

## مقایسه روش‌های طبقه‌بندی، شبکه عصبی مصنوعی و رگرسیون چندمتغیره در برآورد بازایی فلز از بلوک کانسنگ

جواد غلام نژاد<sup>۱</sup>، رضا لطفیان<sup>۲</sup>، یوسف میرزائیان لرد کیوان<sup>۳</sup>

۱- دانشیار، دانشکده مهندسی معدن و متالورژی، دانشگاه یزد، یزد

۲- دانشجوی دکتری، دانشکده مهندسی معدن و متالورژی، دانشگاه یزد، یزد

۳- استادیار، دانشکده مهندسی معدن و متالورژی، دانشگاه یزد، یزد

(دریافت ۱۳۹۸/۰۳/۲۶، پذیرش ۱۳۹۸/۰۸/۲۰)

### چکیده

با توجه به نقش بازایی در محاسبه ارزش اقتصادی بلوک کانسنگ و تأثیر مقدار این ارزش بر محاسبات طراحی و برنامه‌ریزی تولید معدن، تعیین بازایی فلز از بلوک کانسنگ ارسالی به کارخانه فرآوری، از اهمیت بالایی برخوردار است. هدف از این پژوهش، بررسی قابلیت برآورد بازایی بلوک کانسنگ به صورت کیفی و با روش‌های مبتنی بر طبقه‌بندی داده‌ها از مجموعه روش‌های داده‌کاوی و به صورت کمی، با دو روش رگرسیون چندمتغیره و مدل هوشمند شبکه عصبی، بر اساس داده‌های آنالیز خوراک ورودی کارخانه است. برای نیل به این هدف، معدن مس میدوک مورد مطالعه قرار گرفت و با استفاده از ۵۸ نمونه آنالیزشده عیار خوراک کارخانه، شامل عیارهای Cu، CuO و CuS و میزان بازایی عنصر Cu در محصول نهایی، فرآیند پیش‌بینی بازایی کل ذخیره به صورت کیفی با روش‌های طبقه‌بندی درخت تصمیم، قانون بیز و الگوریتم نزدیک‌ترین همسایه انجام شد. برای برآورد کمی میزان بازایی ذخیره، مدل رگرسیون چندمتغیره و شبکه عصبی مصنوعی برای شاخص‌های عیاری مذکور و میزان بازایی بین ۴۷ نمونه از ۵۸ نمونه برقرار شد و توسط ۱۱ نمونه آنالیزشده آزمایشی، مدل‌های به‌دست‌آمده اعتبارسنجی شدند. معیارهای میانگین خطا و جذر میانگین مربعات خطا در مدل رگرسیونی به ترتیب ۰/۲۱۷۰۲ و ۰/۲۴۹۷۲ و در مدل شبکه عصبی مصنوعی به ترتیب ۰/۱۵۷۵۳ و ۰/۲۱۴۰۴ محاسبه شدند. بنابراین مدل شبکه عصبی مصنوعی به‌عنوان ابزار دقیق‌تری در پیش‌بینی بازایی نسبت به مدل رگرسیون چندمتغیره عمل می‌کند. نتایج آنالیز حساسیت این مدل نشان داد، عیار Cu مهم‌ترین عامل و عیار CuO و CuS نیز به ترتیب، دیگر عوامل تأثیرگذار بر تغییرات بازایی هستند.

### کلمات کلیدی

بازایی، طبقه‌بندی، رگرسیون چندمتغیره، شبکه عصبی مصنوعی.

## ۱- مقدمه

مدل بلوکی اقتصادی، اساس طراحی محدوده نهایی و برنامه‌ریزی تولید در معادن است. در این روش با در نظر گرفتن شاخص‌های اقتصادی مانند درآمدها و هزینه‌های استحصال فلز برای هر بلوک، شاخص‌های تکنولوژیکی هم‌چون راندمان‌های استخراج، فرآوری، ذوب و پالایش و تناژ بلوک در مدل بلوکی عیاری به دست می‌آید. از بین ویژگی‌های تکنولوژیکی، راندمان کارخانه فرآوری یکی از شاخص‌های مهم برای محاسبه ارزش اقتصادی بلوک‌ها است. این راندمان با توجه به تنظیمات خاص واحد فرآوری و مشخصات کانسنگ ارسالی به کارخانه تعیین می‌شود. در برخی موارد، مانند مورد بررسی شده در این پژوهش، میزان بازیابی واقعی کارخانه فرآوری با توجه به مشخصات مواد ورودی کارخانه به شدت تغییر می‌کند. در این صورت، میزان فلز استحصال شده مورد انتظار و به دنبال آن، سودهای مورد انتظار از فروش محصول، تغییر کرده و محاسبات ارزش اقتصادی بلوک و به دنبال آن محدوده نهایی و حتی فرآیند برنامه‌ریزی تولید دست‌خوش تغییر خواهد شد. با فرض شرایط یکسان کارخانه در فرآوری خوراک ورودی، یکی از عوامل مهمی که می‌تواند میزان بازیابی را تغییر دهد، نسبت‌های خاص و رابطه پنهان بین مجموعه عناصر موجود در خوراک ورودی است. حال، چنان‌چه بتوان با ارایه مدلی با توجه به آنالیز خوراک ورودی، این رابطه پنهان را مشخص و میزان بازیابی محصول را پیش‌بینی کرد، مقادیر واقع‌بینانه‌تری از ارزش اقتصادی بلوک کانسنگ ارسالی به کارخانه و به دنبال آن، محدوده نهایی به دست می‌آید و برنامه‌ریزی تولید مطمئن‌تری حاصل خواهد شد. در این پژوهش، مدل‌های پیش‌بینی برای میزان بازیابی یک ذخیره معدنی، به دو صورت کیفی و کمی ارایه می‌شود. برآورد کیفی بازیابی براساس مفهوم طبقه‌بندی<sup>۱</sup> از مجموعه الگوریتم‌های داده‌کاوی<sup>۲</sup> و با سه درخت تصمیم<sup>۳</sup>، تیوری بیض<sup>۴</sup> و k-نزدیک‌ترین همسایه<sup>۵</sup> صورت می‌گیرد. از دو رویکرد شبکه عصبی مصنوعی<sup>۶</sup> و رگرسیون چندمتغیره نیز برای برآورد کمی میزان بازیابی بلوک کانسنگ استفاده شده است. در علوم مختلف از روش‌های داده‌کاوی مثل الگوریتم‌های طبقه‌بندی [۱-۳] و از رویکرد شبکه عصبی مصنوعی، به‌طور گسترده‌ای استفاده شده است [۴-۸]. در سال‌های اخیر، رویکرد شبکه عصبی مصنوعی برای حل مسایل مهندسی معدن و ژئوتکنیک نظیر عملکرد ماشین حفاری تمام‌مقطع و پیش‌بینی پیامدهای نامطلوب آتشفشانی (عقب‌زدگی)، لرزش زمین، لرزش هوا و

پرتاب سنگ) استفاده شده است [۹-۱۹]. به‌عنوان مثال، "ارمغانی و همکاران" کاربرد روش‌های رگرسیون چندمتغیره، رویکرد شبکه عصبی مصنوعی و سیستم استنتاج فازی-عصبی تطبیقی<sup>۷</sup> را برای تخمین مقاومت فشاری تک‌محوره<sup>۸</sup> (UCS) سنگ‌های گرانیت در یک پروژه احداث تونل آب در مالزی مورد مطالعه قرار دادند. در پژوهش آنها، مقدار UCS سنگ بر اساس مقادیر اندازه‌گیری شده شاخص بار نقطه‌ای ( $I_{s(50)}$ )، چکش اشمیت ( $R_n$ ) و سرعت موج p با روش‌های مذکور، پیش‌بینی شده و در نهایت، کارایی روش‌ها بر اساس مقادیر ضریب تعیین<sup>۹</sup> ( $R^2$ )، واریانس و جذر میانگین مربعات خطا<sup>۱۰</sup> مقایسه شده است [۲۰]. "سقط‌فروش و همکاران" از ترکیب الگوریتم‌های شبکه عصبی مصنوعی و کلونی مورچه برای پیش‌بینی پرتاب سنگ و عقب‌زدگی ناشی از انفجار در معدن آهن دلکن (واقع در شمال شرق ایران) استفاده کردند. در پژوهش آنها، تاثیرگذارترین شاخص‌ها بر روی پرتاب سنگ و عقب‌زدگی شامل بارسنگ، فاصله طولی بین چال‌ها، طول چال، طول گل‌گذاری و خرج ویژه به‌عنوان ورودی‌های مدل در نظر گرفته شدند [۲۱]. "ترابی و همکاران" کاربرد شبکه عصبی مصنوعی و رگرسیون چندمتغیره برای پیش‌بینی مقاومت فشاری تک‌محوره و مدول الاستیسیته را براساس ویژگی‌های فیزیکی سنگ آهک در تشکیلات آسماری (واقع در جنوب غرب ایران) مورد مطالعه قرار دادند. آنها سرعت موج p، وزن مخصوص و تخلخل را به‌عنوان ورودی‌های مدل پیش‌بینی خود در نظر گرفتند [۲۲]. "طاهری و همکاران" ترکیبی از الگوریتم‌های شبکه عصبی مصنوعی و کلونی زنبورعسل را برای پیش‌بینی لرزش زمین در اثر آتشفشانی در معدن مس میدوک در ایران استفاده کردند. آنها در پژوهش خود، حداکثر سرعت ذره‌ای<sup>۱۱</sup> را به‌عنوان شاخصی از لرزش زمین براساس دو عامل حداکثر وزن خرج مصرفی در هر تاخیر و فاصله بین ایستگاه مانیتورینگ و نقاط آتشفشانی برآورد کردند. علاوه بر این، مقدار لرزش زمین در اثر انفجار بر پایه مقادیر اندازه‌گیری شده را با چندین مدل تجربی محاسبه کردند. در نهایت، نتایج مدل‌های پیش‌بینی با داده‌های اندازه‌گیری شده بر اساس مقادیر میانگین قدرمطلق خطا<sup>۱۲</sup>، جذر میانگین مربعات خطا و ضریب همبستگی<sup>۱۳</sup> مقایسه شد [۲۳]. "حسینی‌پناه و همکاران" از روش‌های طبقه‌بندی و درخت رگرسیون<sup>۱۴</sup> (CART)، رگرسیون چندمتغیره و چند مدل تجربی برای پیش‌بینی لرزش زمین استفاده کردند [۲۴]. "امیری و همکاران" نیز از

مدل و شرط داشتن مقادیر متغیرهای مستقل، می توان مقدار متغیر وابسته را برای هر مجموعه داده، پیش بینی و آنها را در دسته های مختلف طبقه بندی کرد. فرآیند ساخت مدل، یک فرآیند دو مرحله ای است که در مرحله اول به کمک مجموعه داده های آموزشی<sup>۱۷</sup> که برچسب کلاس تمام نمونه های آن مشخص است، مدل ساخته می شود. این مرحله به نام مرحله یادگیری<sup>۱۸</sup> شناخته می شود. در مرحله دوم، با کمک مجموعه داده های آزمایشی که در آن برچسب مربوط به کلاس هر داده نامعلوم است، مدل به دست آمده، اعتبارسنجی می شود. در واقع، ارزش یابی مدل با توجه به این که کلاس چه تعداد از نمونه داده های آزمایشی درست تخمین زده شده است، محاسبه می شود [۲۷]. جدول ۱ داده های آنالیز عیار خوراک ورودی کارخانه فرآوری معدن مس میدوک را برای سه ترکیب Cu، CuS و CuO، به عنوان متغیرهای ورودی و میزان بازیابی را به عنوان متغیر وابسته به دو شکل کمی و کیفی نشان می دهد. در بیان میزان بازیابی کیفی، نمونه ها به سه دسته بازیابی بالا، متوسط و پایین تقسیم بندی می شوند. به این صورت که بلوک هایی با بازیابی بزرگ تر و مساوی ۸۷ درصد و کم تر از ۸۹٫۲ درصد در گروه بلوک های با بازیابی متوسط، بلوک های با بازیابی بزرگ تر و مساوی ۸۹٫۲، در گروه بلوک های با بازیابی بالا و کم تر از ۸۷ درصد نیز دسته بلوک های با بازیابی پایین قرار می گیرند.

