

ارزیابی کیفیت اطلاعات مکانی داوطلبانه از طریق ارزیابی اعتماد پذیری کاربران به کمک شبکه عصبی مصنوعی

الهه آذری عسگری^{۱*}، فرهاد حسینعلی^۲

۱- کارشناسی ارشد مهندسی عمران- سیستم های اطلاعات جغرافیایی، گروه مهندسی نقشه برداری، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران

۲- استادیار گروه مهندسی نقشه برداری، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

تاریخ دریافت مقاله: ۱۴۰۲/۰۳/۰۱ تاریخ پذیرش مقاله: ۱۴۰۲/۰۶/۲۷

چکیده

امروزه اطلاعات مکانی داوطلبانه (VGI) یکی از منابع مهم جمع‌آوری اطلاعات مکانی می‌باشد. برخلاف داده‌های رسمی، دقت و صحت اطلاعات مکانی داوطلبانه در هنگام تولید آنها مشخص نیست. تاکنون روش‌های متفاوتی برای ارزیابی کیفیت این اطلاعات به کار گرفته شده که یکی از آنها ارزیابی عملکرد کاربران VGI در ورود اطلاعات است. در این پژوهش به منظور ارزیابی کیفیت اطلاعات مکانی داوطلبانه، از روش ارزیابی عملکرد کاربران با روپردازی برآورد اعتماد پذیری آنها به کمک شبکه‌های عصبی مصنوعی استفاده شده است. در اینجا به کمک یک برنامه کاربردی تلفن همراه تحت سیستم عامل اندروید، چندین متغیر مربوط به اطلاعات زمینه‌ای کاربران همزمان در نظر گرفته شده است و با استفاده از اطلاعات زمینه‌ای کاربران و شبکه عصبی مصنوعی، میزان اعتماد پذیری کاربران برآورد گردیده است. در برنامه ذکر شده، کاربر سؤالاتی نظری جنسیت، سن، تحصیلات، میزان آشنایی با GPS و یا GIS را پاسخ می‌دهد و سپس بخشی از نقشه منطقه مورد نظر را می‌بیند و پاسخ سؤال مطرح شده درباره نقشه را از میان گزینه‌های ارائه شده انتخاب می‌کند. در این پژوهش هر کاربر باید به سؤالات حداقل سه منطقه پاسخ بدهد. پس از اتمام فرآیند جمع‌آوری اطلاعات، پاسخ کاربران برای ۱۱۰۲ منطقه جمع‌آوری گردید و وارد مرحله پردازش با شبکه عصبی شد. ۷۰٪ داده‌ها برای آموخت شبکه عصبی، ۱۵٪ برای اعتبارسنجی و ۱۵٪ برای آزمایش در نظر گرفته شد و بهترین حالت شبکه عصبی با خطای کمترین مربعات ۱۹۹۸۸،۰، انتخاب گردید. با استفاده از شبکه عصبی آموخت دیده، درصد احتمالی پاسخ‌های صحیح کاربران بر اساس اطلاعات پیشینه آنها برآورد می‌گردد و بدین ترتیب هر مشارکتی که کاربر در تولید اطلاعات داوطلبانه داشته باشد می‌تواند با برچسبی از درجه اطمینان مشخص گردد. شبکه عصبی آموخت دیده برای ارزیابی اطلاعات زمینه‌ای مؤثر کاربران در صحت اطلاعات مکانی هم به خدمت گرفته شد و بهترین حالت ممکن برای ارائه پاسخ‌های صحیح برآورد شده است.

کلید واژه‌ها : اطلاعات مکانی داوطلبانه، VGI، سیستم اطلاعات مکانی، GIS، شبکه عصبی مصنوعی (ANN).

* نویسنده مکاتبه کننده: گروه مهندسی نقشه برداری، دانشکده فنی مهندسی واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.

تلفن: +۹۸۲۱۳۳۷۷۲۸۳۱

۱- مقدمه

تحول داده‌ها، یعنی تاریخچه یا پیشرفت آن مورد تجزیه و تحلیل قرار می‌گیرد. مزیت این رویکرد در مقایسه با روش اول این است که به منابع داده رسمی نیاز ندارد و بروز بودن داده‌ها را در نظر می‌گیرد. با این حال، این روش کیفیت داده‌ها را به طور دقیق ارزیابی نمی‌کند، اما در عوض یک تقریب ارائه می‌دهد [۲، ۸ و ۹].

گوچایلد و لی^۶ (۲۰۱۲)، سه رویکرد برای تضمین کیفیت را توصیف کرده‌اند که به آن‌ها رویکردهای جمع‌سپاری، اجتماعی و جغرافیایی گفته می‌شود. رویکردهای جمع‌سپاری «به توانایی گروه برای تأیید و تصحیح خطاهایی که ممکن است یک فرد مرتکب شود اشاره دارد» در حالی که رویکردهای اجتماعی «به سلسله مراتبی از افراد مورد اعتماد که به عنوان تعديل کننده یا دروازه‌بان عمل می‌کنند، متکی است» [۱۰]. از سوی دیگر سناراتن^۷ و همکاران (۲۰۱۷) ادبیات ارزیابی کیفیت داده‌های *VGI* را بررسی کردند. آن‌ها روش‌های ارزیابی کیفیت *VGI* را به سه دسته طبقه‌بندی کردند: مبتنی بر نقشه، مبتنی بر تصویر و مبتنی بر متن. همچنین آمار آن‌ها نشان داد که روش‌های مبتنی بر نقشه بیشترین فراوانی را در بین سه روش داشتند [۱۱]. فرقانی و دلاور (۲۰۱۴) کیفیت داده‌های *OSM* را در تهران با نقشه رسمی شهرداری تهران با استفاده از معیارهایی مانند مساحت و محدوده مقایسه کردند و دریافتند که کیفیت داده‌های *OSM* نسبتاً خوب است [۱۲]. محمدی و ملک (۲۰۱۴) دقت موقعیتی داده‌های داوطلبانه را با استفاده از شاخص‌های ذاتی تخمین زدند و برای این منظور، شاخص‌هایی را که بر دقت موقعیتی *VGI* تأثیر گذار بودند، استخراج کردند. شاخص‌های استخراج شده در یک سیستم مبتنی

در سال‌های اخیر، تمایل به استفاده از وب^۱ برای ایجاد، جمع‌آوری و انتشار اطلاعات مکانی ارائه شده توسعه افراد به صورت داوطلبانه وجود داشته است. همچنین کاربران علاقه‌مند به مشارکت در زمینه اطلاعات مکانی هستند. این موضوع اولین بار توسط گوچایلد^۲ (۲۰۰۷) با نام اطلاعات مکانی داوطلبانه (*VGI*)^۳ بیان گردید [۱]. این پتانسیل را دارد که منبع مهمی از اطلاعات مکانی در سطح زمین باشد. وب سایتهاي مانند *OSM*^۴ و *Wikimapia* مونه‌هایی از سیستم‌های مبتنی بر *VGI* هستند که اغلب ازان ترین منبع و گاهی تنها منبع اطلاعات مکانی هستند. با این حال، ادغام داده‌های *VGI* با داده‌های معتبر، زیرساخت داده‌های مکانی^۵ و سایر داده‌ها به این دلیل که تعیین دقیق کیفیت *VGI* دشوار است به عنوان یک مسئله مطرح است [۲ و ۳]. دو روش اصلی برای ارزیابی کیفیت داده‌های *VGI* وجود دارد. در رویکرد اول، فرض بر این است که داده‌های معتبر رسمی، داده‌هایی با کیفیت هستند و می‌توانند برای اعتبارسنجی داده‌های *VGI* استفاده شوند. مهم‌ترین نقطه ضعف این رویکرد این است که نیازمند دسترسی به داده‌های رسمی می‌باشد که ممکن است همیشه امکان‌پذیر نباشد. همچنین، به دلیل سرعت کم بروزرسانی مجموعه‌های داده معتبر، صحت داده‌های معتبر را نمی‌توان تضمین کرد، در نتیجه بین داده‌ها و واقعیت در طول زمان ایجاد مغایرت می‌شود [۴]. از نمونه تحقیق‌های صورت گرفته در این دسته می‌توان به تحقیق فونته^۶ و همکاران (۲۰۱۵) [۵]، مرادی و همکاران (۲۰۲۲) [۶] و احمد^۷ و همکاران (۲۰۲۲) [۷] اشاره کرد. در رویکرد دوم، سیر

^۶ Fonte^۷ Ahmad^۸ Goodchild and Li^۹ Senaratne^۱ Web^۲ Goodchild^۳ Volunteered Geographic Information (*VGI*)^۴ OpenStreetMap (*OSM*)^۵ Spatial Data Infrastructure (*SDI*)

