ۵- ساختار شبکه بهینه مورد استفاده در این مطالعه برای برآورد عیار مس.



شکل ۶- توزیع اختلاف مقادیر واقعی با مقادیر برآوردی مدل ANN در برآورد عیار مس (توجه شود که از میان تمام داده‌های موجود تعدادی به صورت تصادفی انتخاب و نمودار آنها رسم شده است).



شکل ۷- مقایسه نتایج حاصل از برآورد شبکه عصبی با نتایج واقعی.

جدول ۲- نتایج مربوط به آزمایش الگوهای آموزشی مختلف با تعداد نرون‌های متفاوت برای تشخیص الگوی آموزشی بهینه (N: تعداد نرون‌های لایه مخفی).

N=5				N=8			
Alg.	R _{ave}	R _{max}	Ep.	Alg.	R _{ave}	R _{max}	Ep.
LM	۰/۸۰۱	۰/۸۲۳	۲۰۰	SCG	۰/۸۲۵	۰/۸۴۵	۲۰۰
LM	۰/۷۹۸	۰/۸۲۸	۳۰۰	LM	۰/۸۳۶	۰/۸۵۲	۳۰۰
BFG	۰/۸۰۱	۰/۸۳۱	۴۰۰	LM	۰/۸۲۶	۰/۸۵۵	۴۰۰
LM	۰/۸۱۱	۰/۸۱۹	۵۰۰	LM	۰/۸۴۵	۰/۸۶۱	۵۰۰
LM	۰/۸۰۳	۰/۸۱۴	۶۰۰	BFG	۰/۸۱۲	۰/۸۴۵	۶۰۰
LM	۰/۸۱۶	۰/۸۳۵	۷۰۰	LM	۰/۸۵۳	۰/۸۷۶	۷۰۰
CGB	۰/۸۰۸	۰/۸۳۹	۸۰۰	LM	۰/۸۶۹	۰/۸۹۱	۸۰۰
LM	۰/۸۳۲	۰/۸۵۰	۹۰۰	LM	۰/۸۹۵	۰/۹۰۱	۹۰۰
LM	۰/۸۲۱	۰/۸۵۴	۱۰۰۰	CGB	۰/۸۹۸	۰/۹۰۹	۱۰۰۰
LM	۰/۸۵۵	۰/۸۶۷	۱۱۰۰	BFG	۰/۹۰۱	۰/۹۱۵	۱۱۰۰
LM	۰/۸۹۸	۰/۹۲۰	۱۲۰۰	CGP	۰/۸۷۸	۰/۸۹۸	۱۲۰۰
LM	۰/۸۳۶	۰/۸۶۹	۱۳۰۰	CGP	۰/۸۹۱	۰/۹۰۳	۱۳۰۰
GDX	۰/۸۱۲	۰/۸۴۹	۱۴۰۰	LM	۰/۸۸۹	۰/۹۰۸	۱۴۰۰

N=10				N=12			
Alg.	R _{ave}	R _{max}	Ep.	Alg.	R _{ave}	R _{max}	Ep.
BFG	۰/۸۶۳	۰/۸۹۹	۲۰۰	BFG	۰/۹۳۵	۰/۹۴۴	۲۰۰
BFG	۰/۸۸۴	۰/۹۰۲	۳۰۰	BFG	۰/۹۴۶	۰/۹۵۹	۳۰۰
OSS	۰/۹۰۱	۰/۹۰۸	۴۰۰	BFG	۰/۹۶۱	۰/۹۶۵	۴۰۰
BFG	۰/۹۰۴	۰/۹۱۵	۵۰۰	BR	۰/۹۶۵	۰/۹۷۳	۵۰۰
BFG	۰/۹۱۳	۰/۹۲۳	۶۰۰	BR	۰/۹۶۱	۰/۹۶۶	۶۰۰
BFG	۰/۹۳۵	۰/۹۴۵	۷۰۰	BR	۰/۹۵۸	۰/۹۶۴	۷۰۰
LM	۰/۹۲۵	۰/۹۴۱	۸۰۰	LM	۰/۹۶۵	۰/۹۷۰	۸۰۰
LM	۰/۹۳۸	۰/۹۴۴	۹۰۰	LM	۰/۹۷۲	۰/۹۷۸	۹۰۰
LM	۰/۹۴۲	۰/۹۴۸	۱۰۰۰	LM	۰/۹۵۸	۰/۹۶۵	۱۰۰۰
BR	۰/۹۲۲	۰/۹۳۶	۱۱۰۰	LM	۰/۹۵۱	۰/۹۶۶	۱۱۰۰
LM	۰/۹۲۸	۰/۹۴۰	۱۲۰۰	GDX	۰/۹۴۳	۰/۹۵۸	۱۲۰۰
LM	۰/۹۰۷	۰/۹۱۸	۱۳۰۰	OSS	۰/۹۴۱	۰/۹۵۱	۱۳۰۰
LM	۰/۹۱۴	۰/۹۲۳	۱۴۰۰	OSS	۰/۹۳۳	۰/۹۴۸	۱۴۰۰

