نقشه‌برداری رقومی کلاس‌های خاک با استفاده از انواع روش‌های داده‌کاوی در منطقه‌ی اردکان استان یزد

روح‌اله تقی‌زاده‌مهرجردی[1*]، فریدون سرمدیان[2]، محمود امید[3]، نورایر تومانیان[4]، محمدجواد روستا[5]، محمد حسن رحیمیان[6]

1*- **نویسنده مسؤول:** استادیار دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان، یزد، ایران (rtaghizadeh@ardakan.ac.ir)

2- استاد گروه خاکشناسی، دانشکده مهندسی و فناوری کشارزی، دانشگاه تهران، کرج، البرز، ایران

3- استاد گروه ماشین‌های کشاورزی، دانشکده مهندسی و فناوری کشارزی، دانشگاه تهران، کرج، البرز، ایران

4- استادیار مرکز تحقیقات کشاورزی اصفهان، اصفهان، ایران

5- استادیار مرکز ملی شوری، یزد، ایران

6- مربی مرکز ملی شوری، یزد، ایران

تاریخ دریافت: 1392/12/10 تاریخ پذیرش: 1393/08/06

چکیده

در طی سال‌های گذشته پیشرفت‌های گسترده‌ای در زمینه علم نقشه‌برداری رقومی خاک شده است؛ به‌طوری‌که منجر به تولید نقشه‌های رقومی خاک در سطح ملی و قاره‌ای گردیده است؛ اما به رغم مطالعات بسیار زیادی که در سطح دنیا انجام گرفته است و یا در حال انجام شدن می‌باشد، تعداد اندکی از خاک‌شناسان ایرانی به علم نقشه‌برداری رقومی خاک علاقه نشان داده‌اند؛ لذا در تحقیق حاضر، سعی شده تا از انواع روش‌های داده‌کاوی برای پیش‌بینی مکانی گروه-های بزرگ خاک در اراضی به وسعت 72000 هکتار در منطقه‌ی اردکان بهره گرفته شود. در این منطقه بر اساس روش مربع لاتین مکان جغرافیایی 187 پروفیل خاک مشخص شده و سپس تشریح، نمونه‌برداری و بر اساس سیستم آمریکایی به شش گروه بزرگ و هشت زیرگروه طبقه‌بندی شدند. متغیرهای محیطی یا نمایندگان فاکتورهای خاک‌سازی که در این پژوهش استفاده شدند، شامل ویژگی های توپوگرافی داده‌های تصویر ETM^+ ماهواره لندست 2002 و نقشه‌های پلی‌گونی از جمله سطوح ژئومورفولوژی، زمین‌شناسی و کاربری اراضی است. نتایج نشان داد که مدل درخت تصمیم نسبت به سایر مدل‌ها دارای برتری نسبی می‌باشد؛ به‌طوری‌که توانسته است دقت پیش‌بینی گروه‌های بزرگ خاک را نسبت به روش آنالیز تشخیصی (ضعیف‌ترین مدل) 44% افزایش دهد. در کل نتایج تحقیق، دقت مدل‌های درخت تصمیم، شبکه عصبی مصنوعی، ترکیب شبکه عصبی مصنوعی - الگوریتم ژنتیک، رگرسیون لاجیستیک چندجمله‌ای و آنالیز تشخیصی را جهت پیش‌بینی گروه‌های بزرگ خاک با دقت کلی 70% 65% 65% 55% و 47% به ترتیب، تایید کرد. همچنین درخت تصمیم، کلاس‌های زیرگروه خاک را با دقت 84/2% پیش‌بینی نموده است.

کلید واژه‌ها: درخت تصمیم، آنالیز تشخیصی، رگرسیون لاجیستیک، شبکه عصبی مصنوعی، الگوریتم ژنتیک

مقدمه

همزمان با انقلاب فنآوری اطلاعات و رایانه، حجم عظیمی از اطلاعات وارد عرصه‌های مختلف علمی شده است. بدیهی است که علم خاک‌شناسی هم از این قاعده مستثنی نبوده و تشکیل پایگاه داده‌های محلی، ملی، قاره‌ای و جهانی یکی از نتایج آن است. بنابراین برای پردازش و استفاده‌ی عملی از این اطلاعات، توسعه‌ی

ابزارهای آماری مانند داده‌کاوی[1] و ماشین‌های یادگیری[2] ضروری می‌باشد (گرانوالد[3]، 2009).

خوشبختانه افزایش توانایی ابزارهایی مانند سامانه‌ی اطلاعات جغرافیایی[4]، سیستم تعیین موقعیت جهانی[5]، سیستم‌های ماهواره‌ای و سایر منابع داده‌ها مانند اطلاعات حاصل از مدل رقومی ارتفاع، همزمان با درخواست و نیاز به داده‌های خاک می‌باشد (فینک[6]، 2012)؛ در نتیجه در سطح دنیا ارگان‌ها و شرکت‌های بسیاری در حال استفاده و تحقیق در مورد نحوه‌ی مطالعه‌ی خاک هستند (مک-براتنی و همکاران[7]، 2003). به نظر می‌رسد که به تدریج تولید نقشه‌های رقومی که در مقابل نقشه‌های رقومی‌شده[8] قرار دارد، از مرحله تحقیق (مور و همکاران[9]، 1993) وارد مرحله اجرا شده‌اند. شاهد این مدعا را می‌توان تولید نقشه‌های رقومی در استرالیا (بوی و موران[10]، 2001) یا مجارستان (دوبوس و همکاران[11]، 2000) گرفت. علاوه بر این در کشور هلند نقشه‌برداری با بزرگنمایی مکانی 100 متر برای تمام کشور به اتمام رسیده است. در مقابل، کشور فرانسه با توسعه‌ی اقتصادی بهتر، تنها 0/26 کشور با بزرگنمایی مکانی 500 متر را توانسته پوشش بدهد. این وضعیت برای کشورهای بزرگی مانند استرالیا و برزیل بدتر
می‌باشد. تنها دلیل کمبود اطلاعات مکانی خاک در چنین کشورهایی، سرعت کم و هزینه بالای نقشه‌برداری سنتی است. البته با وجود این‌که سیستم اطلاعات جغرافیایی با داده‌های مکانی کار می‌کند (بوروق و مک‌دانل[12]، 1998)؛ اما این سیستم به خودی خود قادر به نقشه‌برداری رقومی خاک نمی‌باشد؛ لذا به مرور زمان، محققان خاک‌شناس به فکر ارایه‌ی چارچوبی ذهنی برای نقشه‌برداری رقومی کلاس و ویژگی‌های خاک افتادند (مک‌براتنی و همکاران، 2003).

مک‌براتنی و همکاران (2003)، با مروری بر مدل‌های مختلف تخمین و برآورد مکانی خاک و با توجه به نکات کلیدی مذکور اقدام به ارائه‌ی یک چارچوب تعمیم‌یافته برای تخمین و برآورد مکانی خاک نمودند؛ چارچوبی که بعدها تحت عنوان نقشه‌برداری رقومی خاک نامیده شد. آنان با داخل نمودن موقعیت مکانی و دیگر مشخصه‌های خاک، اقدام به بسط و توسعه مدل چند فاکتوری ینی، به منظور برآورد و تخمین یک مشخصه خاک یا کلاس خاک نمود. به دیگر سخن، آنان مدل پنج فاکتوری ینی را به یک مدل هفت فاکتوری از طریق کمی‌سازی و بهره‌مندی از مدل‌های رقومی ارتفاع، سنجش از دور، سنسورهای سنجنده خاک و دیگر تکنیک‌های پهنه‌بندی، بسط و تعمیم دادند (مک‌براتنی و همکاران، 2003). مدل اسکورپن[13] در حال حاضر پایه و اساس رویکرد نوین نقشه‌برداری خاک را تشکیل می‌دهد و به‌طور وسیعی در تحقیقات نقشه‌برداری رقومی خاک در حال استفاده می‌باشد (رزیتر[14]، 2004). یکی از اجزاء اصلی نقشه‌برداری رقومی خاک (اسکورپن) مدل ارتباط دهنده بین متغیرهای محیطی و خاک می‌باشد که در این راستا می-توان از انواع روش‌های داده‌کاوی بهره گرفت.