بیش ر و یانوویچ^۳ (۲۰۱۰) بیان داشتند که مفهوم اعتماد اطلاعاتی با مفهوم تثیت شده اعتماد بین فردی مرتبط می باشد. آن ها پیشنهاد کردند از اعتماد و شهرت اطلاعاتی به عنوان معیاری برای کیفیت اطلاعات استفاده شود [۱۸]. فراتی و کریمی پور (۲۰۱۶) بیان داشتند که مشکل اصلی در اکثر روش های پیشین عدم امکان ارزیابی کیفی عوارض تصحیح شده توسط کاربرانی است که سابقه ای از عملکرد آن ها موجود نیست. در روشی که فراتی و کریمی پور برای کیفیت سنجی داده های اطلاعات مکانی داوطلبانه پیشنهاد کردند، علاوه بر سابقه کاربران سامانه و همچنین ارزیابی های مربوط به ویرایش های هر کاربر، میزان تأثیر گروه حرفه ای و شغلی که کاربر عضو آن هست نیز در نظر گرفته شده است تا تخمین بهتری از کیفیت داده های مکانی حاصل گردد. نتایج اولیه این تحقیق نشان می دهد این روش در طول زمان بهبود پیدا می کند و هرچه تعداد کاربران با گروه های مختلف حرفه ای در سامانه بیشتر شود، میزان کیفیت سنجی نیز بالاتر می رود [۱۹]. خسروی کزا زی و حسینعلی (۲۰۱۷) در تحقیق خود برای تعیین بهترین داده برای هر شیء، ابتدا به هر مشارکت کننده دو امتیاز نفوذ و قطبیت را تخصیص دادند که تخصیص این امتیازها به مشارکت کننده های اطلاعات مکانی داوطلبانه براساس فعالیت هایشان بود. آن ها در راستای ارزیابی روش پیشنهادی خود، یک سامانه آزمایشی اطلاعات مکانی داوطلبانه را با هدف تولید داده های توصیفی یک خیابان در شهر اراک طراحی کردند که توسط ۱۰ مشارکت کننده مورد استفاده قرار گرفت و نتایج آن بیش از ۹۳٪ با واقعیت انطباق داشت [۲۰]. داکستا^۴ (۲۰۱۶) کیفیت داده های ساختمنی OSM را ارزیابی کرد. او ویژگی کامل بودن و همچنین دقت مکانی را

بر هوش مصنوعی ترکیبی برای دستیابی به الگویی برای ارتباط بین شاخص ها و دقت مکانی استفاده شدند. این الگو با استفاده از داده های داوطلبانه با داده های مرجع مربوطه هر منطقه مورد مطالعه به دست آمد. کارایی روش پیشنهادی آن ها برای تخمین دقت مکانی یک مجموعه داده بدون مرجع مربوطه نشان داده شد [۱۳]. فرج الهی و دلاور (۲۰۱۶) در تحقیق خود جهت ارزیابی کیفیت VGI، از روش اعتبار سنجی افراد با توجه به پیشینه مشارکتی آن ها استفاده کردند. آن ها این روش را روشی دقیق برای تعیین اعتبار مشارکت کننده های و در نتیجه اعتبار اطلاعات دانستند [۱۴]. یانگ^۱ و همکاران (۲۰۱۶) مشارکت کننده های OSM را در آلمان، فرانسه و بریتانیا با معیار حرفه ای بودن آنان، ارزیابی کردند. آن ها به این نتیجه رسیدند که بیشترین مشارکت کننده های OSM از نظر تعداد مبتدی هستند. همچنین سه شاخص تمرین، مهارت و انگیزه می توانند بر کیفیت داده های تولید شده تأثیر بگذارند [۱۵]. علی و اشميد^۲ (۲۰۱۴) از روش های طبقه بندی و سلسله مراتبی برای بررسی یکپارچگی VGI بر اساس داده های OSM استفاده کردند. آن ها بر طبقه بندی پارک ها و باغ ها در آلمان، بریتانیا و استرالیا تمرکز کردند و نشان دادند که دقت داده ها در آلمان و بریتانیا بسیار بهتر از استرالیا است. آن ها اذعان کردند که نتایج به دست آمده ممکن است به دلیل استفاده نکردن از داده های رسمی کاملاً صحیح نباشد [۱۶]. واحدی طرقه و آل شیخ (۲۰۱۶)، دقت توصیفی عوارض داده های داوطلبانه OSM (در قالب نام عارضه) را با مقایسه آن ها با داده های شهرداری شهر تهران مورد بررسی قرار دادند. طبق نتایج حاصل، ۳۳٪ دارای نام صحیح، ۴۴٪ نام تقریباً صحیح و ۲۳٪ نام نادرست و دقت توصیفی کل داده ها برابر ۷۷٪ ارزیابی شد [۱۷].

^۴ Da Costa^۱ Yang^۲ Ali and Schmid^۳ Bishr and Janowicz

منطقه نسبت به روش‌های تجربی سنتی نشان می‌دهد [۲۵].

آنچه از مطالعات پیشین استخراج می‌گردد این است که اطلاعات زمینه‌ای (پیشینه) مشارکت‌کنندگان *VGI* و گروه حرفه‌ای و شغلی آن‌ها در ارزیابی کیفیت اطلاعات آن تأثیرگذار است. اما مشکل روش‌های پیشین این است که برای ارزیابی مشارکت‌کننده جدید که سابقه عملکردی از او در دسترس نیست، روش کاملی وجود ندارد. از طرفی استفاده از شبکه عصبی مصنوعی برای مدل‌سازی و برآورد داده‌های *VGI* نتایج مناسبی داشته است. در این تحقیق برای رفع مشکلات پژوهش‌های پیشین روش برآورد اعتمادپذیری مشارکت‌کنندگان *VGI* با استفاده از شبکه عصبی مصنوعی و پیشینه اجتماعی، شغلی و شناختی مشارکت‌کنندگان پیشنهاد گردیده است. در روش پیشنهادی، اطلاعات پیشینه اجتماعی، شغلی و شناختی کاربر جدید اخذ می‌گردد. سپس به وسیله شبکه عصبی آموزش دیده و براساس اطلاعات اخذ شده میزان اعتمادپذیری او برآورد می‌گردد. کاربرد این روش این است که کاربر در هنگام ثبت نام در سامانه‌های *VGI* مشخصات اولیه از خود را وارد می‌کند و شبکه عصبی مصنوعی براساس مشخصات وارد شده، میزان اعتبار کاربر را پیش‌بینی کند. به این ترتیب می‌توان یک شناخت از کیفیت *VGI* به دست آورده. در ادامه مواد و روش‌های استفاده شده، پیاده‌سازی، جمع‌آوری داده، روش اجرا و محاسبات، بحث و نتیجه‌گیری بیان می‌گردد.

۲- مواد و روش‌ها

در این تحقیق، با توجه به مطالعه‌های صورت گرفته، روش ارزیابی و اعتبارسنجی کیفیت اطلاعات مکانی داوطلبانه از طریق اعتباردهی به کاربران و

مطالعه کرد و به این نتیجه رسید که کیفیت داده‌ها در مراکز شهرها به طور قابل توجهی بهتر است اما با فاصله گرفتن از مراکز شهری کاهش می‌باید [۲۱]. مارتلا^۱ و همکاران (۲۰۱۹) یک برنامه کاربردی^۲ اندرویدی^۳ برای جمع‌آوری اطلاعات مکانی از طریق بازی توسعه دادند. افراد مشارکت‌کننده در این تحقیق کسانی بودند که در منطقه مورد مطالعه زندگی می‌کردند و از برنامه بازی استفاده می‌کردند. از این رو مردم و به ویژه دانشجویان، اطلاعات نقشه آن منطقه (واقع در آلمان) را جمع‌آوری نمودند. کارایی روش با استفاده از پرسشنامه‌های تکمیل شده توسط کاربران تأیید شد [۲۲]. کن^۴ و همکاران (۲۰۱۹) از *VGI* برای استخراج رانش زمین از عکس‌های گرفته شده توسط داوطلبان استفاده کردند. آن‌ها یک شبکه عصبی مصنوعی^۵ کانولوشن^۶ برای این منظور توسعه دادند و *VGI* را منبع خوبی برای جمع‌آوری این نوع داده‌ها یافتند [۲۳]. الغنیم^۷ و همکاران (۲۰۲۱) از ویژگی‌های جاده و رفتار مشارکت‌کنندگان برای ارزیابی کیفیت انواع جاده در استفاده کردند. آن‌ها برای این منظور از روش *OSM* یادگیری ماشین استفاده کردند. نتایج تحقیق نشان داد که دقیت ارزیابی مدل به کار رفته زمانی که از یک مجموعه داده قابل اعتماد استفاده شود، ۸۷٪/۷۵٪ است و زمانی که از مجموعه داده غیرقابل اعتماد استفاده شود، ۹۸٪/۵۷٪ می‌باشد [۲۴]. شی^۸ و همکاران (۲۰۲۱)، یک مدل شبکه عصبی مصنوعی چند لایه را برای پیش‌بینی مطلوبیت منطقه‌ای با داده‌های *VGI* طراحی کردند. تحقیق آن‌ها امکان سنجی و اثربخشی روش‌های مبتنی بر داده (به عنوان مثال، شبکه‌های عصبی مصنوعی) را در مدل‌سازی روابط ناشناخته پنهان و دستیابی به عملکرد بهتر برای پیش‌بینی مطلوبیت