#### ۲-۱- طبقه بندی به روش درخت تصمیم

درخت تصمیم، روشی برای طبقه بندی داده ها است و داده ها را براساس مجموعه ای از آزمون های تعریف شده در هر گره به زیربخش های کوچک تر تقسیم می کند. درخت های تصمیم از یک گره اصلی (ریشه)، یک مجموعه گره میانی و گره های انتهایی (برگ) تشکیل شده اند. با شروع از گره ریشه، درخت تصمیم به شاخه هایی در چپ و راست منشعب می شود که هر شاخه، دامنه ای از مقادیر را نشان می دهد. این فرآیند با استفاده از معیار کم ترین مربعات خطا تا زمانی که خطای تخمین مربوط به متغیر وابسته کمینه شود، ادامه می یابد. شکل ۱ نمونه ای از یک درخت تصمیم را نشان می دهد. هر گره در درخت تصمیم دارای فقط یک مولد<sup>۱۹</sup> و دو یا بیش تر، فرزند<sup>۲۰</sup> است. ویژگی برجسته الگوریتم های درخت تصمیم، تفسیر پذیری ساده آنها با استفاده از یک ساختار درخت سلسله مراتبی است. در مقایسه با دیگر روش های یادگیری ماشین،

شاخص های ورودی یکسان با مدل طاهری و همکاران، برای مدل پیش بینی ترکیبی ANN-KNN<sup>۱۵</sup> خود در برآورد لرزش زمین و بیش فشار هوای ناشی از آتشفشانی در پروژه سد رودخانه شور در استان کرمان ایران استفاده کردند [۲۵]. "صیامی و دیندارلو" مصرف سوخت کامیون های معدنی را با استفاده از رویکرد شبکه عصبی پیش بینی کردند. مدل پیش بینی آنها، مصرف سوخت هر سیکل را بر اساس ظرفیت ترابری، مدت زمان بارگیری، مدت زمان تلف شده برای کامیون های پر، زمان سفر کامیون های پر، زمان سفر کامیون های خالی و زمان تلف شده برای کامیون های خالی تخمین می زند [۲۶].

هدف از این پژوهش، ساخت مدلی برای برآورد بازیابی کارخانه فرآوری در کانسار مس میدوک با توجه به تغییرات عیار عناصر تشکیل دهنده آن است. به این منظور، ابتدا با معرفی سه الگوریتم طبقه بندی درخت تصمیم، تیوری بیز و نزدیک ترین همسایه، برآورد کیفی بازیابی انجام و مقایسه ای بین این سه روش صورت می گیرد. سپس با معرفی شبکه عصبی مصنوعی در برآورد کمی میزان بازیابی، کاربرد آن در پیش بینی بازیابی، بررسی و نتایج حاصل با روش رگرسیون چندمتغیره مقایسه می شود. در نهایت، با اعتبارسنجی مدل ها، بهترین مدل در هر یک از رویکردهای کیفی و کمی پیش بینی بازیابی، انتخاب می شوند. ایجاد مدل پیش بینی بازیابی و اعتبارسنجی آن بر اساس میزان عیار Cu، CuS و CuO آنالیز شده در ۵۸ نمونه از بلوک های کانسنگ ورودی به کارخانه و میزان بازیابی اندازه گیری شده Cu در محصول نهایی کارخانه انجام شده است. به این منظور، تعداد ۵۸ نمونه در دوره های زمانی مختلف از خوراک ورودی کارخانه و کنسانتره نهایی حاصل، نمونه برداری و پس از آماده سازی برای آنالیز و دستیابی به داده های مورد نیاز به آزمایشگاه کارخانه فرآوری ارسال شده است. در نهایت پس از جمع آوری اطلاعات و داده ها و مطالعه روی آنها، کارایی روش های طبقه بندی، شبکه عصبی مصنوعی و رگرسیون چندمتغیره خطی برای پیش بینی بازیابی هر بلوک استخراجی بررسی شد.

#### ۲-۲- برآورد کیفی بازیابی با استفاده از مفهوم طبقه بندی

مفهوم طبقه بندی در علم داده کاوی به روش نظارت شده ای<sup>۱۶</sup> اطلاق می شود که تلاش می کند رابطه ای میان چند متغیر مستقل و یک متغیر وابسته را کشف کند. این رابطه با یک ساختار به عنوان "مدل"، نمایش داده می شود. به کمک

جدول ۱: داده‌های آنالیز عیاری خوراک ورودی کارخانه فرآوری معدن مس میدوک

نمونه	Cu (%)	CuO (%)	CuS (%)	بازیابی (%)	طبقه بازیابی	نمونه	Cu (%)	CuO (%)	CuS (%)	بازیابی (%)	طبقه بازیابی
۱	۰.۶۱۱۹	۰.۳۳۳	۰.۵۷۸۶	۸۸.۰۴۸۹	متوسط	۳۰	۰.۶۳۵۵	۰.۴۶۷	۰.۵۸۸۸	۸۹.۲۵۸۸	بالا
۲	۰.۷۷۵۷	۰.۳۳۳	۰.۷۴۲۳	۹۲.۸۲۱۳	بالا	۳۱	۰.۶۹۳۰	۰.۵۳۲	۰.۶۳۹۸	۸۶.۷۵۸۶	پایین
۳	۰.۷۷۴۱	۰.۳۰۰	۰.۷۴۴۱	۹۰.۱۱۳۴	بالا	۳۲	۰.۶۸۲۰	۰.۴۶۷	۰.۶۳۵۲	۸۵.۶۵۷۸	پایین
۴	۰.۶۷۴۵	۰.۳۰۰	۰.۶۴۴۵	۹۱.۸۲۱۸	بالا	۳۳	۰.۶۱۸۵	۰.۴۰۰	۰.۵۷۸۵	۸۸.۷۵۸۳	متوسط
۵	۰.۶۴۱۴	۰.۳۰۰	۰.۶۱۱۴	۸۸.۹۳۴۷	متوسط	۳۴	۰.۷۳۲۲	۰.۵۳۶	۰.۶۷۸۵	۸۸.۰۶۹۵	متوسط
۶	۰.۶۱۹۷	۰.۳۳۹	۰.۵۸۵۸	۹۲.۵۵۳۸	بالا	۳۵	۰.۷۲۳۲	۰.۵۰۱	۰.۶۷۳۱	۸۸.۷۵۸۵	متوسط
۷	۰.۶۴۰۸	۰.۳۰۰	۰.۶۱۰۸	۹۴.۴۰۲۷	بالا	۳۶	۰.۶۳۹۱	۰.۴۰۷	۰.۵۹۸۴	۸۷.۹۸۴۰	متوسط
۸	۰.۸۳۹۱	۰.۳۰۰	۰.۸۰۹۱	۹۴.۶۴۰۳	بالا	۳۷	۰.۶۹۵۲	۰.۵۷۶	۰.۶۳۷۶	۸۶.۶۰۰۱	پایین
۹	۰.۸۲۴۲	۰.۳۲۹	۰.۷۹۱۲	۹۳.۰۲۸۳	بالا	۳۸	۰.۶۲۳۵	۰.۴۰۰	۰.۵۸۳۵	۸۱.۲۹۵۸	پایین
۱۰	۰.۶۵۸۸	۰.۲۰۰	۰.۶۳۸۸	۹۰.۳۴۰۶	بالا	۳۹	۰.۶۵۹۵	۰.۳۹۱	۰.۶۲۰۳	۸۷.۹۴۲۳	متوسط
۱۱	۰.۷۲۵۱	۰.۴۳۴	۰.۶۸۱۶	۸۹.۲۰۴۷	بالا	۴۰	۰.۷۹۹۷	۰.۵۶۶	۰.۷۴۳۱	۸۸.۶۱۲۷	متوسط
۱۲	۰.۷۲۲۷	۰.۴۳۷	۰.۶۷۹۰	۸۹.۵۰۹۷	بالا	۴۱	۰.۶۸۰۱	۰.۵۶۷	۰.۶۲۳۳	۸۸.۵۲۷۵	متوسط
۱۳	۰.۶۴۵۸	۰.۵۳۳	۰.۵۹۲۵	۸۶.۵۱۶۴	پایین	۴۲	۰.۶۴۷۰	۰.۵۶۵	۰.۵۹۰۵	۸۷.۸۰۳۶	متوسط
۱۴	۰.۷۵۶۷	۰.۶۰۷	۰.۶۹۵۹	۸۲.۶۹۸۹	پایین	۴۳	۰.۶۸۳۱	۰.۴۰۰	۰.۶۴۳۱	۹۰.۴۸۲۱	بالا
۱۵	۰.۷۴۰۷	۰.۶۳۰	۰.۶۷۷۷	۸۷.۰۹۷۹	متوسط	۴۴	۰.۶۷۰۴	۰.۴۳۳	۰.۶۲۷۰	۸۸.۳۲۵۳	متوسط
۱۶	۰.۹۳۴۶	۰.۷۷۷	۰.۸۵۶۹	۸۷.۶۴۶۸	متوسط	۴۵	۰.۶۳۳۵	۰.۵۰۰	۰.۵۸۳۵	۸۵.۰۲۲۷	پایین
۱۷	۰.۸۱۷۴	۰.۵۳۰	۰.۷۶۴۴	۹۰.۰۸۰۵	بالا	۴۶	۰.۶۰۰۹	۰.۴۶۷	۰.۵۵۴۲	۸۹.۰۹۵۹	متوسط
۱۸	۰.۷۴۰۲	۰.۵۰۷	۰.۶۸۹۴	۹۰.۱۸۰۱	بالا	۴۷	۰.۶۶۲۹	۰.۴۶۲	۰.۶۱۶۶	۸۸.۸۰۱۷	متوسط
۱۹	۰.۵۳۵۷	۰.۴۶۳	۰.۴۸۹۴	۸۳.۶۶۳۸	پایین	۴۸	۰.۷۱۲۸	۰.۵۰۰	۰.۶۶۲۸	۷۵.۵۲۶۶	پایین
۲۰	۰.۷۱۹۳	۰.۴۳۰	۰.۶۷۶۳	۸۸.۵۱۶۷	متوسط	۴۹	۰.۷۲۷۴	۰.۵۳۲	۰.۶۷۴۲	۸۳.۸۲۳۴	پایین
۲۱	۰.۶۶۶۸	۰.۶۳۳	۰.۶۰۳۵	۸۴.۸۶۳۲	پایین	۵۰	۰.۶۵۱۰	۰.۵۰۰	۰.۶۰۱۰	۸۳.۲۶۳۵	پایین
۲۲	۰.۷۲۶۶	۰.۴۹۴	۰.۶۷۷۲	۸۷.۸۷۰۲	متوسط	۵۱	۰.۶۲۹۶	۰.۵۳۷	۰.۵۷۵۹	۸۵.۹۱۴۵	پایین
۲۳	۰.۷۵۴۴	۰.۵۳۳	۰.۷۰۱۰	۸۵.۵۹۷۰	پایین	۵۲	۰.۶۵۵۲	۰.۴۳۴	۰.۶۱۱۷	۸۴.۶۸۱۴	پایین
۲۴	۰.۷۵۲۱	۰.۵۷۰	۰.۶۹۵۰	۸۳.۰۱۰۱	پایین	۵۳	۰.۷۱۸۱	۰.۴۳۷	۰.۷۳۸۰	۸۶.۱۸۲۹	پایین
۲۵	۰.۶۹۹۱	۰.۶۴۲	۰.۶۳۴۹	۸۵.۸۲۰۲	پایین	۵۴	۰.۷۲۳۳	۰.۴۲۹	۰.۶۸۰۳	۸۸.۶۹۷۳	متوسط
۲۶	۰.۶۳۴۹	۰.۵۰۰	۰.۵۸۴۹	۹۰.۱۸۹۳	بالا	۵۵	۰.۷۹۱۵	۰.۵۳۶	۰.۷۳۷۸	۹۱.۶۹۹۳	بالا
۲۷	۰.۶۴۵۶	۰.۵۰۰	۰.۵۹۵۶	۸۸.۲۴۶۱	متوسط	۵۶	۰.۶۰۴۹	۰.۵۰۲	۰.۵۵۴۷	۸۹.۸۱۳۴	بالا
۲۸	۰.۷۵۸۱	۰.۶۲۶	۰.۶۹۵۴	۸۹.۴۲۸۲	بالا	۵۷	۰.۷۱۴۴	۰.۴۰۱	۰.۶۷۴۴	۸۷.۰۴۵۰	متوسط
۲۹	۰.۶۳۳۶	۰.۶۶۵	۰.۵۶۷۱	۷۹.۴۵۷۳	پایین	۵۸	۰.۷۳۷۶	۰.۴۶۹	۰.۶۹۰۷	۸۵.۸۹۷۳	پایین

بالا در برخی از الگوریتم‌های ایجاد درخت، یکی از معایب الگوریتم‌های درخت تصمیم برای داده‌های حجیم است. علاوه بر این، در برخی از روش‌ها، هنگامی که تعداد مثال‌ها یا رکوردهای هر طبقه کم باشد، نرخ خطا افزایش می‌یابد [۳۰].