در حقیقت داده‌کاوی یک حوزه میان رشته‌ای و با رشد سریع است که حوزه‌های مختلفی همچون پایگاه داده، آمار، یادگیری ماشین و سایر زمینه‌های مرتبط را با هم تلفیق کرده است تا اطلاعات و دانش ارزشمند نهفته در حجم بزرگی از داده‌ها را استخراج نماید. بر اساس دیدگاهی دیگر، داده‌کاوی بخشی از فرآیند کشف دانش

1- Data Mining
2- Learning Machine
3- Grunwald
4- Geographic Information System (GIS)
5 Global Position System (GPS)
6- Finke
7- McBratney *et al.*
8- Digitized Soil Maps
9- Moore *et al.*
10- Bui and Moran
11- Dobos *et al.*
12- Burrough and McDonnell
13- SCORPAN
14- Rossiter *et al.*

می‌باشد. کشف دانش شامل مراحل متعددی می‌شود؛ گام اول: درک حوزه‌ی کاربرد مورد نظر و نحوه رابطه‌بندی مسئله است. این گام به وضوح پیش‌نیاز استخراج دانش مفید و انتخاب روش‌های داده‌کاوی مناسب در گام سوم، با توجه به هدف کاربرد و طبیعت داده‌هاست؛ گام دوم: جمع‌آوری و پیش‌پردازش داده شامل انتخاب منابع داده، حذف نقاط پرت و مغشوش و کاهش داده‌هاست؛ گام سوم: داده‌کاوی است که هدف آن استخراج الگوها یا مدل‌های مخفی در داده‌هاست؛ گام چهارم: تفسیر دانش کشف شده است. تجربه نشان داده است که همیشه مدل‌های کشف شده از داده‌ها، مفید و جالب نیستند؛ بنابراین داده‌کاوی یک فرآیندی تکراری می‌باشد. یک راه استاندارد ارزیابی قواعد استنتاج شده تقسیم داده‌ها به دو محدوده برای آموزش و آزمون است؛ گام پنجم: استفاده عملی از دانش کشف شده است.

انواع روش‌های داده‌کاوی که برای پیش‌بینی کلاس خاک می‌توان استفاده کرد عبارت‌اند از: رگرسیون لاجیستیک[1] (هنگل و همکاران[2]، 2007)، شبکه‌های عصبی مصنوعی[3] (دو و همکاران[4]، 2008)، طبقه‌بندی درختی یا تصمیم‌گیری درختی[5] (مینازنی و ومک‌براتنی[6]، 2007 a, b) و آنالیز تطبیقی متعارفی[7] (لیو و همکاران[8]، 2008)؛ همچنین بعضی از تحقیقات در زمینه‌ی نقشه‌برداری رقومی سعی در مقایسه انواع روش‌های داده‌کاوی جهت دست‌یافتن به روش مناسب در زمین‌نما مورد نظر دارند (گرانوالت، 2009). برای مثال، ریوروو همکاران[9] (2007) سه نوع روش زمین‌آماری؛ واسکیوس و همکاران[10] (2008) پنج روش را برای پیش‌بینی ماده آلی خاک؛ مینازنی و مک‌براتنی (b 2007) چهار روش را برای پیش‌بینی درصد رس، محصول پنبه، واکنش خاک؛ بهرنس و اشلوتن[11] (2007) ده روش داده‌کاوی را برای پیش‌بینی کلاس خاک در آلمان؛ مونجون[12] و همکاران (2010) تکنیک شبکه عصبی مصنوعی و درخت تصمیم را برای پیش‌بینی واحدهای خاک؛ لوتو و جرت[13] (2005) روش‌های رگرسیون، تصمیم‌گیری درختی و شبکه عصبی را برای نقشه‌برداری سطوح ژئومورفیک در فنلاند؛ اوده و انس[14] (2008) روش‌های رگرسیون‌کریجینگ و کریجینگ شاخص را برای بررسی تغییرات مکانی شوری و سدیمی بودن خاک و پایداری ساختمان خاک در کشور استرالیا؛ جعفری و همکاران[15] (2012) درخت تصمیم و رگرسیون لاجیستیک جهت پیش‌بینی واحدهای خاک در منطقه زرند کرمان؛ هنگل و همکاران (2007) از مدل‌های رگرسیون لاجیستیک، طبقه‌بندی نظارت شده، فواصل تاکسونومیک و رگرسیون کریجینگ را جهت تهیه نقشه رقومی کلاس‌های خاک در ایران بهره گرفتند.

این پژوهش سعی بر این دارد که در ابتدا، انواع روش‌های داده‌کاوی جهت پهنه‌بندی رقومی خاک مقایسه کرده و با استفاده از بهترین روش، نقشه کلاس‌های خاک را تا سطح گروه بزرگ در منطقه‌ی اردکان استان یزد تهیه نماید. در مرحله‌ی بعد، برای به‌دست آوردن اطلاعات جزئی‌تر، اقدام به تهیه نقشه رقومی در سطح زیرگروه خاک می‌شود.

مواد و روش‌ها

نمونه‌برداری

بر اساس روش مربع لاتین، محل 187 پروفیل در منطقه‌ی اردکان استان یزد به وسعت 72000 هکتار انتخاب شد (شکل1). این روش، یک طرح تصادفی طبقه‌بندی شده‌ای است که باعث نمونه‌برداری مؤثری به کمک توزیع چند متغیر می‌شود. یک شبکه مربع حاوی موقعیت‌های نمونه، یک

1- Logistic Regression
2- Hengl *et al*.
3- Artificial Neural Networks
4- Du *et al.*
5- Decision Tree
6- Minasny&McBratney
7- Canonical Correspondence Analysis (CCA)
8- Liu *et al*.
9- Rivero *et al.*
10- Vasques *et al.*
11- Behrens & Scholten
12- Moonjun
13- Luoto and Hjort
14- Odeh and Onus
15 - Jafari et al.

شبکه لاتین است اگر و تنها اگر فقط یک واحد نمونه‌برداری در هر ردیف و هر ستون وجود داشته باشد. یک هایپرکیوب لاتین، تعمیم این مفهوم به تعداد اختیاری از ابعاد است، به‌طوری‌که هر واحد نمونه‌برداری فقط در یک صفحه آن قرار گرفته باشد. LHS شامل نمونه‌برداری n مقدار از توزیع تشریح شده هر متغیر می‌باشد. توزیع تجمعی هر متغیر به n فاصله با احتمال مساوی تقسیم می‌شود و یک مقدار از هر فاصله به‌طور تصادفی انتخاب می‌گردد. سپس، n مقدار به‌دست آمده برای هر متغیر با متغیرهای دیگر جفت می‌شود. این روش، پوشش کاملی از هر متغیر را ایجاد می‌کند (مینازنی و مک‌براتنی، 2006). از لایه‌های ژنتیکی تمامی مقاطع، نمونه‌برداری صورت گرفت و برای تجزیه به آزمایشگاه منتقل شدند. تعدادی از ویژگی‌های فیزیکی و شیمیایی نمونه‌ها از جمله هدایت الکتریکی ظاهری، ماده آلی، بافت خاک، کاتیون و آنیون‌ها، گچ و آهک اندازه‌گیری شدند و در نهایت بر اساس سیتم طبقه‌بندی آمریکا، خاک‌ها به شش گروه بزرگ و هشت زیرگروه طبقه‌بندی شدند (گروه شناسایی خاک[1]، 2010). همچنین در مطالعه حاضر، از انواع متغیرهای پیش‌بینی کننده از قبیل نقشه‌های پلی‌گونی (نقشه‌های ژئومورفولوژی، زمین‌شناسی، کاربری اراضی که شامل باغات پسته، اراضی بایر و مراتع بودند)، اجزاء توپوگرافی (شیب، ارتفاع، ارتفاع بالای شبکه زهکشی، مساحت حوزه‌های اصلاح شده، موقعیت میانی شیب، عمق دره، شاخص خیسی، شاخص صافی قسمت میانی دره با درجه تفکیک بالا، شاخص بالای پشته با درجه تفکیک بالا و شیب حوضه) و شاخص‌های حاصل از تصویر ماهواره لندست مربوط به سال 2002 (مولفه‌های اصلی اول تا سوم، شاخص گیاهی نرمال‌شده، شاخص رس، شاخص کربنات، شاخص گچ، شاخص شوری و شاخص روشنایی) استفاده گردید (شکل1). مدل رقومی ارتفاع و تصویر ماهواره مورد استفاده، دارای قدرت مکانی 10 و 30 متر می‌باشند. برای انجام فرآیند مدل‌سازی تمام لایه‌های اطلاعاتی به اندازه پیکسل 30 متر نمونه‌گیری مجدد شدند.