^۱ *Artificial Neural Network (ANN)*

^۲ *Convolution*

^۳ *Alghanim*

^۴ *Shi*

^۱ *Martella*

^۲ *Application*

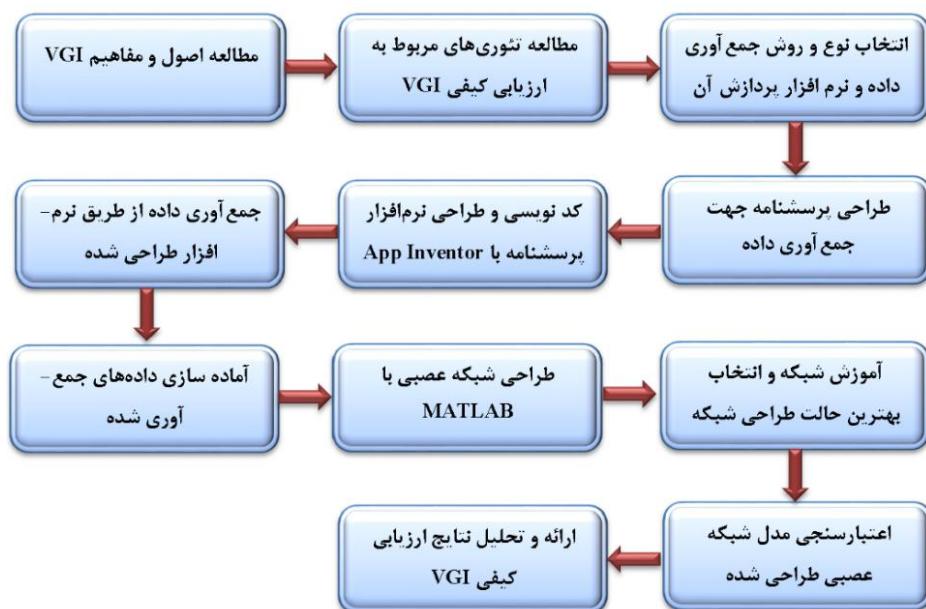
^۳ *Android*

^۴ *Can*

تعدادی سؤال در خصوص مشخصات اولیه کاربر و سؤالاتی در خصوص اطلاعات نقشه شهر تهران بود و از طریق شبکه‌های اجتماعی در اختیار کاربران قرار گرفت تا به صورت داوطلبانه نسبت به پاسخ‌گویی به آن اقدام نمایند. سپس مشخصات جمع‌آوری شده کاربران از طریق روش شبکه عصبی مورد ارزیابی قرار گرفت. در این پژوهش، برای توسعه برنامه کاربردی تلفن هوشمند، ابزار آپ‌اینورتر (*App Inventor*) و برای بخش برنامه‌نویسی شبکه عصبی نیز نرم‌افزار متلب (*MATLAB*) مورد استفاده قرار گرفت. به طور کلی، الگوریتمی که برای اجرای این تحقیق در نظر گرفته شده است، در شکل (۱) نشان داده شده است.

مشارکت‌کنندگان در سامانه‌های *VGI* براساس پیشینه کاربران درنظر گرفته شد. برای این منظور، از شبکه عصبی مصنوعی (*ANN*) برای پیش‌بینی اعتبار کاربران *VGI* براساس ویژگی‌های فردی، تحصیلی، شغلی و غیره استفاده گردید.

برای اجرای این روش نیاز است که برخی مشخصات اولیه مربوط به پیشینه کاربران *VGI* در اختیار باشد. به دلیل اینکه دسترسی به این اطلاعات امکان‌پذیر نیست، ابتدا باید سامانه‌ای طراحی گردد که ویژگی‌های یک سامانه *VGI* یعنی مرتبط با اطلاعات مکانی بودن و داوطلبانه بودن را تأمین نماید و مشخصات اولیه را نیز از کاربر دریافت کند. بنابراین پرسش‌نامه‌ای در قالب برنامه کاربردی تلفن هوشمند طراحی گردید که شامل



شکل ۱: الگوریتم روش تحقیق پیشنهادی

وصل شده‌اند. هر واحد دارای یک مشخصه ورودی/خروجی می‌باشد و محاسبه یا عملی جزئی را اجرا می‌کند. خروجی هر واحد، با توجه به مشخصه آن، اتصال‌های درونیش به سایر واحدها و ورودی‌های خارجی آن تعیین می‌گردد. شبکه عصبی مصنوعی متشکل از یک شبکه نیست، بلکه خانواده‌ای متشکل از

۱-۲- شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی الگوهایی برای پردازش اطلاعات هستند که با تقلید از شبکه عصبی مغزی انسان ساخته شده‌اند [۲۶]. شبکه‌های عصبی مصنوعی، ساختاری (شبکه‌ای) هستند متشکل از تعدادی واحد (نرون (*Neuron*) مصنوعی) که در داخل شبکه به هم

یادگیری نظارت شده^۹ است که باستفاده از مشاهده و اعمال داده‌های آموزشی طبقه‌بندی شده از قبل، انجام می‌شود. در این روش، شبکه‌های عصبی قادر خواهند بود تا با دقت بسیار بالایی و در بسیاری شرایط، از یک مجموعه از داده‌های آموزشی به یک مجموعه کامل از ورودی‌های ممکن، عمومی‌سازی نمایند. روش دیگر، روش یادگیری نظارت نشده^{۱۰} است که وزن‌ها و بایاس‌ها تنها در مقابل ورودی شبکه اصلاح می‌شوند و در واقع هیچ هدفی وجود ندارد. این الگوریتم‌ها اکثراً برای عملیات دسته‌بندی استفاده می‌شوند. آن‌ها الگوهای ورودی را به تعداد محدودی از کلاس‌ها دسته‌بندی می‌کنند. روش دیگر، روش آموزش ترکیبی^{۱۱} است. از آموزشی ترکیبی در موقعي که تعداد نمونه‌ها و داده‌ها خیلی زیاد است، استفاده می‌شود. آموزش ترکیبی به‌این صورت است که ابتدا داده‌ها را براساس روش نظارت نشده، دسته‌بندی می‌کنیم و سپس از هر دسته، نمونه‌هایی انتخاب و براساس روش نظارت شده آموزش انجام می‌شود. پارامترهای شبکه باید طوری تنظیم شوند که پاسخ واقعی شبکه، هر چه بیشتر به پاسخ مطلوب نزدیک شود [۲۶].

شبکه‌های عصبی مصنوعی، چیزی به جز نرون‌های به هم پیوسته نیستند. شکل اتصال نرون‌ها به هم یکی از متغیرهای اصلی در طراحی شبکه‌های عصبی مصنوعی و تنظیم نوع شبکه است. اگر ماتریس وزنی ارتباطات داخلی، فقط به فعالیت بعدی منحصر باشد، یعنی هیچ ارتباط بازخورد و خود اتصالی وجود نداشته باشد، این شبکه را شبکه عصبی پیش‌خور^{۱۲} می‌نامند. شبکه‌های

شبکه‌های گوناگون می‌باشد [۲۷]. شبکه‌های عصبی مصنوعی از قسمت‌های مختلفی تشکیل شده‌اند. اجزای یک شبکه عصبی عبارتند از:

- ورودی‌ها : سیگنال‌های ورودی X_1 تا X_n ، معادل سیگنال‌های عصبی ورودی می‌باشند و در مجموع ورودی نرون را تشکیل می‌دهند.

- بردار وزن : وزن‌های W_{in} تا W_{i2} میزان تأثیر ورودی X_i بر خروجی y ، توسط مشخصه وزن اندازه‌گیری می‌شود.

- تابع جمع^۱ : با رابطه (۱) تعریف می‌شود و عملیات پردازش نرون را انجام می‌دهد:

$$X = \sum_{j=1}^n x_j W_{ij} \quad (1)$$

- تابع فعالیت : محدوده وسیعی از مقادیر ورودی را به مقدار خاص خروجی نگاشت می‌کند.

- خروجی : منظور از خروجی، پاسخ مسئله است.