#### ۲-۲- طبقه‌بندی بر اساس تئوری بیز

یکی از قوانین مهم احتمال، قانون احتمال "بیز" است که به کمک آن می‌توان احتمال برچسب طبقه یک نمونه از داده‌ها را تخمین زد. در این روش، فرض بر این است که تاثیر مقدار

الگوریتم‌های درخت تصمیم به‌عنوان روش‌های جعبه سفید<sup>۲۱</sup> که درک رابطه بین متغیرهای ورودی و خروجی را آسان‌تر می‌کنند، شناخته می‌شوند.

مزیت این روش نسبت به سایر روش‌های نظارت‌شده، این است که روشی غیرپارامتریک است و نیاز به در نظر گرفتن فرضیاتی در رابطه با توزیع داده‌های ورودی ندارد. همچنین این طبقه‌بندی‌کننده می‌تواند روابط غیرخطی بین داده‌ها را مدنظر قرار دهد و برای انواع مختلف داده‌ها از قبیل داده‌های عددی و طبقه‌ای قابل استفاده است [۲۹، ۲۸]. نیاز به حافظه

### ۲-۳- طبقه بندی بر اساس الگوریتم $k$ -نزدیک ترین همسایه

در روش  $k$ -نزدیک ترین همسایه، از یک تابع چگالی احتمالی ساده تر نسبت به روش بیزی استفاده می شود. برای این منظور، احتمال این که  $k$  نمونه از تعداد کل  $N$  نمونه در داخل فضایی به حجم  $V$  و به مرکز نمونه جدید  $x$  قرار گیرند، از رابطه ۲ محاسبه می شود:

$$\hat{p}(x) = \frac{k}{NV} \quad (2)$$

در این روش، با ثابت فرض کردن نسبت  $k/N$  (یعنی ثابت فرض کردن  $k$ )، حجم طوری تغییر می کند تا  $k$  نمونه در داخل آن قرار گیرند. حال، اگر فرض شود از  $k$  نمونه، نمونه  $k_i$  متعلق به  $\omega_i$  و  $k = \sum_i k_i$  باشد، آن گاه از رابطه ۳، احتمال شرطی به دست خواهد آمد:

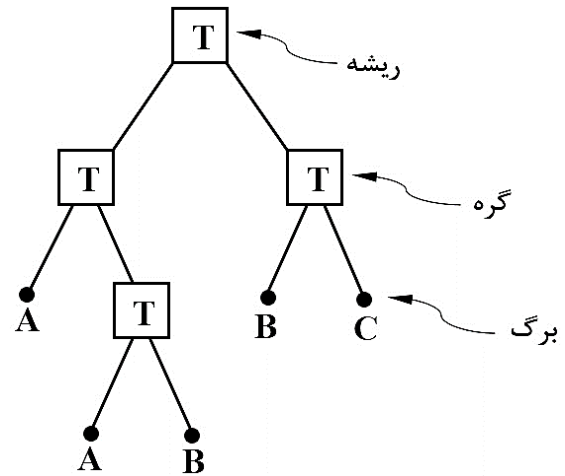
$$g_i(x) = \hat{p}(x | \omega_i) \times p(\omega_i) = \frac{k_i}{n_i V} \times \frac{n_i}{N} = \frac{k_i}{NV} \quad (3)$$

بنابراین نمونه جدید به طبقه ای نسبت خواهد داشت که دارای تعداد نمونه های بیش تری در داخل حجم  $V$  باشد. الگوریتم روش  $k$ -نزدیک ترین همسایه، محاسبه فاصله نمونه جدید از تمامی نمونه های آموزشی و مرتب کردن آنها به ترتیب صعودی و سپس، انتخاب  $k$  نمونه اول خواهد بود [۳۳]. اگر چه روش نزدیک ترین همسایه، روش ساده و موثری است ولی سرعت پردازش آن، پایین است. اگر اندازه مجموعه آموزشی  $D$  و  $k=1$  باشد، طبقه بندی نمونه جدید از مرتبه زمانی  $D$  یعنی  $O(D)$  خواهد بود. برای اجرای این روش در داده های حجیم باید از رایانه هایی با قدرت محاسباتی بالا استفاده کرد [۳۰].

### ۳- برآورد کمی بازیابی

#### ۳-۱- مدل سازی به روش شبکه عصبی مصنوعی

روش شبکه عصبی مصنوعی برای اولین بار توسط "مک کولاج" [۲۴] و "پیتز" [۲۵] ابداع شد [۳۳]. اساس شبکه عصبی، شبیه سازی رفتار مغز انسان است [۳۳]. در بین تمام خواص مهم شبکه عصبی، خاصیت یادگیری شبکه، اهمیت بالایی دارد. شبکه های عصبی، به دلیل انعطاف و قابلیت یادگیری بالا، به عنوان سیستم های یادگیر دارای این توانایی هستند که از گذشته بیاموزند و رفتار خود را هم زمان با یادگیری، بهبود بخشند. شبکه های عصبی، مدل های محاسباتی هستند



شکل ۱: نمونه ای از یک درخت تصمیم [۲۸]

یک صفت خاصه<sup>۲۲</sup> بر روی برجسب طبقه، مستقل از مقادیر دیگر صفات خاصه است و این فرض، "استقلال شرطی"<sup>۲۳</sup> طبقه نامیده می شود. در صورت وجود  $C$  طبقه  $\omega_1, \omega_2, \dots, \omega_c$  و با هدف نسبت دادن نمونه جدید  $x$  به یکی از این طبقه ها، باید احتمال های شرطی  $P(\omega_i | x) (i=1, 2, \dots, c)$  محاسبه شود. بدیهی است نمونه جدید  $x$  به طبقه ای تعلق دارد که دارای بیش ترین احتمال شرطی باشد. ولی از آنجا که محاسبه احتمال های شرطی به صورت مستقیم میسر نیست، از تئوری بیزین مطابق با رابطه ۱ استفاده می شود:

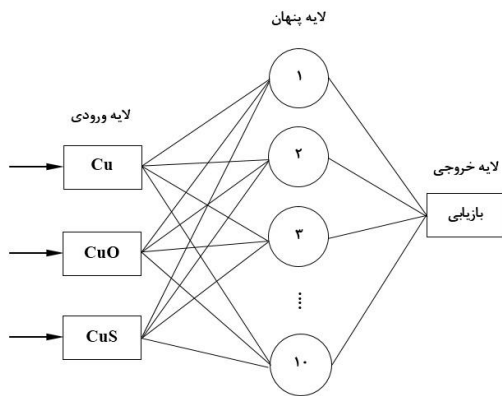
$$p(\omega_i | x) = \frac{p(x | \omega_i) p(\omega_i)}{p(x)} \quad (1)$$

که در آن:

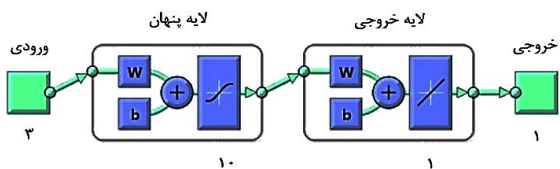
$P(x)$ : احتمال غیرشرطی

$P(\omega_i)$ : احتمال اولیه هر طبقه است.

احتمال اولیه هر طبقه، از تقسیم تعداد نمونه های هر طبقه بر تعداد کل نمونه ها به دست می آید. بنابراین نمونه جدید  $x$  به طبقه ای تعلق خواهد داشت که  $g_i(x) = P(x | \omega_i) P(\omega_i)$  بالاتری داشته باشد [۳۱، ۳۲]. استفاده از این قانون برای طبقه بندی، دقت و سرعت خوبی را در پایگاه داده های بزرگ به همراه دارد. علاوه بر این، اجرای آسان و نتایج خوب از دیگر نقاط قوت این الگوریتم هستند. شایان ذکر است در مواردی که فرض استقلال شرطی طبقه ها برقرار نیست، دقت مدل کاهش می یابد و این مورد، مهم ترین نقطه ضعف روش بیز است [۳۰].



شکل ۲: ساختار شبکه عصبی مصنوعی برای پیش‌بینی بازایی



شکل ۳: ساختار شبکه عصبی مصنوعی در نرم‌افزار MATLAB

### ۲-۲- مدل‌سازی به روش رگرسیون چندمتغیره

رگرسیون خطی چندمتغیره یکی از رایج‌ترین ابزارهای آماری برای انطباق یک معادله خطی بین دو متغیر مستقل یا بیش‌تر و یک متغیر وابسته است. این مدل به‌طور گسترده‌ای توسط پژوهشگران در حل مسایل مختلف مهندسی استفاده شده است [۳۷-۳۹]. به‌طور کلی، مدل رگرسیون خطی چندمتغیره می‌تواند به‌صورت رابطه ۴ بیان شود:

$$Y = P_0 + P_1 X_1 + \dots + P_n X_n \quad (۴)$$

که در آن:

$X_i (i=1, \dots, n)$ : متغیرهای مستقل

$Y$ : متغیر وابسته

$P_i (i=0, 1, \dots, n)$ : ضرایب رگرسیون را بیان می‌کند.

آنالیز رگرسیون چندمتغیره بین مقادیر عیاری Cu، CuO و CuS ۴۷ نمونه از ۵۸ نمونه خوراک ورودی کارخانه، و مقدار بازایی فلز، به روش رگرسیون گام به گام توسط نرم‌افزار SPSS ورژن ۱۶ انجام شد. در این روش، ابتدا تمامی متغیرهای مستقل وارد فرآیند تحلیل می‌شوند. سپس متغیر یا متغیرهای مستقلی که در روند محاسبات، تأثیری محسوس

که می‌توانند رابطه بین ورودی‌ها و خروجی‌های یک دستگاه فیزیکی را با شبکه‌ای از گره‌های متصل به هم، تعیین کنند. لازمه گسترش و توسعه یک مدل شبکه عصبی، طراحی اجزای فنی تشکیل‌دهنده آن است. یک شبکه عصبی از عناصر پردازش (نورون)<sup>۲۶</sup> تشکیل شده است. ساختار شبکه‌های عصبی به شکلی است که نورون‌ها در دسته‌هایی که لایه نام دارند، مرتب می‌شوند. این ساختار متشکل از سه لایه ورودی، پنهان و خروجی است. البته یک شبکه می‌تواند چندین لایه پنهان داشته باشد [۳۴]. در این پژوهش، شبکه‌های چندلایه "پرسپترون" (MLP)<sup>۲۷</sup> که یکی از پرکاربردترین انواع شبکه‌های عصبی مصنوعی برای حل مسایل مهندسی است، مورد استفاده قرار گرفته است [۳۵].