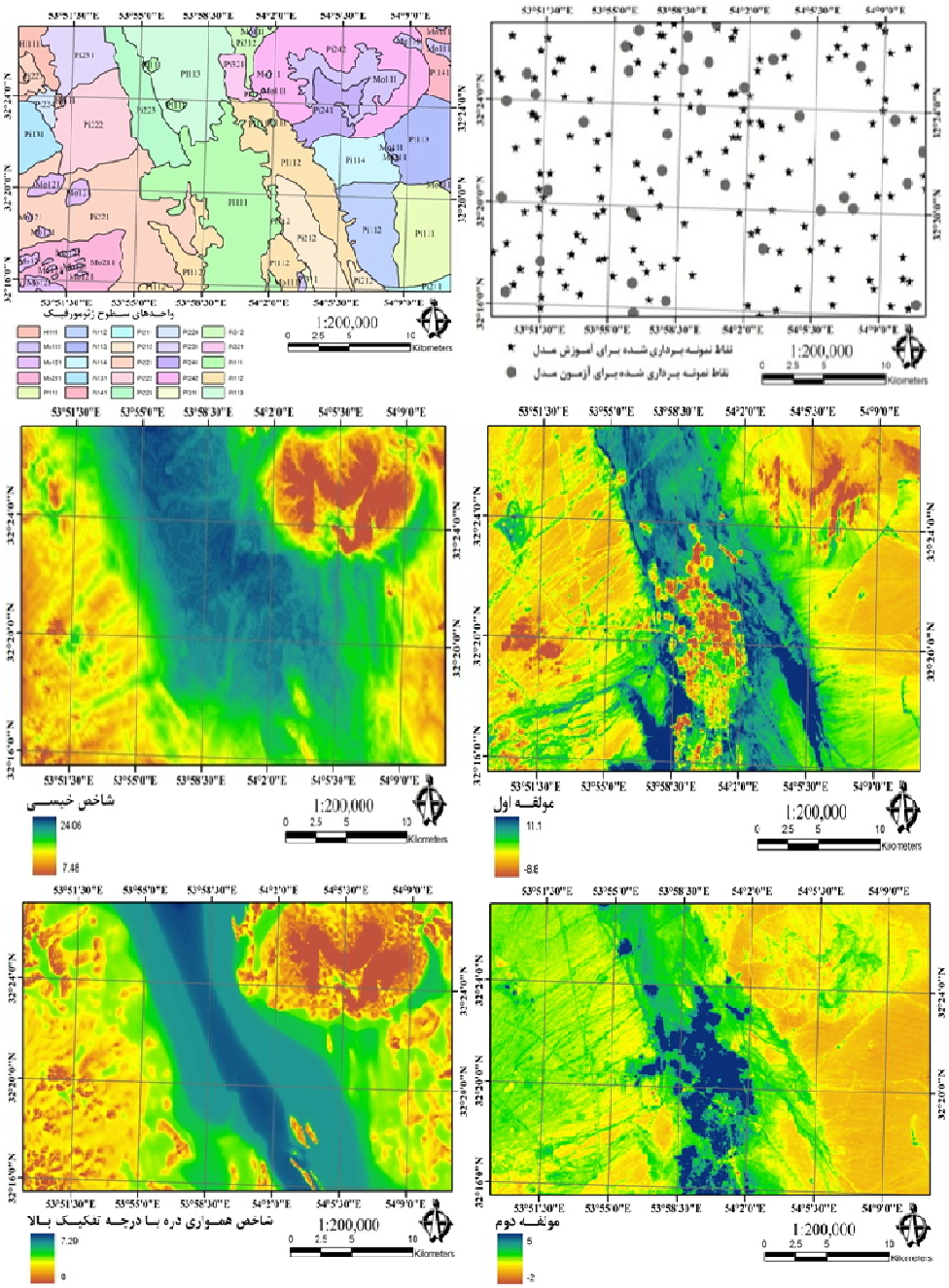
مدل‌سازی

الگوریتم ژنتیک: الگوریتم ژنتیک یک تکنیک برنامه‌نویسی است که از تکامل ژنتیکی به عنوان الگوی حل مسئله استفاده می‌کند (احمد و سیمونوویچ[2]، 2005). در این روش ابتدا برای مجموعه‌ای از داده‌ها که جمعیت نامیده می‌شود، پارامترهای هدف بصورت اتفاقی تولید و افراد در برابر این مجموعه از داده‌ها، مورد آزمایش قرار گرفته و مناسب‌ترین نسل جدید را شکل می‌دهند و این فرآیند برای نسل‌های بعدی تا ارضای معیار همگرایی تکرار می‌شود. در این تحقیق تقریبا مشابه پیش فرض برنامه، تعداد جمعیت اولیه 50، حداکثر نسل 100 و تعداد گردش برابر با 100 دور، در نظر گرفته شده است. در این مطالعه، از روش الگوریتم ژنتیک به عنوان روش بهینه‌سازی وزن‌های شبکه عصبی مصنوعی استفاده گردید.

درخت تصمیم: مدل درختی طبقه‌بندی و رگرسیون، یک روش ناپارامتری الگوریتمی است که قادر به پیش‌بینی متغیرهای کمی یا متغیرهای طبقه‌بندی شده بر اساس مجموعه‌ای از متغیرهای پیش‌بینی کننده کمی و کیفی است. در این روش، مجموعه‌ای از شرط‌های منطقی به صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی یا پیش‌بینی کمی یک متغیر به کار می‌رود. ایجاد درخت تصمیم شامل دو مرحله است. مرحله‌ی اول ایجاد و رشد درخت است. این مرحله شامل پیوند و انشعاب می‌باشد. مرحله‌ی دوم، مرحله توقف و هرس است. هدف از این مرحله به حداقل رساندن خطای پیش‌بینی است. در تحقیق حاضر از نرم‌افزار سی 5 (کینلان[3]، 2001) برای ساختن درخت تصمیم و در نهایت پیش‌بینی کلاس‌های خاک استفاده شد.