- خروجی نرون از رابطه (۲) مشخص می‌شود [۲۶]:

$$Y = f\left(\sum_{i=1}^n x_i W_{ij}\right) \quad (2)$$

شبکه‌های عصبی مصنوعی با توانایی قابل توجه خود در استنتاج نتایج از داده‌های پیچیده و مبهم، می‌توانند در استخراج الگوها و شناسایی گرایش‌های مختلفی که برای انسان و کامپیوتر شناسایی آن‌ها بسیار دشوار است، استفاده شوند. از مزایای شبکه عصبی می‌توان قابلیت یادگیری^۲، قابلیت تعمیم^۳، پردازش موازی^۴، مقاوم بودن^۵، قابلیت کاربردی^۶، تشخیص داده‌های اشتباہ^۷، تحمل خطأ^۸، غیر خطی بودن و تصویر کردن ورودی - خروجی را نام برد [۲۶ و ۲۸].

یکی از روش‌های اصلی یادگیری در شبکه‌های عصبی،

^۸ Fault Tolerance

^۹ Supervised Learning

^{۱۰} Unsupervised Learning

^{۱۱} Hybrid Learning

^{۱۲} Feedforward Neural Network

^۱ Adder

^۲ Learning Ability

^۳ Generalization Ability

^۴ Parallel Processing

^۵ Robustness

^۶ Applicability

^۷ Fault Recognition

اختلاف بین پاسخ مطلوب و پاسخ واقعی شبکه. مقدار خطا پس از محاسبه، در مسیر برگشت از لایه خروجی و از طریق لایه‌های میانی شبکه در کل شبکه توزیع می‌گردد و چون این توزیع در خلاف مسیر ارتباطات وزنی صورت می‌گیرد، کلمه پس‌انتشار خطا جهت توضیح رفتار الگوریتم انتخاب شده است. پارامترهای شبکه، طوری تنظیم می‌شوند که پاسخ واقعی شبکه هرچه بیشتر به پاسخ مطلوب نزدیک شود. استفاده از خروجی‌های واقعی در این نوع از شبکه‌ها، سبب قرار گرفتن این شبکه‌ها در گروه آموزشی نظارت شده است [۲۶].

۳- پیاده‌سازی، جمع‌آوری داده و محاسبه

جهت پیاده‌سازی ابتدا باید پرسشنامه، برنامه کاربردی تلفن هوشمند و بهترین چیدمان شبکه عصبی مصنوعی مورد نظر طراحی گردد. سپس جمع‌آوری داده‌ها و محاسبات صورت گیرد.

۱-۳- طراحی پرسشنامه و نقشه‌ها

برای طراحی پرسشنامه تعدادی سؤال همراه با پاسخ طراحی گردید تا کاربران از میان پاسخ‌های موجود، پاسخ مورد نظر خود را انتخاب نمایند. ۱۰ متغیر سن، جنسیت، شغل، مقطع تحصیلی، گروه تحصیلی، آشنایی با نقشه‌برداری، مهندس نقشه‌برداری بودن، آشنایی با GIS^۳، آشنایی با GPS^۴ و نحوه آشنایی برای طراحی پرسشنامه در نظر گرفته شد. طبق مطالعات پیشین، پیشینه فرد در اعتمادپذیری او تأثیرگذار است. بنابراین در طراحی سؤال‌ها علاوه بر سن و جنسیت، پیشینه تحصیلی و شغلی کاربر به عنوان عوامل تأثیرگذار در نظر گرفته شد. همچنین فرض بر این گذاشته شد که آشنایی فرد با نقشه‌برداری، GPS و GIS می‌تواند در کیفیت پاسخ دهی و همچنین کار با سامانه VGI تأثیر داشته باشد. همچنین فرض شد میزان آشنایی فرد با منطقه مورد پرسش می‌تواند در پاسخ‌دهی مؤثر باشد.

^۳ Geospatial Information System (GIS)

^۴ Global Positioning System (GPS)

عصبی پیش‌خور به نوعی آنی عمل می‌کند. به این معنی که خروجی بلا فاصله بعد از ارائه یک ورودی تأیید می‌شود. نوعی خاص از شبکه‌های پیش‌خور، گروه شبکه‌های لایه‌ای هستند، که آن‌ها را شبکه‌های پرسپترون^۱ چند لایه می‌نامند. پرسپترون‌های شده‌اند. لایه وسط را لایه در لایه‌های جداگانه‌ای تنظیم شده‌اند. لایه وسط را لایه میانی یا پنهان می‌نامند و لایه‌های ورودی و خروجی نیز در این شبکه حضور دارند [۲۶].

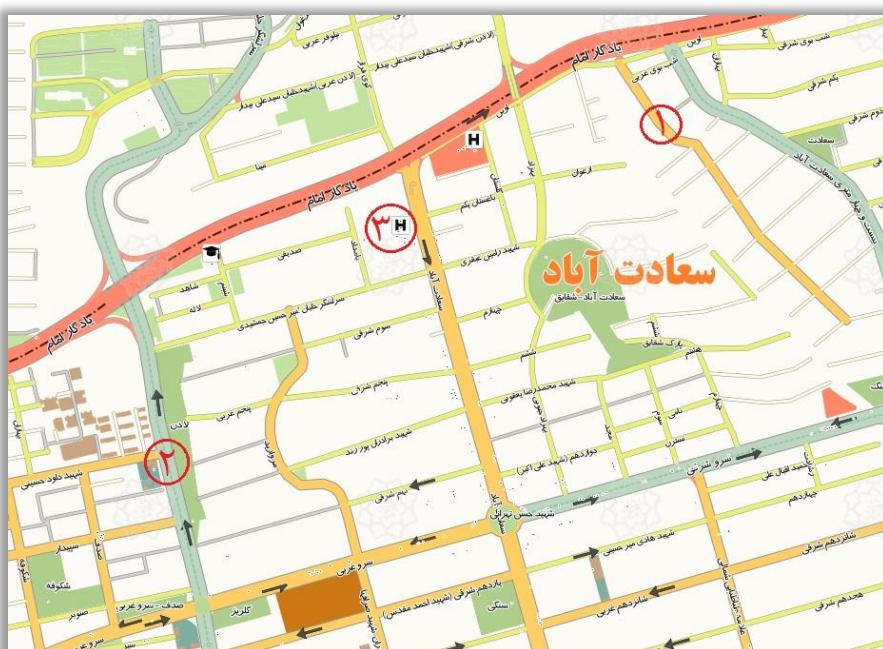
شبکه پرسپترون حداقل سه لایه دارد، یک لایه ورودی، یک لایه خارجی و یک یا چند لایه بین آن‌ها که مستقیماً به داده‌های ورودی و نتایج متصل نیست. در واقع این لایه را لایه پنهان می‌نامند. هر واحد در لایه پنهان و لایه خروجی مانند یک پرسپترون عمل می‌کند [۲۶]. قاعده فراگیری پرسپترون چند لایه را قاعده پس‌انتشار^۲ گویند. عمده‌ترین کاربرد قانون یادگیری پس‌انتشار، در آموزش شبکه‌های عصبی چند لایه پیش‌خور است که عموماً شبکه‌های چند لایه پرسپترون هم نامیده می‌شوند. الگوریتم یادگیری پس‌انتشار بر قانون یادگیری اصلاح خطای مبتنی می‌باشد. به عبارتی تپیلوژی شبکه‌های چند لایه پرسپترون با قانون یادگیری پس‌انتشار تکمیل می‌گردد. این قانون از دو مسیر اصلی تشکیل شده است. مسیر اول موسوم به مسیر رفت می‌باشد. در این مسیر، بردار ورودی به شبکه اعمال شده، تأثیراتش از طریق لایه میانی، به لایه خروجی انتشار می‌یابد. بردار خروجی ایجاد شده در لایه خروجی، پاسخ واقعی شبکه می‌باشد. در این مسیر، پارامترهای شبکه ثابت و بدون تغییر در نظر گرفته می‌شوند. مسیر دوم، موسوم به مسیر برگشت می‌باشد. در این مسیر، بر عکس مسیر رفت، پارامترهای شبکه تغییر پیدا کرده و تنظیم می‌شوند. این تنظیم مطابق با قانون اصلاح خطای صورت می‌گیرد. سیگنال خطای در لایه خروجی شبکه تشکیل می‌شود. بردار خطای برابر است با

^۱ Perceptron Network

^۲ Back Propagation Algorithm

نظر گرفته شد. یک گزینه پاسخ صحیح، یک گزینه پاسخ غلط و یک گزینه پاسخ «نمی‌دانم». همچنین نحوه آشنایی با منطقه نیز پرسیده شد. برای نحوه آشنایی با منطقه، چهار گزینه محل زندگی، محل کار، محل تردد و هیچ کدام در نظر گرفته شد. هیچ کدام حالتی پنداشته می‌شود که فرد آشنایی خوبی با منطقه ندارد. شکل (۲) تغییرات در نقشه منطقه سعادت آباد را نشان می‌دهد.