طراحی این نوع از مدل‌های شبکه عصبی مصنوعی نیازمند دسته‌داده‌های آموزشی، اعتبارسنجی و آزمون است. داده‌های آموزشی به‌منظور یافتن رابطه بین ورودی‌ها و خروجی‌های اندازه‌گیری‌شده، داده‌های اعتبارسنجی برای کنترل و نظارت بر یادگیری صحیح شبکه و داده‌های آزمون برای ارزیابی عملکرد شبکه پیشنهادی استفاده می‌شود. در این پژوهش، ۷۰، ۱۵ و ۱۵ درصد از کل ۴۷ نمونه آنالیزشده به ترتیب برای آموزش، اعتبارسنجی و آزمون مدل اختصاص یافت. در طراحی شبکه، مقادیر عیاری حاصل از آنالیز خوراک ورودی کارخانه فرآوری شامل عیار Cu، CuO و CuS، به‌عنوان ورودی و مقدار بازایی Cu، به‌عنوان خروجی شبکه در نظر گرفته شدند. شکل ۲، ساختار شبکه عصبی مصنوعی برای پیش‌بینی بازایی با استفاده از ویژگی‌های عیاری خوراک ورودی کارخانه را نشان می‌دهد. این شبکه دارای ۳ نورون در لایه ورودی، ۱۰ نورون در لایه مخفی و یک نورون در لایه خروجی است. برای اجرای شبکه عصبی از نرم‌افزار MATLAB [۳۷]، نسخه R2016a استفاده شد.

شکل ۳، ساختار شبکه عصبی در نرم‌افزار MATLAB را نشان می‌دهد، که  $w$  وزن بین هر نورون ورودی و نورون خروجی و  $b$  سوگیری<sup>۲۸</sup> هر نورون را نشان می‌دهد. فرآیند آموزش شامل تغییر وزن بین لایه‌های مختلف در طول آموزش است و تا جایی‌که تفاوت بین داده‌های واقعی (برای داده‌های آزمون) و داده‌های پیش‌بینی‌شده به حداقل برسد، ادامه می‌یابد و در نهایت، بهترین ساختار شبکه برای بازایی فلز براساس کم‌ترین میزان میانگین مربعات خطا (MSE) انتخاب خواهد شد. ۱۱ داده آزمایشی هم برای پیش‌بینی دقت مدل به‌کار گرفته می‌شوند.

## ۴- نتایج

## ۴-۱- نتایج برآورد بازیابی به صورت کیفی

الگوریتم درخت تصمیم برای ۵۸ داده جدول ۱، توسط نرم افزار MATLAB اجرا شد و نتایج در شکل ۴، نمایش داده شده است. علاوه براین، روش های بیزین ساده و  $k$ -نزدیک ترین همسایه (برای  $k=4$ ) برای مجموعه داده های جدول ۱ اجرا شده و ساخت مدل های پیش بینی کننده برای این سه روش انجام و مدل های حاصل، اعتبارسنجی شدند. همان طور که پیش تر بیان شد، اعتبارسنجی مدل ها با توجه به این که در طبقه، چه تعداد از داده های آزمایشی (۱۱ نمونه) درست تخمین زده شده است، محاسبه می شود. به عبارت دیگر، مدل ابتدا با استفاده از نمونه های آموزشی، طراحی و سپس بر پایه عملکرد نمونه های آزمایشی، ارزیابی می شود. از آن جا که قسمتی از داده ها برای آزمایش کنار گذاشته شده است، احتمال این که مدل تولید شده، سنجش مناسبی بر روی داده های آزمایشی نداشته باشد، وجود دارد. چون ممکن است یک برچسب کلاس در داده های آزمایشی وجود داشته باشد که نتوان آن را در داده های آزمایشی پیدا کرد. بنابراین برای اعتبارسنجی مدل از روش  $k$ -fold cross-validation استفاده شده است. در این روش، هر نمونه از داده های اصلی به تعداد دفعات یکسانی

در ارتباط با متغیر وابسته نداشته باشند، با حذف از تحلیل، وارد معادله نمی شوند [۴۰]. برای اعتبارسنجی مدل، از ۱۱ نمونه آنالیز شده آزمایشی استفاده شد. معیارهای میانگین قدرمطلق خطا (MAE)، جذر میانگین مربعات خطا (RMSE) و ضریب تعیین ( $R^2$ )، به ترتیب با استفاده از روابط ۵ تا ۷ قابل محاسبه هستند [۲۰]:

$$MAE = \frac{\sum_{i=1}^n (|R^*(i) - R(i)|)}{n} \quad (5)$$

$$RMSE = \left( \frac{\sum_{i=1}^n (R^*(i) - R(i))^2}{n} \right)^{1/2} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (R(i) - R^*(i))^2}{\sum_{i=1}^n (R(i) - m(i))^2} \quad (7)$$

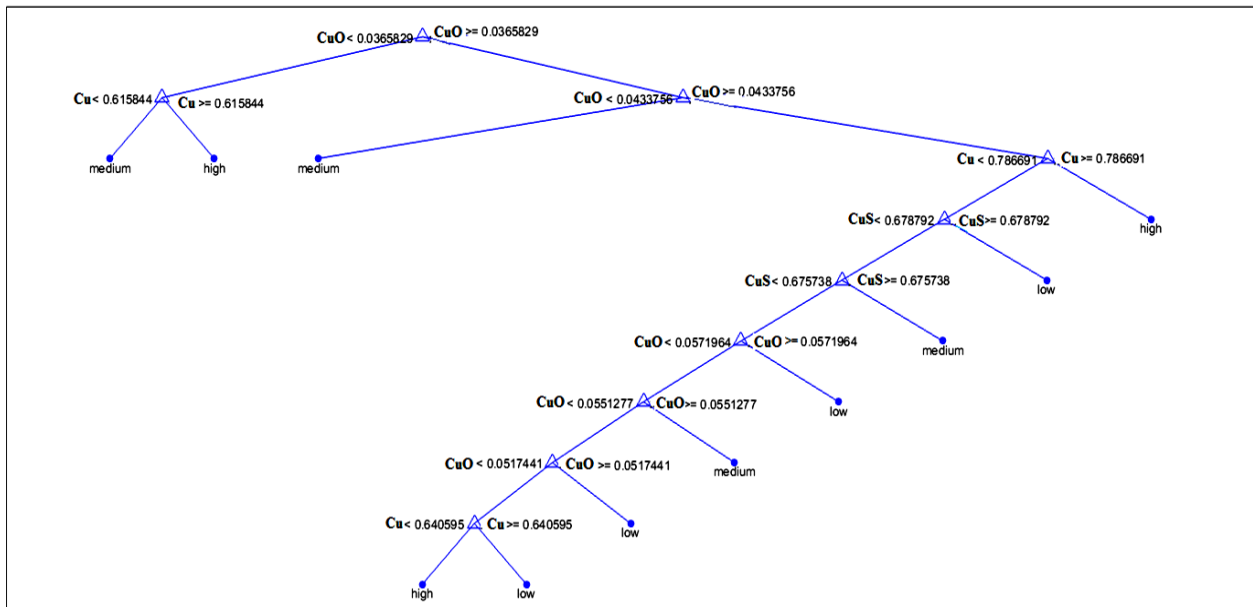
که در آن:

$R^*(i)$ : مقدار بازیابی پیش بینی شده با مدل

$R(i)$ : مقدار بازیابی آنالیز شده

$m(i)$ : میانگین مقدار آنالیز شده

$n$ : تعداد نمونه ها است.



شکل ۴: درخت تصمیم برای داده های آنالیز شده

می‌دهد. همان‌طور که در شکل ۶ نشان داده شده است، مدل مزبور توانسته است ۷۷ درصد از تغییرات مقدار بازیابی فلز در محصول نهایی را توجیه کند.

نتایج اعتبارسنجی مدل، مقدار MAE برابر با ۰٫۰۲۱۷۰۲ نشان می‌دهد که مقدار قابل قبولی بوده و این امر بیان‌گر ناریب بودن برازش توسط مدل ایجاد شده است.

مقدار  $RMSE = ۰٫۰۲۴۹۷۲$  نسبتاً پایین مدل نیز نشان‌دهنده دقت مناسب و قابل قبول مدل است. رابطه ۸ نشان می‌دهد، رابطه بین مقدار بازیابی با عیار Cu، مستقیم و با عیار CuO، معکوس است.

نکته جالب توجه این‌که، عدم حضور متغیر CuS، در رابطه محاسبه بازیابی به‌دست‌آمده از مدل رگرسیونی، نشان‌دهنده عدم تاثیر محسوس مقدار عیار CuS خوراک ورودی بر مقدار بازیابی فلز در محصول نهایی، براساس این مدل است.

#### ۴-۲-۲- نتایج مدل‌سازی شبکه عصبی مصنوعی

در شکل ۷ که تغییرات مقدار میانگین مربعات خطا (MSE) شبکه در گام‌های مختلف را نشان می‌دهد، بهترین خطای شبکه که با خط‌چین نمایش داده شده است، برابر با ۰٫۰۰۰۷۰۹۰۳ و در گام ۲۵ است؛ ضمن این‌که مقدار میانگین مربعات خطای شبکه، مقدار ۰٫۰۰۰۳۰۴ را نشان می‌دهد. زمانی مسیر اجرای آموزش شبکه، صحیح و قابل اعتماد خواهد بود که مقدار میانگین مربعات خطای شبکه کم‌تر از بهترین خطای شبکه باشد و دو خطی که روند اعتبارسنجی و آموزش را نشان می‌دهند، به هم نزدیک‌تر باشند و در این پژوهش، این مهم محقق شده است. شکل ۸، نمودارهای رابطه بین ورودی و خروجی مربوط به داده‌های مرحله طراحی شبکه عصبی مصنوعی (۴۷ نمونه) را نشان می‌دهد. همان‌طور که

در مجموعه آموزشی قرار می‌گیرد و فقط، یک‌بار برای آزمایش انتخاب می‌شود. بنابراین با انتخاب  $k=5$ ، مجموعه داده‌ها به ۵ قسمت تقسیم شد که چهار قسمت اول، ۱۱ عضو و قسمت پنجم، ۱۴ عضو دارد. هر بار یک قسمت از این ۵ قسمت، برای آزمایش کنار گذاشته و با کمک ۴ قسمت دیگر، مدل طراحی شد. به‌این ترتیب دقت یا سطح اطمینان مدل با محاسبه میانگین این ۵ تکرار حاصل و سطح اطمینان الگوریتم درخت تصمیم، بیز ساده و  $k$ -نزدیک‌ترین همسایه به ترتیب ۷۵، ۶۰ و ۶۶ درصد به‌دست آمد. در نتیجه، مشخص شد که الگوریتم درخت تصمیم، می‌تواند بهتر از دو مدل دیگر، بازیابی را به‌صورت کیفی تخمین بزند. اکنون با دراختیار داشتن عیارهای Cu، CuO و CuS برای کل مدل بلوکی ذخیره که پیش‌تر با استفاده از روش‌های زمین‌آماری محاسبه شده است، می‌توان میزان بازیابی هر بلوک را تخمین زد. شکل ۵، مقطعی از مدل بلوکی کلاسه‌بندی‌شده با روش درخت تصمیم را که دارای بیش‌ترین سطح اطمینان است، نشان می‌دهد.