1 -Soil Survey Staff

2- Ahmad and Simonovic
3- Quinlan



شکل1) پراکنش نقاط نمونه‌برداری شده بر حسب مربع لاتین و بعضی از پارامترهای کمکی

آنالیز تشخیصی: آنالیز تشخیصی، راهکاری است برای آن‌که متغیرها را در قالب گروه‌های مجزا از یکدیگر تفکیک کنیم، به صورتی‌که هر گروه در عین این‌که با گروه دیگر شباهت و همبستگی دارد، از انسجام لازم نیز برخوردار باشد. در واقع آنالیز تتشخیصی، اعضای جامعه مورد تحقیق را در قالب گروه‌های مجزا از یکدیگر طبقه‌بندی و تفکیک می‌کند، علاوه بر این، تابع تشخیص، معادله‌ای است که با داشتن مشخصات هر فرد از جامعه، می‌توان با قرار دادن این مشخصات در آن معادله، پیش‌بینی کرد که وی به کدام گروه تعلق دارد، به سخن دیگر منظور از آنالیز تشخیصی، گروه‌بندی داده‌ها به گروه‌های متجانس است، به گونه‌ای که مشاهدات هر گروه با یکدیگر شبیه باشند و مشاهدات گروه‌های مختلف نسبت به یکدیگر کمترین شباهت را داشته باشند. برای انجام این آنالیز، کلاس‌های خاک به عنوان متغیر مستقل و اطلاعات محیطی (به‌جز اطلاعات کیفی مثل ژئومورفولوژی) وارد نرم‌افزار متلب[33] گردید.

شبکه عصبی مصنوعی: در این تحقیق شبکه عصبی مصنوعی پرسپترون چند لایه با یک لایه مخفی که دارای تابع فعال‌سازی سیگموئید در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی بوده و تعداد نرون‌های آن از دو تا 10 نرون متغیر بوده و بهترین تعداد نرون به صورت سعی و خطا تعیین شد، مورد بررسی قرار گرفت. همچنین به علت کارایی، سادگی و سرعت بالا، در این تحقیق الگوریتم آموزشی لونبرگ مارکوارت استفاده شد (امینی و همکاران[34]، 2005). در تحقیق حاضر از نرم‌افزار نروسلوشین[35] برای ساختن شبکه عصبی مصنوعی و در نهایت پیش‌بینی کلاس خاک استفاده گردید.

رگرسیون لاجیستیک چندمتغیره: مدل لاجیستیک مورد خاصی از مدل خطی تعمیم‌یافته است که به دو صورت قابل اجرا است: مدل لاجیستیک دوتایی و مدل لاجیستیک چندجمله‌ای. در مدل لاجیستیک دوتایی، متغیر وابسته به شکل حضور یا عدم حضور (صفر و یک) است؛ مثلاً حضور یا عدم حضور افق مشخصه. در حالی که در مدل لاجیستیک چندجمله‌ای، متغیر وابسته دارای چند کلاس یا طبقه است؛ مثلاً کلاس‌های خاک در یک منطقه. در نهایت، برای بررسی کارائی مدل‌های مورد استفاده، از دقت کلی استفاده شد. لازم به توضیح است که برای بررسی دقت مدل برای پیش‌بینی کلاس خاک، داده‌ها به دو دسته آموزش (80 درصد از داده‌های نمونه-برداری شده) و آزمون (20درصد از داده‌های نمونه-برداری شده) به طور تصادفی تقسیم شدند (شکل 1).

نتایج و بحث

طبقه‌بندی خاک‌ها

آنالیز نمونه‌های خاک نشان داد که خاک‌های منطقه بسیار شور، دارای میزان گچ و آهک زیاد و ماده‌آلی بسیار کم می‌باشد، که همگی از ویژگی‌های خاک‌های مناطق خشک است. کلاس بافتی خاک بسیار متغیر می‌باشد، به‌طوری‌که خاک‌هایی که در نواحی مرتفع‌تر هستند، دارای بافت درشت‌تر هستند و هر چه به سمت انتهای حوزه و نواحی با ارتفاع کمتر پیش می‌رویم بافت خاک سنگین‌تر می‌شود. خاک‌های توری‌ارتندز[36] اکثرا در نواحی با ارتفاع و شیب زیاد (شمال شرق) و اراضی فرسایشی (جنوب غرب) مشاهده شدند. یعنی در مناطقی که امکان و زمان کافی برای تشکیل شدن افق مشخصه زیرسطحی وجود ندارد. علاوه بر این، تغییرات شوری اندازه‌گیری شده در این خاک‌ها به کمتر از حتی یک دسی‌زیمنس بر متر می‌رسد. خاک‌هایی با افق کلسیک[37]

33- MATLAB
34 - Amini *et al.*
35- Neurosolutions-5

36- Torriorthents
37- Calcic Horizon

اکثراً در قسمت‌های میانی منطقه مورد مطالعاتی (باغ‌های پسته) مشاهده گردید. تجمعات آهک در این زیرگروه بیشتر در عمق 75-15 سانتیمتری به صورت نودول[38] و پندانت[39] مشاهد شد. به همراه هاپلوکلسیدز، تعداد اندکی (2/6% کل مشاهدات) خاک‌هایی با طبقه‌بندی هاپلوکمبیدز[40] نیز در باغات پسته مشاهده گردید. کلسی‌ژیپسیدز[41] خاک‌هایی هستند که دارای یک افق کلسیک در عمق 50-15 سانتی‌متری و یک افق مشخصه ژیپسیک[42] در عمق 100-50 سانتی‌متری می‌باشند. که این حالت قرارگیری افق‌های یاد شده نشان‌دهنده‌ی شستشوی از بالا به پایین می‌باشد؛ به‌طوری‌که ذرات آهک که دارای حلالیت کم‌تر نسبت به گچ هستند، در بالای پروفیل به صورت پندانت و نودول رسوب کرده‌اند. این زیرگروه خاک در قسمت‌های غربی منطقه مشاهده شد. هاپلوسالیدز[43] تنها یک افق مشخصه سالیک دارند و دارای بافت سنگین و ریزدانه هستند و میزان شوری خاک بخصوص در افق‌های سطحی به بیش از 60 دسی‌زیمنس بر متر می‌رسد. هاپلوژیپسیدز در قسمت‌های شمال غربی منطقه‌ی مطالعاتی دیده شد. تجمع گچ در این خاک‌ها اکثراً به صورت پندانت مشاهده گردید.

مدل‌سازی مکانی

شبکه عصبی مصنوعی: توپولوژی شبکه، یک فاکتور اساسی در طراحی شبکه عصبی مصنوعی می‌باشد. تعداد لایه‌های مخفی و تعداد نرون‌ها جزء اجزای اصلی شبکه‌های پرسپترون چندلایه می‌باشد. بجز توپولوژی شبکه عصبی مصنوعی که ذکر گردید تعداد اپوک در مرحله یادگیری بسیار حائز اهمیت می‌باشد. تعداد نرون‌ها در لایه‌های ورودی و خروجی ثابت می‌باشد و بسته به تعداد ورودی (متغیرهای مستقل) و خروجی (کلاس‌های خاک) مدل به کار گرفته شده دارد. لایه‌ی ورودی شبکه شامل 15 ورودی می‌باشد. لایه‌ی خروجی شامل شش نرون می‌باشد. تعداد نرون‌ها و لایه‌های مخفی بستگی به پیچیدگی مسئله مورد نظر دارد. در اینجا یک لایه مخفی برای ایجاد شبکه پرسپترون چندلایه مورد استفاده قرار گرفت. تعداد نرون‌های لایه مخفی و همچنین تعداد اپوک[44] در فرآیند آموزش و به‌صورت سعی و خطا به دست آمدند. به طور اختصاصی در این مسئله‌ی مورد نظر، تعداد نرون دو تا 10 و تعداد اپوک 100 تا 1000 مورد آزمایش قرار گرفت. برای تعیین بهترین ترکیب از شاخص صحت کلی استفاده گردید. نتایج نشان داد که تعداد نرون نه و تعداد اپوک 200 دارای بهترین ترکیب برای پیش‌بینی گروه بزرگ خاک می‌باشد. در نتیجه ساختار 6-9-15 برای پیش‌بینی کلاس‌های خاک در سطح گروه بزرگ در منطقه مورد مطالعه پیشنهاد می‌گردد. همچنین نتایج ماتریس ابهام برای مرحله‌ی آموزش با 150 داده نشان داد که کلاس هاپلوسالیدز (96/15%) دارای بالاترین صحت طبقه‌بندی می‌باشد؛ در حالی‌که مدل در مرحله آموزش نتوانسته هیچ یک از هاپلوکمبیدها (0%) را به‌درستی طبقه‌بندی کند، که دلیل آن را می‌توان در تعداد کم داده مشاهداتی جستجو کرد (جعفری و همکاران، 2013). دقت کلی شبکه عصبی مصنوعی جهت پیش‌بینی کلاس خاک در سطح گروه بزرگ در مرحله آزمون 65% می‌باشد. همچنین نتایج آنالیز حساسیت شبکه عصبی مصنوعی نشان داد که پارامتر ژئومورفولوژی بالاترین اثر برای پیش‌بینی کلاس‌های خاک در سطح گروه بزرگ را دارد. بعد از سطوح ژئومورفیک پارامترهای سرزمین از جمله شاخص خیسی و همواری دره با درجه تفکیک بالا حائز اهمیت می‌باشند. محققان مختلفی از شبکه‌های عصبی مصنوعی برای پیش‌بینی کلاس خاک استفاده