برای طراحی سوال‌های مربوط به نقشه، از نقشه شهر تهران در وبسایت نقشه تهران (<https://map.tehran.ir>) که توسط شهرداری تهران تهیه شده است، استفاده گردید. برای این منظور، ۱۷ منطقه از شهر تهران که پر تردد بودند، به صورتی که در سطح شهر تهران پراکندگی مناسبی داشته باشند، انتخاب گردید. در هر منطقه، محدوده‌ای در نظر گرفته و سه تغییر ایجاد شد. محل مورد تغییر در نقشه با شماره سؤال مربوطه مشخص و برای پاسخ آن سه گزینه در



شکل ۲: نقشه تغییر یافته منطقه سعادت آباد

میان گزینه‌های موجود، گزینه مورد نظر خود را انتخاب نماید. سپس از میان مناطق موجود، منطقه‌ای که آشنایی بیشتری با آن دارد را برگزیند. کاربر باید حداقل به سؤال‌های سه منطقه پاسخ دهد. پس از انتخاب منطقه، کاربر وارد صفحه آن منطقه شده و به سؤال‌های آن پاسخ می‌دهد. در شکل‌های (۳) و (۴) به ترتیب صفحه سؤال‌های پرسش‌نامه و صفحه پاسخ‌گویی به سؤال‌های منطقه سعادت آباد نشان داده شده است.

۲-۳- طراحی ابزار کاربردی تلفن هوشمند با اپاینورتر

اپاینورتر یک محیط برنامه‌نویسی تحت وب و متن باز^۱ است. با استفاده از اپاینورتر، یک برنامه کاربردی برای سیستم عامل اندرورید تلفن‌های هوشمند طراحی گردید. این برنامه ابتدا یک نام به عنوان نام کاربری از کاربر دریافت می‌کند. سپس کاربر باید به سؤال‌های اولیه در مورد پیشینه یا اطلاعات زمینه‌ای خود پاسخ دهد و از

^۱ Open Source

(ب)

-انتخاب-	سن
جنس	-انتخاب-
شفل	جنسیت
زیر دیپلم	شفل
میزبان	میزان تحصیلات
دیپلم	گروه تحصیلی
گروه	آیا با رشته هندسه نقشه برداری آشنایی دارد؟
کاردادانی	بله <input type="checkbox"/> خیر <input type="checkbox"/>
آشنایی	آیا تحصیلات شما در زمینه هندسه نقشه برداری است؟
کارشناسی	بله <input type="checkbox"/> خیر <input type="checkbox"/>
آیا ت نقشه	آیا با سیستم GPS آشنایی دارد؟
کارشناسی ارشد	بله <input type="checkbox"/> خیر <input type="checkbox"/>
آیا با	آیا با سیستم GIS آشنایی دارد؟
دکتری تخصصی	بله <input type="checkbox"/> خیر <input type="checkbox"/>
آیا با	
دکتری حرفه ای	

صفحه بعد

(الف)

شکل ۳: (الف) صفحه پاسخ‌گویی به سوال‌های پیشینه‌ای پرسش‌نامه (ب) انتخاب مقطع تحصیلی براساس لیست طراحی شده



شکل ۴: صفحه پاسخ‌گویی به سوال‌های منطقه سعادت آباد

۴-۳- یافتن بهترین چیدمان شبکه عصبی مصنوعی

با توجه به مطالعه‌های صورت گرفته و نوع داده‌ها، شبکه عصبی پرسپترون چند لایه با ویژگی پسانشان، برای آموزش شبکه انتخاب گردید. به عبارتی، این شبکه در مرحله اول آموزش، حرکت رو به جلو دارد. سپس با مقایسه خروجی شبکه و خروجی واقعی، خطأ و انحراف شبکه محاسبه می‌شود و در مرحله بعد شبکه به عقب بازگشته و وزن‌ها و انحراف‌ها را بهنگام‌سازی می‌کند. تکرار ادامه پیدا می‌کند تا جایی که شبکه برای تمام داده‌های آموزشی به نزدیک‌ترین مقدار خود برسد. به این ترتیب، شبکه یاد گرفته و از این به بعد می‌تواند با مشاهده ویژگی‌های یک داده جدید، خروجی آن را تشخیص دهد.

برای طراحی شبکه از ابزار شبکه عصبی برنامه متلب استفاده شد. برای آموزش شبکه از روش آموزش دسته‌ای کاهش شیب با گشتاور^۲، برای تابع آموزش از الگوریتم لونبرگ-مارکوارد^۳، برای تابع عملکرد از خطای میانگین مربعات^۴ (*MSE*) و برای تابع فعال‌سازی، از تابع تانژانت سیگموئید^۵ استفاده گردید. شبکه عصبی در دو حالت، یکبار با یک لایه پنهان و یکبار با دو لایه پنهان طراحی شد. تعداد نرون‌ها برای هر لایه بین ۲ تا ۹ نرون و تعداد تکرار آموزش شبکه در هر حالت ۱۰ مرتبه در نظر گرفته شد. به عبارت دیگر، در حالت یک لایه پنهان، ۸۰ مرتبه آموزش شبکه و در حالت دو لایه پنهان، ۶۴۰ مرتبه آموزش شبکه، مجموعاً ۷۲۰ مرتبه آموزش شبکه در نظر گرفته شد. همچنین به صورت تصادفی، ۷۰٪ داده‌ها برای آموزش شبکه، ۱۵٪ داده‌ها برای اعتبارسنجی شبکه و ۱۵٪ داده‌ها برای آزمایش شبکه در نظر گرفته شد. در هر مرتبه آموزش شبکه، ترکیب داده‌های آموزش، اعتبارسنجی و آزمایش شبکه،

۳-۳- جمع‌آوری و آماده سازی داده‌ها

پس از طراحی کامل برنامه، فایل نصب برنامه کاربردی طراحی شده بین افراد مختلف در شبکه‌های اجتماعی قرار گرفت تا کاربران از هر رده سنی، شغلی و تحصیلی، به صورت داوطلبانه، به سؤال‌های این برنامه پاسخ دهند. پس از جمع‌آوری، داده‌ها، باید برای ورود به شبکه عصبی آماده گردند. روش درنظر گرفته شده برای این کار، بی‌مقیاس کردن داده‌ها می‌باشد. برای این منظور، ابتدا داده‌ها در محیط برنامه اکسل^۱ وارد شده و براساس نام کاربری و منطقه پاسخ داده شده، مرتب شدند. سپس هر متغیر براساس حالت‌های مختلفی که برای آن درنظر گرفته شده، بین بازه ۰ تا ۱ بی‌مقیاس شدند. به این معنی که برای هر حالت متغیر، بین عدد ۰ تا ۱، یک عدد بصورت کد درنظر گرفته شد که هر کد نمایانگر حالت خاصی از متغیر است. سپس با توجه به پاسخ صحیح سؤال‌های مربوط به نقشه مناطق، به پاسخ کاربران امتیاز داده شد. برای پاسخ صحیح عدد ۱، پاسخ غلط عدد ۰ و برای پاسخ نمی‌دانم عدد ۰. در نظر گرفته شد. میانگین پاسخ‌ها محاسبه و به صورت درصد پاسخ صحیح در یک ستون جدید ذخیره گردید. به این ترتیب پاسخ‌های داده شده در بازه ۰ تا ۱۰۰٪ قرار گرفتند. ۱۰۰٪ به معنی پاسخ صحیح به هر ۳ سؤال و ۱۰۰٪ به معنی پاسخ غلط به هر ۳ سؤال می‌باشد. مجموعاً ۲۷۵ نفر از افراد مختلف جامعه در این طرح شرکت کردند و هر فرد حافظل به پرسش‌های سه منطقه پاسخ داده است. از آنجا که نحوه آشنایی با هر منطقه (که جزء سؤال‌ها می‌باشد) متفاوت بوده و در پاسخ کاربران تأثیرگذار است، می‌توان پاسخ به پرسش‌های هر منطقه را به عنوان یک نمونه مجزا در نظر گرفت. بنابراین مجموعاً ۱۱۰۲ نمونه جمع‌آوری، مرتب و بی‌مقیاس گردیدند.