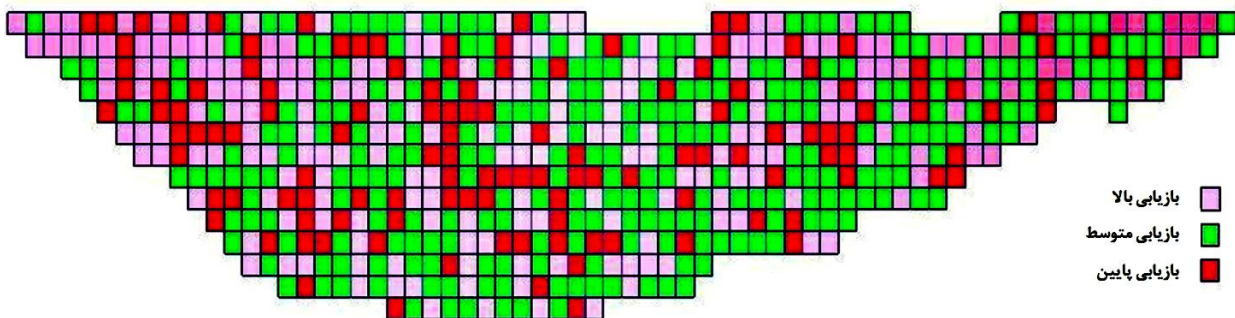
#### ۴-۲-۴ نتایج برآورد بازیابی به‌صورت کمی

#### ۴-۲-۴-۱ نتایج رابطه رگرسیونی بین مقادیر عیار خوراک ورودی کارخانه و مقدار بازیابی

رابطه ۸ مدل روش رگرسیون چندمتغیره برای پیش‌بینی مقدار بازیابی را نشان می‌دهد:

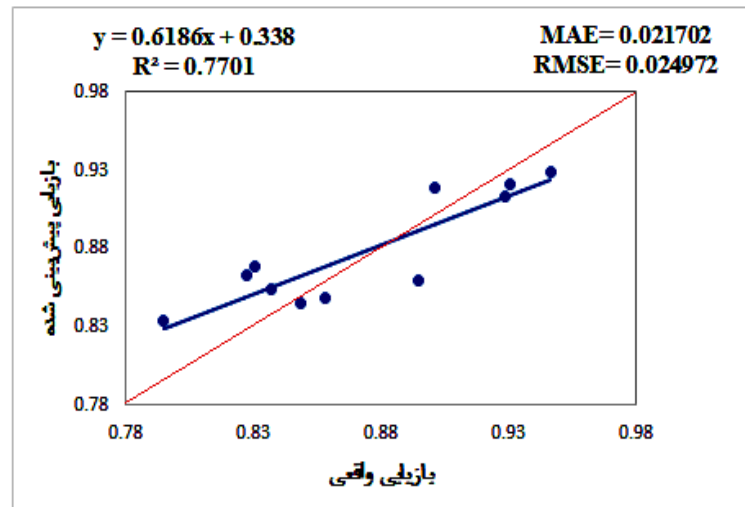
$$R = 0.851 - 1.725CuO + 0.154Cu \quad (8)$$

که  $R$ ، مقدار بازیابی را نشان می‌دهد. شکل ۶ نیز ارتباط بین بازیابی آنالیزشده (واقعی) و بازیابی پیش‌بینی‌شده در مدل رگرسیون چندمتغیره را برای ۱۱ نمونه آزمایشی نشان

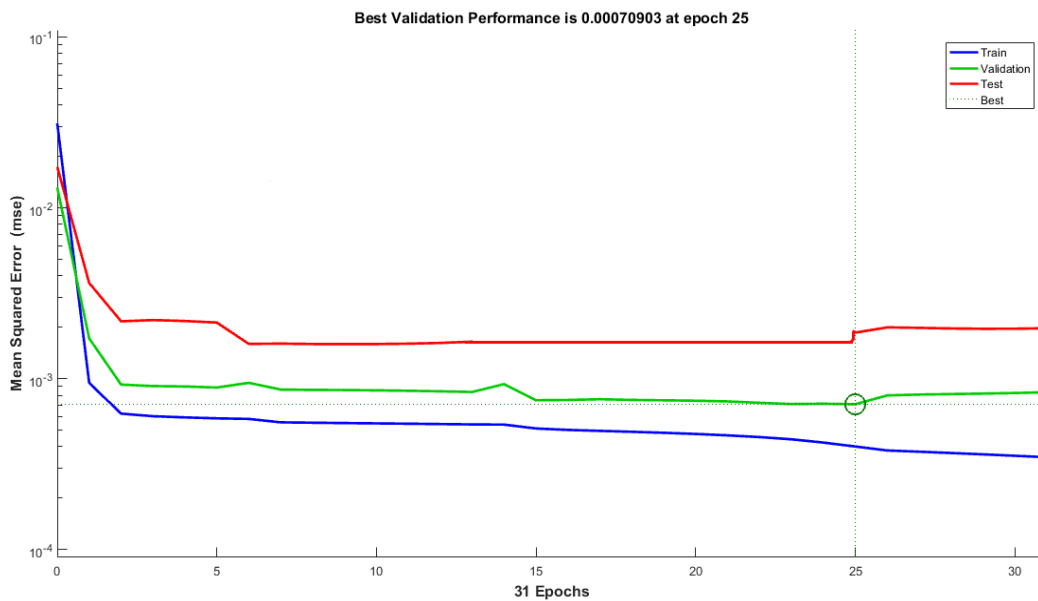


شکل ۵: مقطعی از مدل بلوکی کلاسه‌بندی‌شده با روش درخت تصمیم





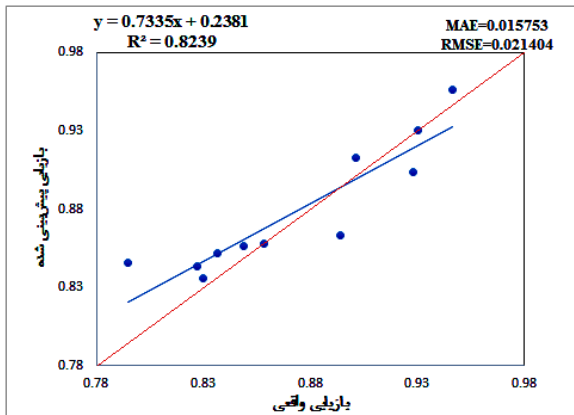
شکل ۶: ارتباط بین بازیابی آنالیز شده و بازیابی پیش‌بینی شده در مدل رگرسیون



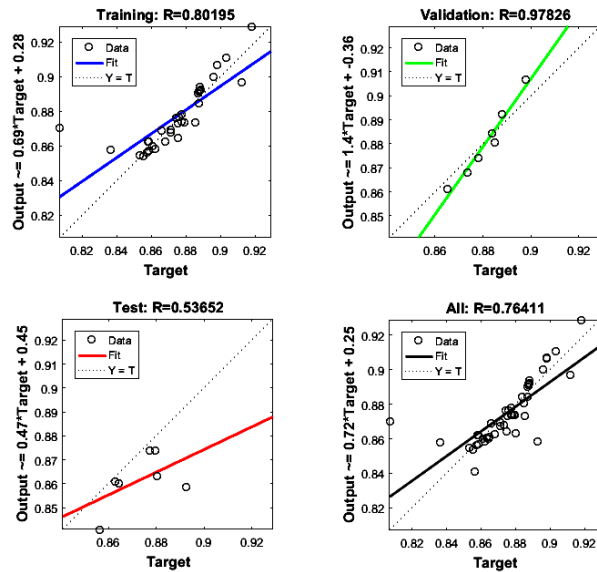
شکل ۷: تغییرات مقدار میانگین مربعات خطا (MSE) شبکه در گام‌های مختلف

درصد، ۹۷ درصد و ۵۳ درصد است. برای جلوگیری از معضل بیش‌برازش<sup>۲۹</sup> نیز در کد شبکه عصبی نوشته شده، دستور توقف فرآیند آموزش براساس رسیدن به تعداد مشخصی خطای اعتبارسنجی<sup>۳۰</sup> تعبیه شده است. بنابراین با توجه به همبستگی بالای حاصل شده به‌ویژه در مراحل آموزش و اعتبارسنجی، می‌توان اطمینان حاصل کرد که شبکه طراحی شده، فرآیند آموزش را به‌صورت مطلوبی طی کرده و در مسیر آموزش، بیش‌برازش و کم‌برازش<sup>۳۱</sup> رخ نداده است.

پیش‌تر توضیح داده شد، در فرآیند مدل‌سازی این پژوهش، ۷۰ درصد نمونه‌ها (۳۳ نمونه) برای آموزش، ۱۵ درصد (۷ نمونه) برای اعتبارسنجی هم‌زمان با آموزش و ۱۵ درصد باقیمانده (۷ نمونه) نیز برای آزمون، تفکیک شدند. این تفکیک توسط شبکه و به صورت تصادفی روی داده‌های مرحله طراحی انجام شد، اما درصد نمونه‌های اختصاص یافته به هر مرحله، توسط کاربر تعیین شد. شکل ۸ نشان می‌دهد، ضریب همبستگی مرحله آموزش، اعتبارسنجی و آزمون به ترتیب برابر با ۸۰



شکل ۹: ارتباط بین بازایی آنالیز شده و بازایی پیش‌بینی شده در مدل شبکه عصبی



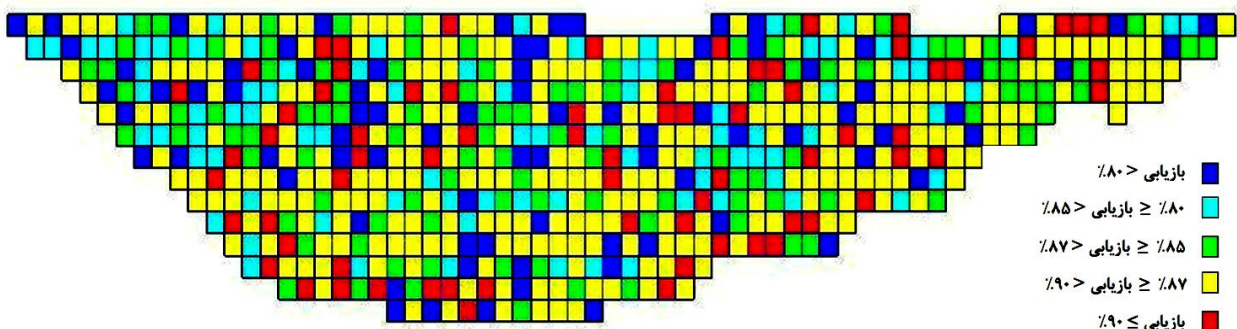
شکل ۸: نمودارهای رابطه بین ورودی و خروجی مربوط به داده‌های مرحله طراحی شبکه عصبی مصنوعی

جدول،  $R$  نشان‌دهنده میزان بازایی بر حسب درصد است. شکل ۱۰ نیز مقطعی از مدل بلوکی کلاسه‌بندی شده با روش شبکه عصبی مصنوعی را نشان می‌دهد.

#### ۴-۲-۳- آنالیز حساسیت در روش شبکه عصبی

فرآیند آنالیز حساسیت، اطلاعات ارزشمندی درباره میزان حساسیت مدل به متغیرهای ورودی آن در اختیار طراح مدل قرار می‌دهد. با شناسایی میزان تاثیر متغیرهای ورودی بر دقت پیش‌بینی مدل، می‌توان متغیرهای کم‌اثر را از شبکه، حذف و مدل ساده‌تری را گسترش و توسعه داد. در این پژوهش برای انجام آنالیز حساسیت مدل از روش "آمار نرم" استفاده شد. در این روش مقدار ضریب حساسیت هر متغیر ورودی از تقسیم کردن خطای کل شبکه در غیاب آن متغیر، بر خطای کل شبکه در حضور تمام متغیرهای ورودی محاسبه می‌شود [۴۱]. به‌طور معمول برای تعیین مقدار تاثیر شاخص‌های

علاوه بر این، ارتباط بین مقدار بازایی آنالیز شده و مقدار بازایی پیش‌بینی شده در مدل شبکه عصبی مصنوعی برای داده آزمایشی در شکل ۹ نشان داده شده است. این مدل بیان‌گر آن است که مدل شبکه عصبی توانسته است ۸۲٫۴ درصد از تغییرات مقدار بازایی را توجیه کند. در نتیجه، برای پیش‌بینی میزان بازایی در ذخیره مس میدوک، مجموعه داده مدل بلوکی عیاری این ذخیره شامل مقادیر عیار  $Cu$ ،  $CuO$  و  $CuS$  به‌عنوان ورودی به مدل شبکه عصبی، وارد می‌شوند و بازایی تمام بلوک‌های واقع در مدل بلوکی ذخیره، براساس متغیرهای عیاری محاسبه می‌شود. نتایج برآورد بازایی ذخیره به روش رگرسیون و شبکه عصبی در جدول ۲ نشان داده شده است. در این



شکل ۱۰: مقطعی از مدل بلوکی کلاسه‌بندی شده با روش شبکه عصبی

جدول ۲: نتایج برآورد بازیابی ذخیره به روش شبکه عصبی و رگرسیون

روش	$R \leq 86.7$	$86.7 < R < 89.2$	$89.2 \geq R$	مجموع	میانگین کل بازیابی
رگرسیون	89.54	76.73	66.69	23296	87.9
شبکه عصبی مصنوعی	97.29	80.67	55.00	23296	85.8
مشترک در هر دو روش	78.34	70.09	52.12	20055	