38- Nodules
39- Pendants
40- Haplocambids
41- Calcigypsids
42- Gypsic
43- Haplosalids

44- Epoch

کرده‌اند؛ برای نمونه لهمن و همکاران[45] (1999)، و بهرنس و همکاران (2005) با استفاده از داده‌های کمکی (شیب، انحنای شیب، ارتفاع، جهت شیب و تابش خورشیدی، ژئومرفولوژی و زمین‌شناسی) و مدل مکانی شبکه عصبی مصنوعی پیش‌بینی مکانی واحدهای خاک با دقت 77% را انجام دادند. ژائوو همکاران[46] (2009) نیز توانستند با استفاده از شبکه عصبی مصنوعی، نقشه رقومی بافت خاک با دقت کلی 88% را در کشور چین به‌دست آورند.

شبکه عصبی مصنوعی- الگوریتم ژنتیک: در این روش ابتدا برای مجموعه‌ای از داده‌ها که جمعیت نامیده می‌شود، پارامترهای هدف به‌صورت اتفاقی تولید و افراد در برابر این مجموعه از داده‌ها، مورد آزمایش قرار گرفته و مناسب‌ترین نسل جدید را شکل می‌دهند و این فرآیند برای نسل‌های بعدی تا ارضای معیار همگرایی تکرار می‌شود. در این روش از الگوریتم ژنتیک به عنوان روش بهینه‌سازی وزن‌های شبکه عصبی مصنوعی استفاده گردید. بر اساس نتایج در نسل پنجم، شبکه عصبی مصنوعی به عملکرد بهینه دست یافته است. دقت کلی مدل برای مرحله آزمون 65% می‌باشد و بهترین عملکرد مربوط به طبقه هاپلوسالیدز (96/15%) و بدترین پیش‌بینی مربوط به هاپلوکمبیدز (0%) می‌باشد. بر خلاف انتظار نتایج نشان داد که ترکیب دو تکنیک فوق نمی‌تواند دقت طبقه‌بندی در مورد کلاس‌های گروه بزرگ خاک را نسبت به روش شبکه عصبی مصنوعی افزایش دهد. در حقیقت استفاده از الگوریتم ژنتیک زمانی سودمند خواهد بود که شبکه عصبی مصنوعی در حین فرآیند آموزش در کمینه محلی قرار بگیرد، این در حالی است که اگر شبکه عصبی مصنوعی به خوبی آموزش داده شود، ممکن است به کمینه اصلی دست پیدا کند. در چنین شرایطی استفاده از الگوریتم ژنتیک کمک معنی‌داری در افزایش دقت مدل نهایی نخواهد کرد (احمد و سیمونوویچ، 2005). همچنین نتایج آنالیز حساسیت این مدل نشان داد که مهم‌ترین پارامتر ورودی، اجزاء سرزمین از جمله شاخص همواری دره با درجه تفکیک بالا می‌باشد. بعد از آن، پارامتر زمین‌شناسی عامل تاثیرگذار در شناسایی خاک‌های منطقه تا سطح گروه بزرگ می‌باشد.

رگرسیون لاجیستیک: در اجرای مدل لاجیستیک چندجمله‌ای، گروه بزرگ خاک به عنوان متغیر هدف و متغیرهای کمکی (پلی‌گونی، شاخص‌های حاصل از داده‌های ماهواره‌ای و اجزاء سرزمین) به عنوان متغیرهای پیش‌بینی کننده در نظر گرفته شدند. نتایج آماری نشان داد که مدل رگرسیونی برازش داده شده در سطح 1% معنی‌دار می‌باشد. دقت کلی مدل در مرحله 55% برای پیش‌بینی گروه‌های بزرگ خاک می‌باشد. بهترین طبقه‌ی خاک متعلق به گروه بزرگ توری‌ارتندز و بدترین نتایج مربوط به هاپلوژیپسیدز می‌باشد. در این مرحله علاوه بر شاخص کلی مدل، مشخصه عامل گیرنده[47] نیز محاسبه گردید. این منحنی، امکان تعیین بهترین ترکیب حساسیت و ویژگی[48] در تشخیص حضور یا عدم حضور را فراهم می‌کند (ریچاردسون و ویگاند[49]، 1977). سپس، سطح زیر منحنی[50] (پترسون و همکاران[51]، 1954) به عنوان شاخصی از کیفیت کلی نقشه تخمین زده شد. سطح زیر منحنی مساوی یک، نشان دهنده قدرت بالای روش در تشخیص حضور و عدم حضور و سطح زیر منحنی مساوی 0/5 نشان می دهد که روش مورد استفاده هیچ قدرت تمایزی در تشخیص حضور و عدم حضور ندارد و در واقع تشخیص به صورت تصادفی است. نتایج نشان داد که سطح زیر منحنی برای هر شش کلاس

45- Lehmann *et al.*
46- Zhao *et al*.
47- Receiver Operating Characteristic (ROC)
48- Sensitivity and Specificity
49- Richardson and Wiegand
50- Area Under Curve (AUC)
51- Peterson *et al.*

خاک نزدیک به یک شده است که نشان‌دهنده دقت مدل رگرسیونی می‌باشد. همچنین نتایج نشان داد که این مدل از نقشه‌ی ژئومورفولوژی، ارتفاع بالای شبکه زهکشی، شیب میانی، عمق دره، شاخص خیسی، شاخص رس و شاخص شوری به عنوان مهم‌ترین متغیرهای پیش-بینی کننده استفاده کرده است. همان‌طورکه نتایج نشان داد مدل رگرسیون لاجیستیک، سطوح ژئومورفولوژی را به عنوان یک پیش‌بینی کننده قوی شناخته است. این موضوع مؤید نقش مؤثر فرآیندهای ژئومورفولوژی در توسعه‌ی خاک منطقه‌ی مورد مطالعه و از طرفی نشان دهنده‌ی انتخاب مناسب این متغیر پیش‌بینی‌کننده نیز می‌باشد. این نتیجه همچنین در مورد مدل‌های شبکه عصبی مصنوعی نشان داده شد. بنابراین، براساس نتایج به دست آمده از مدل‌سازی، نقش انکارناپذیر فرآیندهای ژئومرفولوژی در این مطالعه تأیید می‌شود؛ البته نتایج مدل رگرسیون لاجیستیک در انتخاب ورودی‌ها تا اندازه‌ای با نتایج شبکه عصبی مصنوعی و همچنین شبکه عصبی مصنوعی-الگوریتم ژنتیک متفاوت است. به‌طوری‌که در این مدل از شاخص‌های ماهواره‌ای نیز استفاده شده است؛ در حالی‌که دو مدل قبلی از هیچ یک از داده‌های مستخرج از تصویر +ETM استفاده نکرده بودند. محققان مختلفی از جمله گیاسون و همکاران[52] (2008)؛ هنگل و همکاران (2007)؛ دابلاگیلو و اتزلمولر[53] (2009) و جعفری و همکاران (2012) نیز از مدل‌های رگرسیون لاجیستیکی (دوتایی و چندتایی) برای پیش‌بینی کلاس‌های خاک استفاده کرده‌اند و همگی بیان داشته‌اند که مهم‌ترین عامل در بالا بردن دقت مدل، همبستگی و ارتباط خوب بین متغیرهای محیطی و هدف می‌باشد. البته یکی از حسن‌های روش رگرسیون لاجیستیک تولید نقشه‌های احتمال می‌باشد. در حقیقت نقشه خروجی احتمال حضور یک کلاس از خاک یا افق مشخصه را نشان می‌دهد که برای بررسی عدم قطعیت مدل و همچنین بررسی دقیق‌تر ارتباط بین متغیرهای محیطی و کلاس خاک به‌کار می‌رود.