^۴ Mean Squared Error (*MSE*)

^۵ Tangent Sigmoid

^۱ Excel

^۲ Batch Gradient Descent with Momentum

^۳ Levenberg-Marquardt (LM)

حالت، یک لایه پنهان و دو لایه پنهان، با ۲ تا ۹ نرون و هر بار ۱۰ مرتبه آموزش، مجموعاً ۷۲۰ مرتبه آموزش داده شده و هر بار نتایج MSE آموزش ثبت گردید. در هر مرحله میانگین MSE مربوط به ۱۰ بار آموزش شبکه محاسبه گردید. بهترین عملکرد برای شبکه سه لایه، با ۷ نرون در لایه پنهان اول و ۴ نرون در لایه پنهان دوم و مقدار میانگین ۰,۱۹۹۸۰ به دست آمد.

جدول (۱)، مقادیر میانگین MSE بهترین عملکرد اعتبارسنجی شبکه در دو حالت معماری مختلف شبکه را نشان می‌دهد. شکل (۵) معماری شبکه عصبی طراحی شده منتخب و شکل (۶)، نمودار بهترین عملکرد اعتبارسنجی در بدست آوردن MSE در حالت منتخب را نشان می‌دهند.

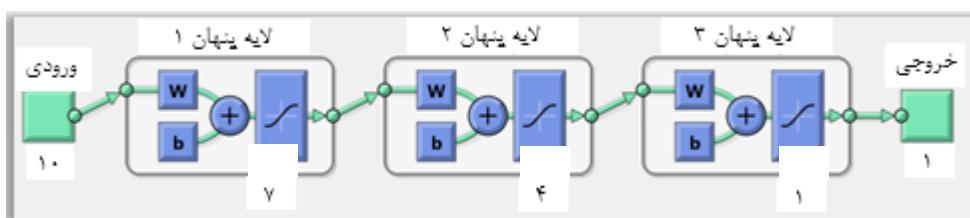
به صورت تصادفی توسط ابزار شبکه عصبی متلب تغییر داده شد. در هر بار تکرار آموزش برای یک معماری، مقدار MSE بهترین عملکرد اعتبارسنجی شبکه ذخیره گردیده و میانگین آن برای ۱۰ بار تکرار آموزش، محاسبه شده تا بهترین حالت معماری شبکه مشخص شود.

۵-۳- محاسبات

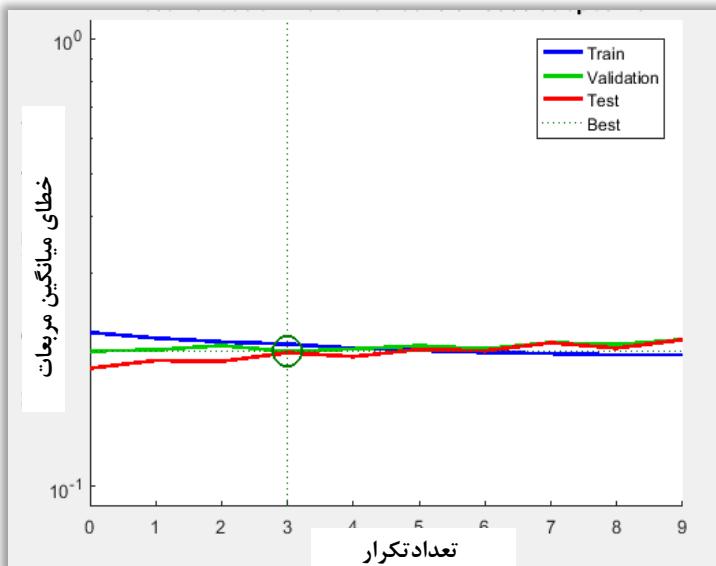
داده‌های کدگذاری شده به دو دسته به صورت جداگانه تقسیم شدند. یک دسته داده ورودی (*Input Data*) شامل اطلاعات جمع آوری شده و دسته دوم داده هدف (*Data Target*) که فقط شامل درصد پاسخ صحیح می‌باشد. داده‌های ورودی و هدف به شبکه عصبی طراحی شده معرفی شدند. سپس شبکه عصبی در دو

جدول ۱: مقادیر میانگین MSE بهترین عملکرد اعتبارسنجی شبکه در دو معماری مختلف شبکه

نوع معماری شبکه	تعداد نرون	میانگین MSE شبکه در ۱۰ بار تکرار آموزش	دو لایه (یک لایه پنهان)	سه لایه (دو لایه پنهان)
	تعادل نرون		۸	۷ و ۴
		۰,۱۹۹۸۰	۰,۲۰۹۳۲	



شکل ۵: معماری شبکه عصبی سه لایه با دو لایه پنهان و ۷ نرون در لایه پنهان اول و ۴ نرون در لایه پنهان دوم

شکل ۶: نمودار بهترین عملکرد اعتبارسنجی با شاخص MSE

اطلاعات داوطلبانه مربوط به سامانه‌های پرکاربرد نباشد و یا مربوط به مناطق نسبتاً دورافتاده باشد و یا توسط کاربرانی با مشارکت کم تعداد تهیه شده باشد، روش‌های مزبور کارایی ندارد. روش مورد استفاده در این تحقیق کوشیده است ارتباطی بین اطلاعات زمینه‌ای افراد و صحت اطلاعات نقشه‌ای آنان برقرار نماید. این روش در مطالعات پیشین مسحوق به سابقه است [۲۹ و ۳۰]. اما در این تحقیق اطلاعات زمینه‌ای بیشتری مورد استفاده قرار گرفته و روش سنجش اطلاعات نقشه‌ای افراد نیز کاملاً ابتکاری بوده است. مضافاً اینکه از شبکه عصبی مصنوعی برای نیل به این هدف استفاده شده است. همچنین گفتنی است که صحت اطلاعاتی که در این پژوهش سنجیده شده است می‌تواند متأثر از دو منشأ باشد که یکی میزان اطلاعات افراد از منطقه‌ای است که به صورت داوطلبانه انتخاب کرده‌اند و دومی میزان صداقت آنها در درج اطلاعات صحیح است. بدیهی است که ممکن است قلیلی از افراد به انگیزه‌های مختلف دست به تولید اطلاعات مکانی ناصحیح بزنند. در هر صورت نمی‌توان تعیین نمود که ناصحیح بودن اطلاعات وارد شده به دلیل اطلاعات ناقص بوده است و یا قصدی

۴- نتایج و بحث

اطلاعات مکانی داوطلبانه سبب کاهش بسیاری از مشکلات ناشی از نبود و کمبود داده شده است. با این حال ارزیابی کیفیت و میزان اعتمادپذیری این داده‌ها همواره یک چالش اساسی بوده است. تنها راه مطمئن برای یافتن کیفیت این نوع اطلاعات، مقایسه آنها با داده‌های رسمی است اما در بسیاری از موارد داده‌های رسمی متناظر با اطلاعات مکانی داوطلبانه وجود ندارد چه اینکه ارزش بسیاری از این داده‌ها در پر کردن خلاصه‌های رسمی می‌باشد. در اینگونه موارد راه حل‌های گوناگون و بعضی ابتکاری به مسئله باز می‌شود. یکی از این راه حل‌ها که این تحقیق به دنبال آن بوده است، تعیین میزان اعتبار تولیدکنندگان اطلاعات مکانی داوطلبانه است. در واقع باید سنجید که هر داوطلب، تا چه حد اطلاعات دقیق و صحیحی وارد نموده است. در سامانه‌های شناخته شده *VGI* نظری *OSM* می‌توان میزان اعتبار اطلاعات وارد شده توسط کاربران را به شیوه‌های مختلفی از جمله مقایسه با داده‌های رسمی، مقایسه با یکدیگر و یا تأیید اطلاعات توسط کاربران دیگر، با دقت خوبی تعیین نمود. با این حال چنانچه

هدف از این کار این است که بتوان حدس زد چه کاربری با چه اطلاعات زمینه‌ای ممکن است بالاترین یا پایین‌ترین صحت اطلاعات ورودی را داشته باشد. برای این کار پس از به دست آوردن بهترین حالت شبکه، ماتریسی از تمام حالت‌های مختلف متغیرهای ورودی به صورت جایگشت تشکیل داده شد. با توجه به حالت‌های مختلفی که برای پاسخ متغیرها در نظر گرفته شده، نهایتاً ماتریسی با ۳۲۵۶۰ ردیف و ۱۰ ستون تشکیل شد. از این ماتریس که از جایگشت به دست آمده است، حالت‌های نامتعارف، به عنوان مثال مهندس نقشه‌بردار با مدرک تحصیلی دیپلم و سن کمتر از ۱۸ سال، از میان داده‌ها حذف گردیدند و باقی موارد به عنوان ورودی به شبکه منتخب معرفی شده و خروجی آن، یعنی درصد پاسخ صحیح به دست آمد. سپس متغیرهای ورودی و خروجی به ترتیب بالاترین درصد پاسخ صحیح، مرتب شدند. بالاترین و پایین‌ترین درصد پاسخ صحیح برای افرادی با مشخصات مذکور مطابق جدول (۲) به دست آمد. بنابر نتایج، اگر فردی برای اولین بار مشخصات اولیه خود را مطابق پارامترهایی که شبکه عصبی منتخب با آن‌ها آموزش دیده است، وارد سامانه *VGI* مربوطه نماید، شبکه عصبی آموزش دیده منتخب می‌تواند درصد پاسخ صحیح فرد مذکور را برآورد نماید. درواقع درصد پاسخ صحیح برآورده شده میزان اعتقادپذیری شخص را مشخص می‌کند.