### ۵- بحث

نتایج برآورد بازیابی به صورت کیفی نشان می دهد که روش درخت تصمیم با توجه به سطح دقت محاسبه شده، دارای بیشترین سطح اطمینان است. بنابراین از بین سه روش بررسی شده در این پژوهش، به عنوان قابل اعتمادترین روش در تخمین بازیابی به صورت کیفی و با استفاده از متغیرهای کلامی انتخاب می شود. این برآورد کیفی برای تخمین ها و پیش بینی های اولیه از میزان استحصال فلز در استخراج کل ذخیره مورد استفاده قرار می گیرند و به دلیل ماهیت کیفی شان، به سادگی قابل بیان و درک هستند. برای پیش بینی دقیق تر، دو مدل شبکه عصبی و مدل رگرسیونی ساخته شده و این دو روش بر اساس سه شاخص ضریب تعیین  $R^2$ ، MAE و RMSE بررسی می شوند. مدلی که دارای ضریب تعیین بیشتر و MAE و RSME کمتری باشد، دارای دقت بیشتر و قابل اعتمادتر خواهد بود. بنابراین با توجه به مقادیر  $R^2$ ، MAE و RMSE که برای مدل شبکه عصبی به ترتیب برابر  $0.8239$ ،  $0.15753$  و  $0.21404$  و برای مدل رگرسیونی برابر با  $0.7701$ ،  $0.21702$  و  $0.24972$  است، شبکه عصبی مصنوعی نسبت به روش رگرسیونی در پیش بینی بازیابی فلز محصول نهایی کارخانه فرآوری، دقت بالاتری را نشان می دهد. دلیل این امر، در نظر گرفتن روابط غیرخطی بین شاخص های ورودی و خروجی در مدل شبکه عصبی است. در ادامه بررسی معیارهای ارزیابی، شایان ذکر است که معیار MAE برای مدل حاصل از شبکه عصبی مصنوعی که دقیقاً میزان خطای مدل را نشان می دهد، مقدار  $1.57$  درصد به دست آمد و می توان این گونه تفسیر کرد به طور میانگین، بین مقدار پیش بینی شده توسط مدل شبکه عصبی و مقادیر واقعی،  $1.57$  درصد خطا در تخمین مقدار بازیابی وجود دارد؛ یعنی با وارد کردن متغیرهای ورودی بلوک ها شامل عیارهای Cu، CuO و CuS هر بلوک به مدل پیش بین، مقدار بازیابی تمام بلوک های موجود در پیت با میانگین خطای  $\pm 1.57$  درصد،

ورودی در متغیر وابسته با حذف هر یک از شاخص های ورودی، فرآیند آموزش برای متغیرهای مستقل باقیمانده و متغیر وابسته انجام و تغییرات مقدار خطای کل شبکه بررسی می شود. بنابراین طبق رابطه ۹، ضریب حساسیت نسبی برای هر شاخص محاسبه می شود:

$$RSC_k = \frac{e_k}{E} \quad (9)$$

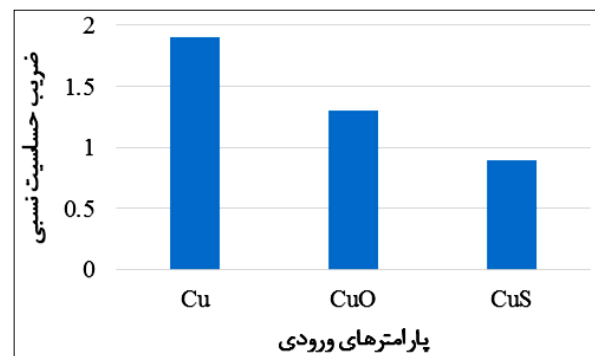
که در این رابطه:

$RSC_k$ : ضریب حساسیت نسبی<sup>۳۳</sup>

$e_k$ : خطای شبکه در غیاب متغیر ورودی  $k$

$E$ : خطای کل شبکه در حضور تمامی متغیرهای ورودی است.

به این ترتیب متغیری که ضریب حساسیت نسبی آن بالاتر باشد، تاثیر بیشتری بر متغیر وابسته دارد. نتایج آنالیز حساسیت در شکل ۱۱، نشان می دهد عیار Cu با بیشترین ضریب حساسیت، مهم ترین عامل در تغییرات بازیابی است و عیارهای CuO و CuS به ترتیب در رده های بعدی هستند.



شکل ۱۱: نتایج آنالیز حساسیت بازیابی بر اساس روش شبکه عصبی

بازیابی با تغییر ترکیب عناصر تشکیل دهنده کانسنگ، تغییر می‌کند. از این رو می‌توان به‌طور تقریبی، میزان بازیابی را با توجه به مقادیر عیاری عناصر کانسنگ تخمین زد.

در این پژوهش، به دلیل اهمیت بالای این موضوع تعدادی از نمونه‌های آنالیز شده خوراک ورودی کارخانه معدن مس میدوک، مورد مطالعه قرار گرفت و برای اولین بار، میزان بازیابی برای کل ذخیره به‌صورت کیفی با سه روش طبقه‌بندی درخت تصمیم، تئوری بیز و نزدیک‌ترین همسایه و به‌صورت کمی با دو روش شبکه عصبی مصنوعی و رگرسیون چندمتغیره برآورد شد. در تمامی این روش‌ها، شاخص‌های ورودی مقادیر عیار Cu، CuO و CuS و متغیر خروجی نیز میزان بازیابی است.

در مدل‌های کیفی، روش درخت تصمیم دارای بیش‌ترین سطح اطمینان بود و نتایج مدل‌های کمی، حاکی از دقت بالاتر روش شبکه عصبی مصنوعی در پیش‌بینی میزان بازیابی داشت که طبق آن می‌توان با داشتن مقادیر عیاری Cu، CuO و CuS برای هر بلوک واقع در ذخیره، میزان بازیابی آن را پیش‌بینی کرد و بدین ترتیب پیش‌بینی واقع‌بینانه‌تری از فلز استحصال شده از هر بلوک کانسنگ ارسالی به کارخانه فرآوری به دست‌آورد. وارد کردن معیارهای دیگر مانند نسبتی از مقادیر شاخص‌های ورودی و ساختن مدل بر اساس آنها و بررسی مدل با تعداد آنالیزهای بیش‌تر برای توسعه پژوهش در پژوهش‌های آتی پیشنهاد می‌شود. همچنین استفاده از الگوریتم شبکه عصبی مصنوعی در ترکیب با الگوریتم‌های فراابتکاری، می‌تواند به‌عنوان چشم‌اندازهای توسعه تحقیق مدنظر قرار گیرد.

## ۷- مراجع

- [1] Alonso, S. G., De La Torre-Diez, I., Hamrioui, S., López-Coronado, M., Barreno, D. C., Nozaleda, L. M., and Franco, M. (2018). "Data mining algorithms and techniques in mental health: a systematic review". *Journal of Medical Systems*, 42(9): 161.
- [2] Injadat, M., Salo, F., and Nassif, A. B. (2016). "Data mining techniques in social media: A survey". *Neurocomputing*, 214: 654-670.
- [3] Gandhi, N., and Armstrong, L. J. (2016). "A review of the application of data mining techniques for decision making in agriculture". In 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), IEEE, 1-6.
- [4] Almonacid, F., Fernandez, E. F., Mellit, A., and Kalogirou,

قابل پیش‌بینی است. این میزان دقت در پیش‌بینی با توجه به دامنه تغییرات بازیابی‌ها [۷۵٫۵-۹۴٫۵] در ۵۸ نمونه مورد مطالعه، قابل قبول است. با توجه به این‌که عیارهای Cu، CuO و CuS برای تمام مدل بلوکی ذخیره در دسترس بوده و پیش‌تر با استفاده از روش‌های زمین‌آماري محاسبه شده است، میزان بازیابی تمامی بلوک‌های ذخیره به دو روش شبکه عصبی و رگرسیون برآورد شد. با توجه به جدول ۲، محدوده بازیابی تعداد ۲۰۰۵۵ بلوک از ۲۳۲۹۶ بلوک موجود در ذخیره (۸۶ درصد کل ذخیره)، توسط هر دو روش به‌طور یکسان برآورد شد. عدم اختلاف فاحش در بازیابی برآوردشده یک بلوک با دو روش، نشان می‌دهد که می‌توان با اعتماد و قطعیت بیش‌تری در مورد بازیابی واقعی بلوک مورد نظر قضاوت کرد. به‌این ترتیب، می‌توان بازیابی هر بلوک را ابتدا با روش رگرسیون محاسبه و سپس با میزان بازیابی به‌دست‌آمده از روش شبکه عصبی مقایسه کرد. بدیهی است که برای ۳۲۴۱ بلوک باقی‌مانده که برای آنها تفاوت زیادی بین تخمین حاصل از دو روش وجود دارد، بازیابی حاصل از شبکه عصبی، دقیق‌تر و قابل اعتمادتر است. بنابراین می‌توان از ابزار پیش‌بینی‌کننده کنونی در مقابل سایر روش‌ها مانند روش‌های آزمایشگاهی پرهزینه و غیرعملی استفاده و تغییرات بازیابی در محصول نهایی را کنترل کرد و برنامه‌ریزی‌ها و تمهیدات لازم را در مقابل نوسانات بازیابی به‌کار بست. علاوه بر این، با بررسی آنالیز حساسیت صورت‌گرفته بین معیارهای ورودی مدل شبکه عصبی مشخص می‌شود که میزان بازیابی، حساسیت زیادی روی عیار Cu خوراک ورودی نشان می‌دهد. دیگر شاخص‌های مهم و موثر بر میزان بازیابی، به ترتیب CuO و CuS هستند. این بدین معنا که در غیاب معیار Cu و پس از آن، معیار CuO که ضریب حساسیت نسبی مربوط به آنها بیش‌تر از یک است، خطای مدل تخمین ساخته‌شده، افزایش یافته و بنابراین، مقدار Cu و پس از آن، مقدار CuO، بیش‌ترین نقش را در تخمین بازیابی ایفا می‌کنند. پیش‌تر مدل رگرسیونی نیز با حذف شاخص CuS در فرآیند ساخت مدل، تاثیر نه‌چندان محسوس این معیار را در تغییرات بازیابی پیش‌بینی کرده بود.