آنالیز تشخیصی: در این روش متغیرهای کمکی در قالب گروه‌های مجزا از یکدیگر تفکیک می‌شود، به صورتی‌که هر گروه در عین این‌که با گروه دیگر شباهت و همبستگی دارد، از انسجام لازم نیز برخوردار باشد. علاوه بر این می‌توان با وارد کردن مشخصات متغیرهای محیطی در قسمت‌های مختلف منطقه در تابع تشخیص، پیش‌بینی کرد که هر پیکسل به کدام طبقه از خاک تعلق دارد. نتایج آنالیز تشخیصی مشخص کرد که تابع تشخیص، متغیرها را در گروه‌های مختلف قرار داده است. همچنین نتایج پیش‌بینی در مرحله آزمون نشان داد که دقت کلی مدل آنالیز تشخیصی 47% می‌باشد. بدترین طبقه‌بندی در این روش متعلق به گروه بزرگ هاپلوکلسیدز (29%) و بهترین طبقه‌بندی مربوط به کلاس کلسی‌ژیپسیدز (78%) می‌باشد. نتایج منحنی مشخصه، عدم کارایی مدل را در تشخیص گروه هاپلوکلسیدز (0/69) نشان می‌دهد و بهترین طبقه مربوط به کلاس هاپلوکمبیدز (0/91) می‌باشد. همچنین بر اساس روش گام به گام[54] متغیرهایی که به عنوان پیش‌بینی کننده انتخاب شدند عبارت است از: شاخص هموارى دره با درجه تفکیک بالا، عمق دره، شاخص شوری، ارتفاع، مساحت حوزه اصلاح شده و شاخص رس. لازم به ذکر می‌باشد که در مدل آنالیز تشخیصی متغیرهای کمکی پلی‌گونی (نقشه ژئومورفولوژی) را نمی‌توان وارد کرد (لیو و همکاران، 2008).

درخت تصمیم: در این روش، مجموعه‌ای از شرط‌های منطقی به صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی کلاس‌های خاک به کار می‌رود.

52- Giasson *et al.*

53- Debella-Gilo and Etzelmuller

54- Step by Step

نتایج پیش‌بینی در مرحله آزمون نشان داد که دقت کلی مدل درخت تصمیم 70% می‌باشد. در این گروه داده نیز بالاترین دقت متعلق به گروه هاپلوسالیدز (88%) می‌باشد و بدترین طبقه‌بندی مربوط به کلاس هاپلوکمبیدز خاک می‌باشد که دارای دقت پیش‌بینی صفر درصد است. محققان مختلفی کارایی درختان تصمیم را در پیش‌بینی کلاس‌های خاک تایید کرده‌اند (بوی و همکاران، 1999؛ بوی و موران، 2001). آنالیز درخت تصمیم نشان داد که بعضی از متغیرهای کمکی مانند شاخص خیسی، سطوح ژئومورفولوژی، شاخص همواری دره با درجه تفکیک بالا، مولفه‌ی درجه دوم، ارتفاع و ارتفاع بالای شبکه زهکشی، بالاترین تاثیر را برروی پیش‌بینی گروه‌های بزرگ خاک دارند. نتایج نشان داد که شاخص خیسی موثرترین متغیر کمکی در مدل می‌باشد و این شاخص 100% در مدل استفاده شده است. همچنین متغیر کمکی مهم دیگر، شاخص همواری دره با درجه تفکیک بالا می‌باشد که 27% در مدل وارد شده است. این موضوع می‌تواند مؤید این مطلب باشد که در این منطقه، پستی و بلندی از مهم‌ترین فاکتورهای خاکسازی می‌باشد. اجزاء سرزمین در توزیع مکانی خاک و ویژگی‌های آن بسیار مؤثر هستند؛ زیرا توزیع مکانی تابش خورشیدی، درجه حرارت، رطوبت و جریان مواد را تحت تأثیر قرار می‌دهند. یکی از اجزاء سرزمین، شاخص خیسی یا رطوبتی است که برای تعیین توزیع مکانی نفوذ و جریان آب سطحی به‌کار می‌رود. در واقع تأثیر توپوگرافی بر مکان و میزان تجمع رطوبت در خاک و یا آب در سطح زمین‌نما را نشان می‌دهد. این پارامتر با متغیرهای وابسته‌ای که با رطوبت ارتباط دارند و تحت تأثیر رطوبت قرار دارند، همبستگی بیشتری نشان می‌دهد. به‌طوری‌که در مطالعه حاضر، شاخص خیسی به عنوان یک متغیر پیش‌بینی کننده معنی‌دار در مدل‌های افق مشخصه سالیک وارد گردید. شاخص خیسی و همواری دره با درجه تفکیک بالا، مناطق مستعد برای دریافت رسوبات را (نواحی میانی) نشان می‌دهند و این قسمت از مناطق عمدتا از خاک‌های شور تشکیل شده‌اند. در حقیقت این شاخص‌ها در شناسایی گروه بزرگ هاپلوسالیدز کمک می‌کنند. به‌طورمشابه مور و همکاران (1991) ارتباط معنی‌دار بین خاک‌های شور و شاخص خیسی را گزارش کرده‌اند. جعفری و همکاران (2012) نیز در منطقه‌ی خشک اقدام به پیش‌بینی کلاس‌های خاک کرده‌اند و شاخص خیسی و همواری دره با درجه تفکیک بالا را به عنوان پارامتر مهم برای مدل معرفی نمودند. سطوح ژئومورفولوژی در رده دوم قرار دارد و 84% در مدل استفاده شده است. سطح ژئومورفولوژی نیز یک پارامتر بسیار مهم می‌باشد که این مسئله می‌تواند بدان دلیل باشد که سطوح ژئومورفولوژی اخیراً تشکیل شده است و بنابراین ارتباط خوبی را با فرآیندهای خاکسازی در مناطق خشک نشان می‌دهد. این نتایج در راستای نتایج محققان دیگر می‌باشد. به عنوان مثال اسکال و همکاران[55] (2005) و جعفری و همکاران (2012) تاکید کردند که سطوح ژئومورفولوژی مهم‌ترین متغیر کمکی برای شناسایی خاک‌ها در مناطق بیابانی می‌باشد. نتایج همچنین نشان داد که سایر متغیرهای کمکی تاثیر چندانی برروی مدل برای پیش‌بینی گروه‌های بزرگ خاک ندارند.