بررسی‌هایی که روی اطلاعات خام جمع‌آوری شده انجام شد نشان داد که نحوه آشنایی با منطقه بیشترین تأثیر را بر صحت اطلاعات پاسخ‌های افراد دارد. در ادامه نتایج حاصل از شبکه عصبی مصنوعی طبق جدول (۲) مشخص کرد که کسانی که نحوه آشنایی آنها با منطقه از طریق محل کار بوده است پاسخ‌های صحیح‌تری داده‌اند در حالیکه محل سکونت با اختلاف کم در رتبه بعدی است و سپس محل عبور و هیچ‌کدام قرار گرفته‌اند. بر اساس داده‌های جمع‌آوری شده، عوامل زمینه‌ای دیگر به جز سن، تأثیر ملموسی روی صحت پاسخ‌ها نداشتند و لذا اطمینان به نتایج آنها طبق

در آنها نهفته بوده ولی اطلاعاتی که در این تحقیق حاصل شده حاوی هردوی این موارد به صورت توان بوده است. البته به صورت کلی باید اذعان داشت که درصد اطلاعاتی که عمداً به صورت اشتباه تولید می‌شوند بسیار اندک است.

شبکه عصبی انتخاب شده در این پژوهش پس از آموزش قادر خواهد بود برآورده از میزان صحت اطلاعات یک کاربر که حتی برای اولین بار مبادرت به تولید اطلاعات مکانی داوطلبانه می‌کند ارائه نماید و این کار را بر اساس اطلاعات زمینه‌ای (پیشینه) فرد انجام می‌دهد. یکی از چالش‌ها در این مورد آن خواهد بود که ممکن است افراد از ارائه اطلاعات زمینه‌ای خود امتناع نمایند و یا اینکه آن را به درستی وارد نکنند. این چالشی اساسی است چراکه سامانه‌های معروف *VGI* همچون *OSM* الزاماً برای ورود اطلاعات زمینه‌ای افراد، آن هم با این جزئیات، قرار نداده‌اند. به هر حال ممکن است سامانه‌های دیگر و یا سامانه‌های فعلی در زمان آینده، وجود چنین الزاماً را ضروری تشخیص دهند. صحت اطلاعات زمینه‌ای را نیز می‌توان با روش‌هایی از جمله استفاده از پایگاه‌های اطلاعاتی موجود، ارزیابی نمود. به هر صورت چنانچه اطلاعات زمینه‌ای به صورت گستردگی جمع‌آوری شود؛ به تدریج ارتباط بین اطلاعات زمینه‌ای افراد و صحت اطلاعات مکانی وارد شده توسط آنان پرنگ‌تر می‌شود و از سویی می‌توان با داده‌کاوی، اطلاعات ذیقیمت دیگری را نیز استخراج کرد. ممکن است نبود موارد اطلاعات زمینه‌ای یک فرد در میان گزینه‌ها نیز به عنوان یک چالش تلقی شود ولی این مورد با توجه به انعطاف برنامه به راحتی قابل حل است و چنانچه این روش به صورت گسترده مورد استفاده قرار گیرد به تدریج انواع اطلاعات زمینه‌ای افراد در آن گنجانده خواهد شد. بنابراین نتیجه اول این تحقیق قابلیت برآورد صحت اطلاعات مکانی ورودی کاربران است.

از طرف دیگر با توجه به داده‌های جمع‌آوری شده می‌توان به بررسی جامعه آماری مورد استفاده پرداخت.

نتایج به دست نمی‌تواند این واقعیت را انکار نماید که مقدار اطلاعات جمع‌آوری شده (تعداد مشارکت-کنندگان) علیرغم تلاش بسیار پژوهشگران، برای انجام نتیجه‌گیری‌های محقق، ناکافی است.

جدول (۲) زیاد نیست. تأثیر سن کمی پررنگ‌تر از سایر عوامل زمینه‌ای بود و همانطور که جدول (۲) هم نشان می‌دهد سن‌های بالا پاسخ‌های ناصحیح را بیشتر می‌کند.

جدول ۲: بالاترین و پایین‌ترین درصد پاسخ صحیح پیش‌بینی شده توسط شبکه با داده‌های ساخته شده

درصد پاسخ صحیح	نحوه آنلاین با منطقه	آنلاین با GPS	آنلاین با GIS	مهندس نقشه برداری	آنلاین با نقشه برداری	گروه تحصیلی	قطعه تحصیلی	شغل	جنسیت	سن
۹۹,۷۹	محل کار	آشنا هستم	آشنا هستم	مهندس نقشه بردار نیستم	آشنا با نقشه برداری هستم	فنی و مهندسی	کارشناسی	معلم	مرد	۴۰-۵۰ سال
-۸۴,۳۲	هیچکدام	آشنا نیستم	آشنا نیستم	مهندس نقشه بردار نیستم	آشنا با نقشه برداری نیستم	علوم پایه	زیر دیپلم	سایر	مرد	۶۰-۷۰ سال

VGI نمادین و کوچک می‌باشد. درواقع هدف این است که از این نوآوری و نتایج آن برای ارزیابی کاربران سیستم‌های *VGI* بزرگ و واقعی که بسیار کاربردی هستند و روزانه در سراسر دنیا اطلاعات زیادی در آن‌ها ایجاد، ویرایش یا حذف می‌شود (*Mannd OSM*) استفاده گردد. اگر در یک سیستم *VGI* کاربری در ابتدای ورود به صفحه کاربری خود، مشخصاتی از خود را وارد نماید و توسط شبکه عصبی آموزش داده شده میزان اعتمادپذیری وی تخمین زده شود، سیستم می‌تواند سطح دسترسی مناسب با ارزیابی صورت گرفته برای ایجاد تغییرات را به او بدهد. این روش میزان اعتمادپذیری به اطلاعات تولید شده در سیستم *VGI* را افزایش می‌دهد.

۵- نتیجه‌گیری و پیشنهادها

آنچه از تحقیقات پیشین به دست می‌آید این است که *VGI* اغلب ارزان‌ترین منبع و گاهی تنها منبع اطلاعات مکانی می‌باشد. از طرفی به دلیل اینکه این اطلاعات توسط افراد تهیه می‌شوند که اکثرًا تجربه و دانش

نتایج حاصل از اجرای شبکه عصبی مصنوعی هم مؤید همین مطلب است و همانطور که شکل (۶) نشان داده است، شبکه در تکرار سوم به حالت تقریباً بهینه خود رسیده است. یقیناً افزایش داده‌ها که طبعاً باعث افزایش نوع آنها هم خواهد شد، شبکه عصبی را به کارایی بهتری خواهد رساند. اما سیستم طراحی شده برای جمع‌آوری اطلاعات با موفقیت کار می‌کند و افرادی حاضر شده‌اند که این نرم‌افزار را بر روی تلفن همراه خود نصب و از آن استفاده نمایند. برای استخراج نتایج مطمئن، لازم است که این اطلاعات از هزاران نفر با اطلاعات زمینه‌ای متنوع جمع‌آوری شود اما از آنجا که این سیستم مستقل از سامانه‌های داوطلبانه است، یافتن داوطلب برای آن بسیار دشوار خواهد بود. راه حلی که می‌تواند در این وادی راهگشا باشد، گنجاندن این برنامه در قالب برنامه‌های پرکاربرد تلفن همراه نظریه بازی‌ها است. بدین ترتیب می‌توان اطلاعات بسیار زیاد و پرتنوعی را جمع‌آوری و نتایج احتمالاً ذیقیمتی را از آنها حاصل نمود.