## ۶- نتیجه‌گیری

میزان بازیابی واقعی فلز در محصول نهایی کارخانه فرآوری، عامل مهمی است که می‌تواند درصد دست‌یابی به اهداف اقتصادی یک پروژه معدنی را تحت تاثیر قرار دهد. میزان

- Sciences, 72(10): 3915-3928.
- [16] Singh, T. N. (2004). "Artificial neural network approach for prediction and control of ground vibrations in mines". Mining Technology, 113(4): 251-256.
- [17] Singh, T. N., and Singh, V. (2005). "An intelligent approach to prediction and control ground vibration in mines". Geotechnical & Geological Engineering, 23(3): 249-262.
- [18] Yagiz, S., Gokceoglu, C., Sezer, E., and Iplikci, S. (2009). "Application of two non-linear prediction tools to the estimation of tunnel boring machine performance". Engineering Applications of Artificial Intelligence, 22(4-5): 808-814.
- [19] Sharma, L. K., Singh, R., Umrao, R. K., Sharma, K. M., and Singh, T. N. (2017). "Evaluating the modulus of elasticity of soil using soft computing system". Engineering with Computers, 33(3): 497-507.
- [20] Armaghani, D.J., Mohamad, E.T., Hajihassani, M., Yagiz, S., and Motaghedi, H. (2016). "Application of several non-linear prediction tools for estimating uniaxial compressive strength of granitic rocks and comparison of their performances". Engineering with Computers, 32(2): 189-206.
- [21] Saghatforoush, A., Monjezi, M., Faradonbeh, R. S., and Armaghani, D. J. (2016). "Combination of neural network and ant colony optimization algorithms for prediction and optimization of flyrock and back-break induced by blasting". Engineering with Computers, 32(2): 255-266.
- [22] Torabi-Kaveh, M., Naseri, F., Saneie, S., and Sarshari, B. (2015). "Application of artificial neural networks and multivariate statistics to predict UCS and E using physical properties of Asmari limestones". Arabian journal of Geosciences, 8(5): 2889-2897.
- [23] Taheri, K., Hasanipanah, M., Golzar, S. B., and Majid, M. Z. A. (2017). "A hybrid artificial bee colony algorithm-artificial neural network for forecasting the blast-produced ground vibration". Engineering with Computers, 33(3): 689-700.
- [24] Hasanipanah, M., Faradonbeh, R. S., Amnieh, H. B., Armaghani, D. J., and Monjezi, M. (2017). "Forecasting blast-induced ground vibration developing a CART model". Engineering with Computers, 33(2): 307-316.
- [25] Amiri, M., Amnieh, H. B., Hasanipanah, M., and Khanli, L. M. (2016). "A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure". Engineering with Computers, 32(4): 631-644.
- [26] Siami-Irdemoosa, E., and Dindarloo, S. R. (2015). S. (2017). "Review of techniques based on artificial neural networks for the electrical characterization of concentrator photovoltaic technology". Renewable and Sustainable Energy Reviews, 75: 938-953.
- [5] Li, H., Zhang, Z., and Liu, Z. (2017). "Application of artificial neural networks for catalysis: A review". Catalysts, 7(10): 306.
- [6] Ghaedi, A. M., and Vafaei, A. (2017). "Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: a review". Advances in Colloid and Interface Science, 245: 20-39.
- [7] Mellit, A., and Kalogirou, S. A. (2018). "A Survey on the Application of Artificial Intelligence Techniques for Photovoltaic Systems". In McEvoy's Handbook of Photovoltaics, Academic Press, 735-761.
- [8] Tkáč, M., and Verner, R. (2016). "Artificial neural networks in business: Two decades of research". Applied Soft Computing, 38: 788-804.
- [9] Ebrahimi, E., Monjezi, M., Khalesi, M. R., and Armaghani, D. J. (2016). "Prediction and optimization of back-break and rock fragmentation using an artificial neural network and a bee colony algorithm". Bulletin of Engineering Geology and the Environment, 75(1): 27-36.
- [10] Faradonbeh, R. S., Monjezi, M., and Armaghani, D. J. (2016). "Genetic programming and non-linear multiple regression techniques to predict backbreak in blasting operation". Engineering with Computers, 32(1): 123-133.
- [11] Khandelwal, M., and Singh, T. N. (2006). "Prediction of blast induced ground vibrations and frequency in opencast mine: a neural network approach". Journal of Sound and Vibration, 289(4-5): 711-725.
- [12] Meulenkamp, F., and Grima, M. A. (1999). "Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness". International Journal of Rock Mechanics and Mining Sciences, 36(1): 29-39.
- [13] Khandelwal, M., Kumar, D. L., and Yellishetty, M. (2011). "Application of soft computing to predict blast-induced ground vibration". Engineering with Computers, 27(2): 117-125.
- [14] Mohamed, M. T. (2011). "Performance of fuzzy logic and artificial neural network in prediction of ground and air vibrations". International Journal of Rock Mechanics and Mining Sciences, 48(5): 845.
- [15] Vasović, D., Kostić, S., Ravilić, M., and Trajković, S. (2014). "Environmental impact of blasting at Drenovac limestone quarry (Serbia)". Environmental Earth

Applications, 28(1): 1043-1050.

[40] Myers, R. H., and Myers, R. H. (1990). "Classical and modern regression with applications". Duxbury Press, 2<sup>nd</sup> Edition, pp. 488.

[41] Pilevar, S. A., Ayoubi, S. H., and Khademi, H. (2011). "Comparison of artificial neural network (ANN) and multivariate linear regression (MLR) models to predict soil organic carbon using digital terrain analysis (Case Study: Zargham Abad Semrom, Isfahan Province)".

"Prediction of fuel consumption of mining dump trucks: A neural networks approach". Applied Energy, 151: 77-84.

[27] Han, J., Pei, J., and Kamber, M. (2011). "Data mining: concepts and techniques". Morgan Kaufmann Publishers, pp. 744.

[28] Friedl, M. A., and Brodley, C. E. (1997). "Decision tree classification of land cover from remotely sensed data". Remote Sensing of Environment, 61(3): 399-409.

[29] Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). "Classification and regression trees". Belmont, CA: Wadsworth, International Group, pp. 432.

[30] Ye, N. (2003). "The handbook of Data Mining". Lawrence Erlbaum Associates Publishers, pp. 689.

[31] Theodoridis, S., Pikrakis, A., Koutroumbas, K., and Cavouras, D. (2010). "Introduction to pattern recognition: a matlab approach". 1<sup>st</sup> Edition, Academic Press, pp. 231.

[32] Duda, R. O., Hart, P. E., and Stork, D. G. (2001). "Pattern classification". 2<sup>nd</sup> Edition, John Wiley & Sons, pp. 688.

[33] McCulloch, W. S., and Pitts, W. (1943). "A logical calculus of the ideas immanent in nervous activity". The Bulletin of Mathematical Biophysics, 5(4): 115-133.

[34] Haykin, S. (1994). "Neural networks: A comprehensive foundation". 2<sup>nd</sup> Edition, Publisher: Prentice Hall, pp. 842.

[35] Rumelhart, D. E. (1986). "Parallel distributed processing: Explorations in the microstructure of cognition". Learning Internal Representations by Error Propagation, 1: 318-362.

[36] MathWorks Inc., MATLAB 7.4 R2016a (2016). Software, MathWorks, Inc.

[37] Hasanipanah, M., Armaghani, D. J., Monjezi, M., and Shams, S. (2016). "Risk assessment and prediction of rock fragmentation produced by blasting operation: a rock engineering system". Environmental Earth Sciences, 75(9): 808.

[38] Hasanipanah, M., Naderi, R., Kashir, J., Noorani, S. A., and Qaleh, A. Z. A. (2017). "Prediction of blast-produced ground vibration using particle swarm optimization". Engineering with Computers, 33(2): 173-179.

[39] Hasanipanah, M., Armaghani, D. J., Amnieh, H. B., Majid, M. Z. A., and Tahir, M. M. (2017). "Application of PSO to develop a powerful equation for prediction of flyrock due to blasting". Neural Computing and

<sup>1</sup> Classification

<sup>2</sup> Data mining

<sup>3</sup> Decision tree

<sup>4</sup> Bayes theory

<sup>5</sup> K-Nearest Neighbor (KNN)

<sup>6</sup> Artificial Neural Network (ANN)

<sup>7</sup> Adaptive Neuro-Fuzzy Inference System (ANFIS)

<sup>8</sup> Uniaxial Compressive Strength

<sup>9</sup> Coefficient of determination

<sup>10</sup> Root Mean Squared Error (RMSE)

<sup>11</sup> Peak Particle Velocity (PPV)

<sup>12</sup> Mean Absolute Error (MAE)

<sup>13</sup> Coefficient of correlation

<sup>14</sup> Classification And Regression Tree (CART)

<sup>15</sup> Artificial neural network and k-nearest neighbors

<sup>16</sup> Supervised

<sup>17</sup> Training data

<sup>18</sup> Learning

<sup>19</sup> Parent

<sup>20</sup> Children

<sup>21</sup> White box

<sup>22</sup> Attribute

<sup>23</sup> Conditional Independence

<sup>24</sup> McCulloch

<sup>25</sup> Pitts

<sup>26</sup> Neuron

<sup>27</sup> Multi-Layer Perceptron

<sup>28</sup> Bias

<sup>29</sup> Overfitting

<sup>30</sup> Vlidation fail

<sup>31</sup> Underfitting

<sup>32</sup> Statsoft

<sup>33</sup> Relative sensitivity coefficient



DOI:10.30479/jmre.2019.10997.1284

## Comparison of Artificial Neural Networks and Multivariate Linear Regression Classification Techniques in Metal Recovery Estimation

Gholamnejad J.<sup>1\*</sup>, Lotfian R.<sup>2</sup>, Mirzaeian Lord Keivan Y.<sup>3</sup>

- 1- Associate Professor, Dept. of Mining and Metallurgical Engineering, Yazd university, Yazd, Iran  
j.gholamnejad@yazd.ac.ir
- 2- Ph.D. Student, Dept. of Mining and Metallurgical Engineering, Yazd university, Yazd, Iran  
reza.lotfian11@gmail.com
- 3- Assistant Professor, Dept. of Mining and Metallurgical Engineering, Yazd university, Yazd, Iran  
mirzaeian@yazd.ac.ir

(Received: 16 Jun. 2019, Accepted: 11 Nov. 2019)

**Abstract:** Due to the role of recovery in calculating the economic value of ore blocks and the impact of the block's economic value on the design calculations of the final pit and production planning, determination of the amount of metal recovery from the ore material sent to the processing plant is very important. The aim of this study is to investigate the capability of estimating the recovery rate of ore in qualitative manner with three methods based on data classification from data mining techniques and quantitatively using multivariate regression and artificial neural networks. Hence, the Miduk copper mine was studied using 58 analyzed samples of the feed of the plant, including Cu, CuO and CuS grades, and the recovery rate of Cu in the final product of the plant. The process of predicting the total recovery of the reserve was made qualitatively by decision tree method, classification based on Bayes rule and k-nearest neighbor (kNN) classification algorithm. For quantitative estimation of recovery, multivariate regression and artificial neural network models were established between the mentioned grade parameters and recovery rates (For 47 samples of 58 samples) and with the 11 additional analyzed samples, the obtained models were validated. The coefficient of (R<sup>2</sup>), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) in the regression model were 0.77, 0.027722 and 0.029722, respectively, and in the artificial neural network model, 0.82, 0.015753 and 0.024040, respectively. Therefore, the artificial neural networks model acts as a more accurate tool for predicting recovery versus the multivariable regression model. The results of sensitivity analysis of artificial neural network model showed that Cu grade is the most important factor and grade of CuO and CuS, respectively, as well as other factors influencing the changes in recovery rate.

**Keywords:** Recovery, Classification, Multivariate regression, Artificial neural network.

### INTRODUCTION

The economic block model is the basis of designing the final pit limit and production planning in mines. The economic block model is calculated by taking into account economic parameters such as revenues



and cost of metal extraction for each block and assigning a value to it. The amount of metal recovery from each ore block sent to the processing plant plays an important role in calculating its economic value. In most of the calculations of the block economic value, the recovery factor is defined as the efficiency of the processing plant or the smelting and refining plant in terms of recovering the metal from the input feed. This nominal recovery is determined according to the specific settings of the processing plant. But in some cases, such as the case studies investigated in this study, given the Specific ratios of the elements forming each block, the actual recovery rate in practice is different from its nominal value. In this case, the amount of recovered metal expected and the expected profits from the product sale will not be realized, and the calculations of the economic value of the block, followed by the final pit and the process of production planning, will change. Assuming the same plant conditions in input feed processing, one of the important factors that can affect the recovery rate is the specific ratio and hidden relationship between the set of elements in the feed. Now, if we can provide a model of this hidden relationship and predict the recovery rate, we will have more realistic values of the economic value of the ore block sent to the plant, followed by more reliable the final pit and production planning. In this study, both qualitative and quantitative, the prediction models are provided for the recovery rate of a mineral deposit. The qualitative recovery estimation is based on the concept of classification from a set of data mining algorithms and with three methods of decision tree, Bayes theory and k-nearest neighbor. Artificial neural network approach and multivariate regression are also used to estimate the recovery rate of ore block. Data mining techniques such as classification algorithms [1-3] and the artificial neural network approach have been widely used in various sciences [4-8]. In recent years, artificial neural network approach has been used to solve mining and geotechnical engineering problems such as TBM performance, predicting undesirable impacts of blasting such as backbreak, flyrock, ground vibration and airblast [9-15]. The purpose of this research is to first create a model for the prediction of recovery of the Midok copper ore, due to the variation in the grade of its content elements.