ارزیابی مدل‌ها: برای انتخاب بهترین مدل، اقدام به مقایسه‌ی مدل‌های استفاده شده (شبکه عصبی مصنوعی، شبکه عصبی مصنوعی- الگوریتم ژنتیک، رگرسیون لاجیستیک، آنالیز تشخیصی و درخت تصمیم) بر اساس دقت کلی نقشه گردید. نتایج مقایسه‌ی مدل‌ها برای پیش‌بینی گروه بزرگ خاک نشان داد که مدل درخت تصمیم بهترین عملکرد برای پیش‌بینی گروه بزرگ خاک را دارا می‌باشد. بعد از درخت تصمیم، شبکه‌های عصبی و الگوریتم ژنتیک با دقت کلی 65% در مرتبه دوم قرار می‌گیرند. رگرسیون لاجیستیک (0/55) در رتبه‌ی

55- Scull *et al.*

بعدی قرار می‌گیرد. از بین مدل‌های استفاده شده برای طبقه‌بندی خاک‌ها، آنالیز تشخیصی دارای کمترین دقت کلی 0/47 می‌باشد. علت این موضوع را می‌توان در نوع ورودی‌های این مدل جستجو کرد. همان‌طور که قبلا بیان گردید این مدل قادر به پذیرش داده‌های پلی‌گونی از جمله سطوح ژئومورفولوژی نمی‌باشد، بنابراین احتمالا همین مسئله باعث قرارگیری مدل آنالیز تشخیصی از لحاظ دقت کلی مدل در رده آخر در مقایسه با سایر مدل‌ها برای پیش‌بینی کلاس خاک است. لذا؛ از این مدل برای تهیه نقشه رقومی خاک تا سطح گروه بزرگ استفاده گردید (شکل 2). به‌طورمشابه لوتو و جرت (2005) کارائی بالاتر درخت تصمیم را در مقایسه با شبکه عصبی نشان دادند. با این وجود مونجون و همکاران[56] (2010) اقدام به مقایسه‌ی تکنیک شبکه - عصبی مصنوعی و درخت تصمیم برای پیش‌بینی واحدهای خاک نموده و نشان دادند که تفاوت معنی-داری بین دو تکنیک برای پیش‌بینی وجود ندارد.

نقشه رقومی زیرگروه خاک: در مرحله بعد، برای تهیه نقشه با اطلاعات جزئی‌تر اقدام به تهیه نقشه رقومی برای سطح زیرگروه (8 کلاس) گردید. با توجه به طبقه-بندی خاک‌ها در سطح زیرگروه، متوجه می‌شویم که تنها هاپلوسالیدز به سه کلاس طبقه‌بندی می‌شود و سایر گروه‌ها مشاهدات پروفیلی کمتری را دارند. علاوه بر این هاپلوسالیدز در منطقه از لحاظ مساحت غالب می‌باشد و ما می‌توانیم تنها مدل‌سازی را برای این قسمت از منطقه جداگانه انجام بدهیم. بنابراین نواحی که در نقشه رقومی تهیه شده به عنوان هاپلوسالیدز در نظر گرفته شده بود، خارج گردید و برای آن مدل‌سازی درختی انجام گرفت. مدل درختی در این مرحله از متغیرهای کمکی ارتفاع (100%)، مولفه اول (43%) و مولفه دوم (41%) بیشتر از سایر داده‌های محیطی برای شناسایی سه کلاسی که در هاپلوسالیدز قرار دارد استفاده کرده است. نتایج این تحقیق همچنین نشان داد که دقت کلی برای مرحله آموزش و آزمون 85/7% و 84/2% می‌باشد. مدل درختی به‌کار رفته، کلاس خاک ژیپسیک هاپلوسالیدز را با بالاترین دقت کاربر 100% بر اساس داده‌های آزمون پیش‌بینی کرده است. در نهایت نقشه نهایی خاک در سطح زیر گروه با ترکیب قوانینی که در دو مرحله قبل (گروه بزرگ و زیرگروه‌های هاپلوسالیدز) به‌دست آمده بود، تهیه گردید (شکل 3). با توجه به این شکل، کاملاً مشخص است که ژیپسیک هاپلوسالیدز در منطقه غالب می‌باشد و پتروژیپسیک هاپلوسالیدز را می‌توان در نواحی شرقی مشاهده نمود.

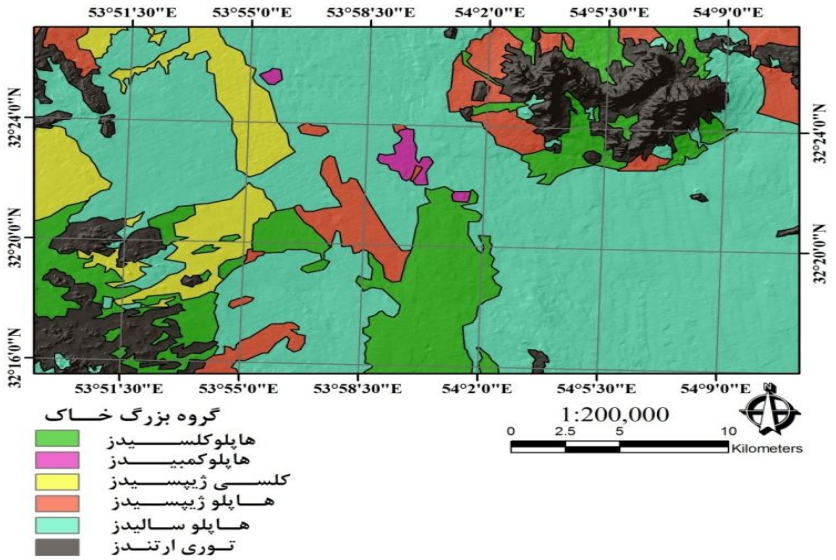
نتیجه‌گیری

نتایج مقایسه‌ی انواع مدل‌ها برای پیش‌بینی کلاس خاک (گروه بزرگ) نشان داد که مدل درخت تصمیم دارای بالاترین کارایی در جداسازی خاک‌ها می‌باشد..

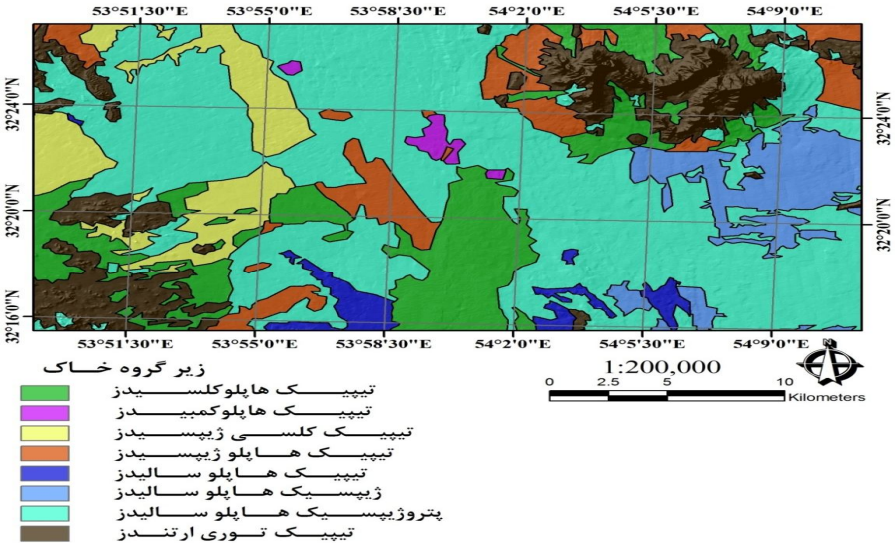
البته سایر مدل‌های استفاده شده از قبیل رگرسیون لاجیستیک و شبکه عصبی مصنوعی نیز به دقت‌های قابل قبولی جهت پیش‌بینی کلاس خاک دست یافتند. نتایج همچنین نشان داد که در منطقه مورد مطالعه، مهم‌ترین متغیر محیطی شاخص خیسی و سطوح ژئومورفولوژی می‌باشند. وجود ارتباط قوی بین داده خاک و پارامترهای محیطی نیز یکی دیگر از فاکتورهای تاثیرگذار برروی دقت می‌باشد. نتایج همچنین اهمیت اثر تعداد خاک‌رخ مشاهده شده در هر کلاس خاک را بر روی صحت طبقه‌بندی نشان داد. در کل می‌توان گفت که تکنیک‌های نقشه‌برداری رقومی یا به عبارت دیگر تکنیک‌های نقشه‌برداری رقومی خاک می‌توانند فرآیند نقشه‌برداری خاک‌ها را در گستره‌ای وسیع و متشکل از هر گونه عوارض طبیعی به یک‌باره انجام داده، روش سنتی را ارتقاء بخشیده، سرعت عمل و کارآمدی نقشه‌ها را در انتقال داده‌ها و اطلاعات افزایش داده و قابلیت استفاده آن‌ها را برای قشر وسیعی از شاخه‌های علمی

56 - Moonjun *et al.*

امکان‌پذیر نماید. البته مطالعات در زمینه نقشه‌برداری رقومی خاک باید ادامه پیدا کند و در مطالعات آینده سعی شود در سطوح پایین رده‌بندی خاک، پیش‌بینی صورت پذیرد.