References

- Andrew, F. W., Anthony, J. H. & AndrewWare, J., 2007- Two Supervised Neural Networks for Classification of Sedimentary Organic Matter Images from Palynological Preparations, *Math Geol*, 39: 657–671.
- Cybenko, G., 1989- Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, 2(4), 303.
- Exploration of Sonajil copper deposit, Iranian company of copper, northwestern report exploration, 1380-1382.
- Funahashi, K. I., 1989- On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2(3), 193.
- Gallagher, M. R., 1999- Multi-layer perceptron error surfaces: Visualization, structure and modeling: Unpublished PhD dissertation, University of Queensland, 225 p.
- Hartman, E. J., Keeler, J. D. & Kowalski, J. M., 1990- Layered neural networks with Gaussian hidden units as universal approximations. *Neural Computation*, 2(2), 210.
- Hassanipak, A. A. & Sharafeddin, M., 2005- Exploration data analysis, Tehran university press, p303.
- Hezarkhani, A. & Ghayouri, K., 2007- Ore Forming Fluid geochemistry Investigations on Sonajil Porphyry Stock, Azarbaijan-Iran, Based on Fluid Inclusion Microthermometry. *Amirkabir Journal of Science and Technology*.
- Hezarkhani, A., 2003- Exploration of Sonajil copper deposit, Iranian company of copper, northwestern report exploration.
- Hezarkhani, A., 2007- Hydrothermal Evolution in Sonajil Porphyry Copper System (East Azarbaijan Province, Iran): The History of an Uneconomic Deposit. *Journal of IGR, Stanford-USA*.
- Hornik, K., Stinchcombe, M. & White, H., 1989- Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359.
- Koike, K. & Matsuda, S., 2003- Characterizing content distributions of impurities in A limestone mine using a feed forward neural network: *Nat. Resour. Res.*, 12, 3: 209–223.
- Koike, K., Matsuda, S. & Gu, B., 2001- Evaluation of interpolation accuracy of neural kriging with application to temperature-distribution analysis: *Math. Geol.*, 33, 4: 421–448.
- Krose, B. & Smagt, P., 1996- An introduction to neural networks, eighth edition, November, Amsterdam.
- Lacassie, J. P., Roser, B. P., Ruiz-del-Solar, J. & Herv'e, F., 2004- Visualization of geochemical datasets by using neural networks: a novel perspective for sedimentary provenance analysis: *Sedimentary Geol.*, 165, 1: 175–191.
- Lacassie, J. P., Solar, J. R., Roser, B. & Herve, F., 2006- Visualization of Volcanic Rock Geochemical Data and Classification with Artificial Neural Networks, *Mathematical Geology*, 38, 6.
- Matias, J. M., Vaamonde, A., Taboada, J. & Gonz'alez-Manteiga, W., 2004- Comparison of Kriging and Neural Networks With Application to the Exploitation of a Slate Mine, *Mathematical Geology*, 36, 4.
- Rizzo, D. M. & Dougherty, D. E., 1994- Characterization of aquifer properties using artificial neural networks: *Neural kriging: Water Resour. Res.*, 30, 2: 483–497.
- Samanta, B., Bandopadhyay, S. & Ganguli, R., 2004- Data segmentation and genetic algorithms for sparse data division in Nome placer gold grade estimation using neural network and geostatistics: *Mining Exploration Geol.*, 11, 1–4: 69–76.
- Samanta, B., Ganguli, R. & Bandopadhyay, S., 2005- Comparing the predictive performance of neural networks with ordinary kriging in a bauxite deposit: *Transactions of Institute of Mining and Metallurgy*, 114:129–139
- Shang, Yi. & Wah, B. W., 1996- Global optimization for neural network training: *IEEE Comput.*, 29, 3:45–54.
- Singer, D. A. & Kouda, R., 1996- Application of a feed forward neural network in the search for Kuroko deposits in the Hokuroku district, Japan: *Math. Geol.*, 28, 8:1017–1023.
- Singer, D. A., 2006- Typing mineral deposits using their associated rocks and grades and tonnages in a probabilistic neural network. *Math Geol* 38(4):465–475.
- Weller, A. F., Corcoran, J., Harris, A. J. & Ware, J. A., 2005- The semi-automated classification of sedimentary organic matter in palynological preparations. *Comput Geosci* 31(10):1213–1223.
- Weller, A. F., Harris, A. J., Ware, J. A. & Jarvis, P. S., 2006- Determining the saliency of feature measurements obtained from images of sedimentary organic matter for use in its classification. *Comput Geosci* 32(9):1357–1367.
- Wu, X. & Zhou, Y., 1993- Reserve estimation using neural network techniques: *Comput. Geosci.*, 19, 4:567–575.
- Yama, B. R. & Lineberry, G. T., 1999- Artificial neural network application for a predictive task in Mining: *Mining Eng.*, 51, 2:59–64.