## METHODS

### Qualitative estimation of recovery using the concept of classification

The concept of classification in the field of data mining is a supervisory approach that seeks to discover the relationship between independent variables and an associated variable. This relationship is represented by a structure as a model. Classification builds up and utilizes a model to predict the categorical labels of unknown objects to distinguish between objects of different classes. These categorical labels are predefined, discrete and unordered. Classification and prediction is the process of identifying a set of common features and models that describe and distinguish data classes or concepts. Common classification techniques include the naive Bayes technique, decision trees and support vector machines [16]. In this research, three methods including decision tree, Bayes theory and k-nearest neighbor are used to estimate the recovery based on the Cu, CuO and CuS grades of the feed of the plant.

### Recovery estimation using multivariate regression and artificial neural network Methods

#### Artificial neural network (ANN) method

One of the subsystems of AI systems is an ANN. The ANN model has been developed since the 1960s. Generally, the structure of an ANN, which is inspired by the human brain, consists of a group of computational units called neurons or nodes. These neurons are highly interconnected with each other. A typical ANN consists of three layers, namely input, hidden and output layers. The mentioned neurons are placed in these layers and linked to each other by weights. On the other hand, problem effective and objective variables are placed in the input and output layers, respectively. Theoretically, there are no restrictions on the No. of hidden layers and No. of neurons in the hidden layers and can be determined based on trial and error procedure. To construct an ANN model, in the first step, ANNs require training to learn and consequently map a relationship from the data. There are many algorithms to train the network, such as Levenberg–Marquardt (LM), conjugate gradient and scaled conjugate gradient algorithms. In the second step, to check the performance capacity of the constructed model, the rest of datasets are used for testing [17].

### Multivariate regression method

Multiple linear regression (MLR) is one of the most common statistical tools to fit a linear equation between two or more independent variables and a dependent variable. This model is extensively used for solving different engineering problems by many researchers.

## FINDING AND ARGUMENT

### Qualitative recovery estimation

In order to predict the recovery rate in the Midok copper mine, the block model of the deposit is entered into the decision tree, Bayesian and k-nearest neighbor models constructed in the training stage. In this way, all blocks in the block model are classified in the three levels of recovery (high, medium and low). Figure 1 shows a cross-sectional view of a classified block model by decision tree method with the highest confidence level.

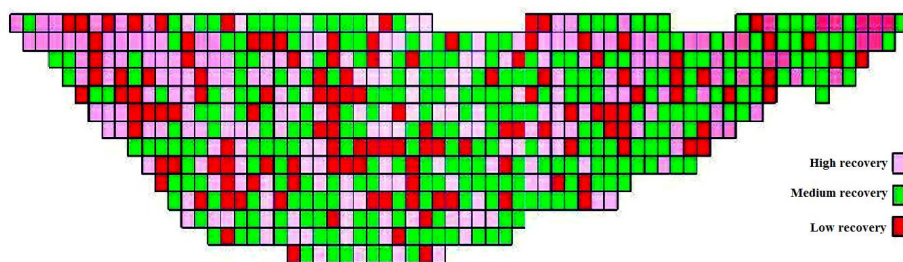


Figure 1. A cross-sectional view of a classified block model by decision tree method

### Quantitative recovery estimation

#### Multivariate regression results

Relation 1 shows the model of multivariate regression method for predicting recovery rate. Figure 2 also shows the relationship between the measured recovery and the predicted recovery in the multivariate regression model for the 11 additional samples.

$$R = 0.851 - 1.725CuO + 0.154Cu \quad (1)$$

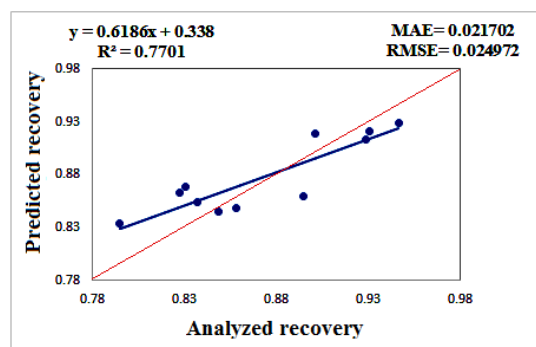


Figure 2. Relationship between the measured and the predicted recovery in the multivariate regression model

### Artificial neural network modeling results

The relationship between the measured recovery and the predicted recovery in the artificial neural network model for the 11 additional samples is shown in Figure 3. Figure 4 shows a cross-sectional view of classified block model by artificial neural network. The results of estimating recovery of the deposit by regression and neural network methods are shown in Table 1. The results of the statsoft sensitivity analysis are shown in Figure 5. Thus, the Cu-grade parameter with the highest sensitivity coefficient is the most important factor in recovery changes, and the CuO and CuS grades are respectively in the following ranks.

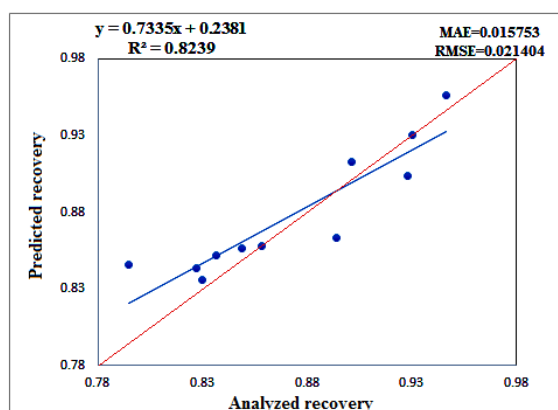


Figure 3. Relationship between the measured and the predicted recovery in the ANN model

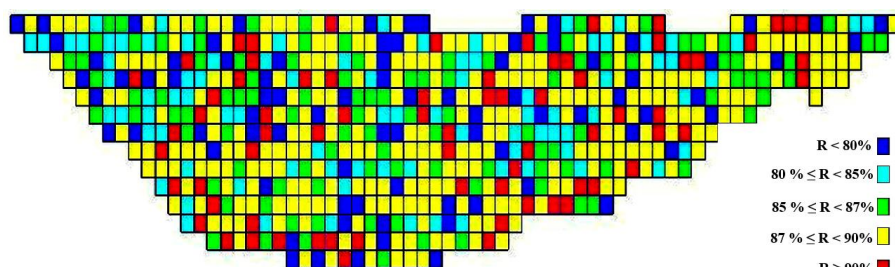


Figure 4. A cross-sectional view of classified block model by artificial neural network

Table 1. Results of estimating recovery of the deposit by regression and neural network methods

Method	$R \leq 86.7$	$86.7 < R < 89.2$	$R \geq 89.2$	Sum
MLR	8954	7673	6669	23296
ANN	9729	8067	5500	23296
Common in both methods	7834	7009	5212	20055

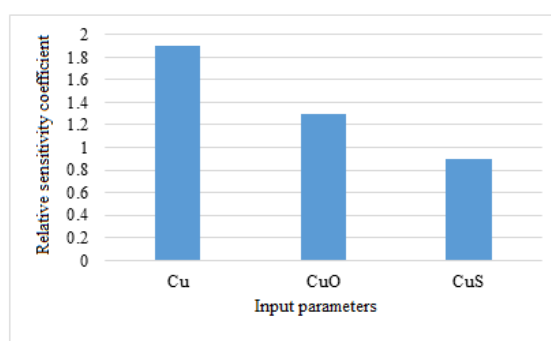


Figure 5. Results of the statsoft sensitivity analysis

## CONCLUSION

The actual recovery rate of the metal in the final product of the processing plant is an important factor that can affect the achievement of the economic goals of a mineral project. The recovery rate changes by changing the composition (ratio) of the ore forming elements. Therefore, it can be estimated the recovery rate by the grades of the ore block. In this research, due to the importance of this issue, a number of analyzed feed samples from the Midok copper mine plant were studied and for the first time, the recovery

rate for total deposit was qualitatively predicted by three methods of decision tree, Bayes theory and nearest neighbor and quantitatively with two methods of artificial neural network and multivariate regression. Input parameters in all of these methods are Cu, CuO, CuS grades, and recovery is output variable. In the quality models, decision tree has the highest level of confidence. The results of quantitative models indicate a higher accuracy of artificial neural network method in predicting recovery rates according to which it is possible to predict the recovery rate for each block of deposit based on CuO, CuO, and CuS grades. It is recommended to enter other parameters such as a ratio of input parameter values and construct a model based on them, and suggest a model with a number of further analyzes for the development of research in future work. Also, the use of artificial neural network algorithm in combination with metaheuristic algorithms can be considered as research development perspectives.

## REFERENCES

- [1] Alonso, S. G., De La Torre-Diez, I., Hamrioui, S., López-Coronado, M., Barreno, D. C., Nozaleda, L. M., and Franco, M. (2018). "Data mining algorithms and techniques in mental health: a systematic review". *Journal of medical systems*, 42(9): 161.
- [2] Injadat, M., Salo, F., and Nassif, A. B. (2016). "Data mining techniques in social media: A survey". *Neurocomputing*, 214: 654-670.
- [3] Gandhi, N., and Armstrong, L. J. (2016). "A review of the application of data mining techniques for decision making in agriculture". In *Contemporary Computing and Informatics (IC3I)*, 2nd International Conference on IEEE, 1-6.
- [4] Almonacid, F., Fernandez, E. F., Mellit, A., and Kalogirou, S. (2017). "Review of techniques based on artificial neural networks for the electrical characterization of concentrator photovoltaic technology". *Renewable and Sustainable Energy Reviews*, 75: 938-953.
- [5] Li, H., Zhang, Z., and Liu, Z. (2017). "Application of artificial neural networks for catalysis: a review". *Catalysts*, 7(10): 306.
- [6] Ghaedi, A. M., and Vafaei, A. (2017). "Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: a review". *Advances in Colloid and Interface Science*, 245: 20-39.
- [7] Mellit, A., and Kalogirou, S. A. (2018). "A Survey on the Application of Artificial Intelligence Techniques for Photovoltaic Systems". In *McEvoy's Handbook of Photovoltaics*, Academic Press, 735-761.
- [8] Tkáč, M., and Verner, R. (2016). "Artificial neural networks in business: Two decades of research". *Applied Soft Computing*, 38: 788-804.
- [9] Ebrahimi, E., Monjezi, M., Khalesi, M. R. and Armaghani, D. J. (2016). "Prediction and optimization of back-break and rock fragmentation using an artificial neural network and a bee colony algorithm". *Bulletin of Engineering Geology and the Environment*, 75(1): 27-36.
- [10] Faradonbeh, R. S., Monjezi, M., and Armaghani, D. J. (2016). "Genetic programming and non-linear multiple regression techniques to predict backbreak in blasting operation". *Engineering with Computers*, 32(1): 123-133.
- [11] Saghatforoush, A., Monjezi, M., Faradonbeh, R. S., and Armaghani, D. J. (2016). "Combination of neural network and ant colony optimization algorithms for prediction and optimization of flyrock and back-break induced by blasting". *Engineering with Computers*, 32(2): 255-266.
- [12] Armaghani, D. J., Mohamad, E. T., Hajihassani, M., Yagiz, S., and Motaghedi, H. (2016). "Application of several non-linear prediction tools for estimating uniaxial compressive strength of granitic rocks and comparison of their performances". *Engineering with Computers*, 32(2): 189-206.
- [13] Hasanipanah, M., Faradonbeh, R. S., Amnieh, H. B., Armaghani, D. J., and Monjezi, M. (2017). "Forecasting blast-induced ground vibration developing a CART model". *Engineering with Computers*, 33(2): 307-316.
- [14] Yagiz, S., Gokceoglu, C., Sezer, E., and Iplikci, S. (2009). "Application of two non-linear prediction tools to the estimation of tunnel boring machine performance". *Engineering Applications of Artificial Intelligence*, 22(4-5): 808-814.

- [15] Sharma, L. K., Singh, R., Umrao, R. K., Sharma, K. M., and Singh, T. N. (2017). "Evaluating the modulus of elasticity of soil using soft computing system". *Engineering with Computers*, 33(3): 497-507.
- [16] Han, J., Pei, J., and Kamber, M. (2011). "Data mining: concepts and techniques". Elsevier.
- [17] Hasanipanah, M., Noorian-Bidgoli, M., Armaghani, D. J., and Khamesi, H. (2016). "Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling". *Engineering with Computers*, 32(4): 705-715.