شکل2- نقشه رقومی خاک در سطح بزرگ خاک بر اساس مدل درخت تصمیم



شکل 3- نقشه رقومی خاک در سطح زیر گروه بر اساس مدل درخت تصمیم

منابع

1. Ahmad, S., and Simonovic, S.P. 2005. An artificial neural network model for generating hydrograph from hydro meteorological parameters. Hydrology, 315: 236-251.

2. Amini, M., Abbaspour, K.C., Khademi, H., Fathianpour, N., Afyuni, M., and Schulin, R. 2005. Neural network models to predict cation exchange capacity in arid regions of Iran. Europian Journal of Soil Science, 53: 748–757.

3. Behrens, T., and Scholten, T. 2007. A comparison of data-mining techniques in predictive soil mapping. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping. An Introductory Perspective. Developments in Soil Science, vol. 31. Elseveier, Amsterdam, pp. 353- 364.

4. Behrens, T., Foster, H., Scholten, T., Steinrucken, U., Spies, E., and Golshmitt, M. 2005. Digital soil mapping using artificial neural networks. Journal of Plant Nutrition Soil Science, 169: 434-443.

5. Bui, E.N., and Moran, C.J. 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. Geoderma, 103: 79– 94.

6. Bui, E.N., Loughhead, A., and Corner, R. 1999. Extracting soil-landscape rules from previous soil surveys. Australian Journal of Soil Research, 37: 495- 508.

7. Burrough, P. A., and McDonnell, R.A. 1998. Principles of geographic information systems. (Revised edition.) Oxford: Clarendon Press.

8. Debella-Gilo, M., and Etzelmuller, B. 2009. Spatial prediction of soil classes using digital terrain analysis and multinomial logistic regression modeling integrated in GIS: Examples from Vestfold County, Norway. Catena, 77: 8-18.

9. Dobos, E., Micheli, E., Baumgardner, M.F., Biehl, L., and Helt, T., 2000. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. Geoderma, 97: 367- 391.

10. Du, C., Linker, R., and Shaviv, A., 2008. Identification of agricultural soils using mid-infrared photoacoustic spectroscopy. Geoderma. 143: 85–90.

11. Finke, P.A. 2012. On digital soil assessment with models and the Pedometrics agenda. Geoderma, 171: 3-15.

12. Giasson, E., Figueiedo, S.R., Tornquist, C.G., and Clarke, R.T. 2008. Digital soil mapping using logistic regression on terrain parameters for several ecological regions in Southern Brazil. In Hartemink, A.E., McBratney, A.B., and Mendonca-Santos, M.L. (eds.), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 3–14.

13. Grunwald, S. 2009. Multi-criteria characterization of recent digital soil mapping andmodeling approaches. Geoderma, 152: 195–207.

14. Hengl, T., Toomanian, N., Reuter, H., and Malakouti, M.J. 2007. Methods to interpolate soil categorical variables from profile observations: Lessons from Iran. Geoderma, 140: 417–427.

15. Jafari, A., Ayoubi, S., Khademi, H., Finke, P.A., and Toomanian, N. 2013. Selection of a taxonomic level for soil mapping using diversity and map purity indices: A case

study from an Iranian arid region. Geomorphology, 201: 86-97.

16. Jafari, A., Finke, P.A., de Wauw, J.V., Ayoubi, S., and Khademi H. 2012. Spatial prediction of USDA- great soil groups in the arid Zarand region, Iran: comparing logistic regression approaches to predict diagnostic horizons and soil types. European Journal of Soil Science, 63: 284– 298.

17. Lehmann, D., Billen, N., and Lenz, R. 1999. Anwendung von Neuronalen Netzen in der Landschaftsokologie-Synthetische Bodenkartierung im GIS, In Strobl, J., and Blaschke, T. (eds.), Anwndung von Bodenkartierung im GIS. Wichmann, Heidelberg, pp. 330-336.

18. Liu, J., Pattey, E., Nolin, M.C., Miller, J.R., and Ka, O. 2008. Mapping within-field soil drainage using remote sensing, DEM and apparent soil electrical conductivity. Geoderma, 143: 261–272.

19. Luoto, M., and Hjort, J. 2005. Evaluation of current statistical approaches for predictive geomorphological mapping. Geomorphology, 67: 299-315.

20. McBratney, A.B., Mendonça-Santos, M.L., and Minasny, B. 2003. On digital soil mapping. Geoderma, 117: 3–52.

21. Minasny, B., and McBratney A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computer & Geosciences, 32: 1378–1388.

22. Minasny, B., and McBratney, A.B. 2007a. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. Geoderma, 142: 285–293.

23. Minasny, B., and McBratney, A.B. 2007b. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. Geoderma, 140, 324–336.

24. Moonjun, R., Farshad, A., Shrestha, D.P., and Vaiphasa, C. 2010. Artificial Neural Network and Decision Tree in Predictive Soil Mapping of Hoi Num Rin Sub-Watershed, Thailand. Digital Soil Mapping, Progress in Soil Science 2, pp. 151- 164

25. Moore, I.D., Turner, A.K., Wilson, J.P., Jenson, S.K., and Band, L.E. 1993. GIS and landsurface-subsurface process modelling. In Goodehild, M.F., Parks, B., and Steyaert, L.T. (eds.), Environmental Modeling with GIS. Oxford University Press, Oxford, pp. 196– 230.

26. Odeh, I.O.A., and Onus, A. 2008, Spatial Analysis of Soil Salinity and Soil Structural Stability in a Semiarid Region of New South Wales, Australia. Environmental Management, 42: 265–278

27. Peterson, W.W., Birdsall, T.G., and Fox, W.C. 1954. The theory of signal detectability. Transactions of the IRE Professinal Group in Information Theory, PGIT, 2-4: 171-212.

28. Quinlan, J.R. 2001. See5: An Informal Tutorial. http://www.rulequest.com.

29. Richardson, A.J., and Wiegand, C.L. 1977. Distinguishing vegetation from soil background information. Remote Sensing, 43: 1541-1552.

30. Rivero, R.G., Grunwald, S., and Bruland, G.L., 2007. Incorporation of spectral data into multivariate geostatistical models to map soil phosphorus variability in a Florida wetland. Geoderma, 140: 428– 443.

31. Rossiter, D.G. 2004. Digital soil resource inventories: status and prospects. Soil Use and Management, 20: 296-301

32. Scull, P., Franklin, J., and Chadwick, O.A., 2005. The application of classification of tree analysis to soil type prediction in a desert landscape. Ecological Modeling, 181: 1- 15.

33. Soil Survey Staff. 2010. Keys to Soil Taxonomy, 10th edition. United States Department of Agriculture, Washington.

34. Vasques, G.M., Grunwald, S., and Sickman, J.O. 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma, 146: 14– 25.

35. Zhao, Z., Chow, T.L., Rees, H.W., Yang, Q., Xing Z., and Meng, F. 2009. Predict soil texture distributions using an artificial neural network model. Computers and Electronics in Agriculture, 65: 36–48.