آنچه در این پژوهش انجام شده بر روی یک سیستم

این شبکه متناسب با داده‌های جمع‌آوری شده طراحی گردید. در ادامه داده‌ها برای ورود و آموزش شبکه آماده سازی شدند. با استفاده از داده‌های جمع‌آوری شده، شبکه عصبی مصنوعی در چیدمان‌های مختلفی از نرون‌ها آموزش داده شد و بهترین چیدمان با MSE برابر با 19988% انتخاب شد. سپس از شبکه منتخب برای تخمین کیفیت اطلاعات *VGI* کاربران جدید استفاده گردید. این پژوهش بر روی یک نمونه کوچک و نمادین از سامانه *VGI* انجام شده و قابل بسط به سامانه‌های بزرگ و کاربردی است. به این ترتیب که هرگاه کاربری جدید که بخواهد وارد یک سامانه *VGI* شود و یا برای اولین بار شروع به ورود اطلاعات مکانی یا مرتبط با مکان در آن نماید، با دانستن مشخصات اولیه او (اطلاعات زمینه‌ای) و با استفاده از شبکه عصبی آموزش دیده، می‌توان تخمینی از میزان اعتبار داده‌های ورودی توسط او در اختیار داشت. بنابراین سامانه *VGI* می‌تواند سطح دسترسی متناسب با سطح اعتبار برآورد شده را در اختیار کاربر قرار دهد و به این صورت اعتبار بیشتری برای داده‌های داوطلبانه وارد شده حاصل گردد.

با توجه به مشکل اصلی این تحقیق که جمع‌آوری داده‌ها بود نهایتاً 110^2 رکورد داده جمع‌آوری شد، ولی کماکان استحصال داده‌های بیشتر برای رسیدن به نتایج قابل اطمینان‌تر در تحقیقات آینده ضروری است. با در نظر گرفتن پژوهش‌های پیشین و نتایج حاصل از این تحقیق، پیشنهاد می‌شود در پژوهش‌های آینده در زمینه نحوه آشنایی با منطقه، عوامل بیشتری در نظر گرفته شود. همچنین برای پاسخ به سوال‌های پرسش‌نامه می‌توان از حالت فازی نیز استفاده گردد.

تهیه اطلاعات مکانی را ندارند، کیفیت آن‌ها دارای ابهام است و از این رو ارزیابی این کیفیت اهمیت زیادی پیدا می‌کند. تاکنون روش‌های مختلفی برای ارزیابی کیفیت *VGI* مورد بررسی قرار گرفته است که به صورت کلی به دو دسته تقسیم می‌شوند. دسته اول که کیفیت اطلاعات *VGI* را با منابع رسمی و با کیفیت معلوم مقایسه می‌کند و دسته دوم که از مشخصات ذاتی اطلاعات *VGI* همچون کاربران، روش‌های آماری و غیره استفاده می‌کند. به این دلیل که منابع رسمی همیشه در دسترس نیستند و حتی گاهی اطلاعات *VGI* کامل تراز منابع رسمی هستند؛ روش‌های دسته دوم، تنوع و دامنه کاربرد وسیع‌تری دارند.

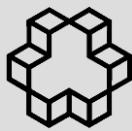
این تحقیق بر آن بود تا روشی برای ارزیابی کیفیت اطلاعات *VGI* در قالب دسته دوم از طریق ارزیابی کاربران *VGI* به کمک شبکه عصبی مصنوعی ارائه دهد. در این روش اطلاعات زمینه‌ای کاربران اهمیت دارد. از طرفی اطلاعات زمینه‌ای کافی از کاربران *VGI* وجود ندارد یا دسترسی به آن‌ها برای کاربران عادی غیر ممکن است. بنابراین جهت شبیه‌سازی سامانه *VGI* و جمع‌آوری اطلاعات زمینه‌ای کاربران آن، یک برنامه کاربردی برای تلفن‌های هوشمند مشابه سامانه *VGI* طراحی گردید و از طریق شبکه‌های اجتماعی در اختیار کاربران قرار گرفت. در این برنامه، از نقشه ۱۷ ناحیه در شهر تهران استفاده شد و اطلاعات زمینه‌ای کاربر شامل سن، جنسیت، شغل، مقطع تحصیلی، گروه تحصیلی، آشنایی با نقشه‌برداری، مهندس نقشه‌برداری بودن، آشنایی با *GPS*. آشنایی با *GIS* و نحوه آشنایی با منطقه جمع‌آوری گردید. سپس برای ارزیابی کیفیت کاربران *VGI*، از شبکه عصبی مصنوعی استفاده شد و

مراجع

- [1] M. F. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211-221, 2007.
- [2] J. Severinsen, M. de Roiste, F. Reitsma, and E. Hartato, "VGTrust: measuring trust for volunteered geographic information," *International Journal of Geographical Information Science*, vol. 33, no. 8, pp. 1683-1701, 2019.
- [3] M. Drews et al., "The utility of using Volunteered Geographic Information (*VGI*)

- for evaluating pluvial flood models," *Science of The Total Environment*, vol. 894, p. 164962, 2023/10/10/ 2023, doi: <https://doi.org/10.1016/j.scitotenv.2023.164962>.
- [4] Z. Chang et al., "An updating of landslide susceptibility prediction from the perspective of space and time," *Geoscience Frontiers*, vol. 14, no. 5, p. 101619, 2023/09/01/ 2023, doi: <https://doi.org/10.1016/j.gsf.2023.101619>.
- [5] C. C. Fonte, L. Bastin, L. See, G. Foody, and F. Lupia, "Usability of VGI for validation of land cover maps," *International Journal of Geographical Information Science*, vol. 29, no. 7, pp. 1269-1291, 2015/07/03 2015, doi: [10.1080/13658816.2015.1018266](https://doi.org/10.1080/13658816.2015.1018266).
- [6] M. Moradi, S. Roche, and M. A. Mostafavi, "Exploring five indicators for the quality of OpenStreetMap road networks: a case study of Québec, Canada," *Geomatica*, pp. 1-31, 2022, doi: [10.1139/geomat-2021-0012](https://doi.org/10.1139/geomat-2021-0012).
- [7] M. Ahmad, M. S. H. Khayal, and A. Tahir, "Analysis of Factors Affecting Adoption of Volunteered Geographic Information in the Context of National Spatial Data Infrastructure," *ISPRS International Journal of Geo-Information*, vol. 11, no. 2, p. 120, 2022.
- [8] V. Antoniou and A. Skopeliti, "Measures and indicators of VGI quality: An overview," *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, vol. 2, p. 345, 2015.
- [9] C. Dai, D. Lin, E. Bertino, and M. Kantarcioglu, "An approach to evaluate data trustworthiness based on data provenance," in *Workshop on Secure Data Management*, 2008: Springer, pp. 82-98.
- [10] M. F. Goodchild and L. Li, "Assuring the quality of volunteered geographic information," *Spatial statistics*, vol. 1, pp. 110-120, 2012.
- [11] H. Senaratne, A. Mobasher, A. L. Ali, C. Capineri, and M. Haklay, "A review of volunteered geographic information quality assessment methods," *International Journal of Geographical Information Science*, vol. 31, no. 1, pp. 139-167, 2017.
- [12] M. Forghani and M. R. Delavar, "A quality study of the OpenStreetMap dataset for Tehran," *ISPRS International Journal of Geo-Information*, vol. 3, no. 2, pp. 750-763, 2014.
- [13] N. Mohammadi and M. Malek, "Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data," *Journal of Spatial Science*, vol. 60, no. 1, pp. 119-135, 2015/01/02 2015, doi: [10.1080/14498596.2014.927337](https://doi.org/10.1080/14498596.2014.927337).
- [14] G. Farajolahi and M. R. Delavar, "Provide a model for quality control of local spatial information for the analysis of traffic accident locations," presented at the The 2nd National Conference on Geospatial Information Technology, Tehran, Iran, 2016. [Online]. Available: <http://fa.seminars.sid.ir/ViewPaper.aspx?ID=36253>.
- [15] A. Yang, H. Fan, and N. Jing, "Amateur or professional: Assessing the expertise of major contributors in OpenStreetMap based on contributing behaviors," *ISPRS International Journal of Geo-Information*, vol. 5, no. 2, p. 21, 2016.
- [16] A. L. Ali and F. Schmid, "Data quality assurance for volunteered geographic information," in *International Conference on Geographic Information Science*, 2014: Springer, pp. 126-141.
- [17] B. Vahedi Torgabe and A. A. Alesheykh, "Assessing the Attribute Accuracy of Volunteered Geographic Information," (in eng), *Journal of Geomatics Science and Technology, Research* vol. 5, no. 3, pp. 49-64, 2016. [Online]. Available: <http://jgst.issge.ir/article-1-348-fa.html>.
- [18] M. Bishr and K. Janowicz, "Can we trust information?-the case of volunteered geographic information," in *Towards Digital Earth Search Discover and Share Geospatial Data Workshop at Future Internet Symposium*, volume, 2010, vol. 640.

- [19] A. M. Forati and M. Karimipour, "Evaluate the quality of voluntary spatial information by accrediting users and their social groups," presented at the Conference on Surveying & Spatial Information, Tehran, 2016. [Online]. Available: <http://fa.seminars.sid.ir/ViewPaper.aspx?ID=89018>.
- [20] A. Khosravi Kazazi and F. HoseinAli, "Increasing the quality of voluntary spatial information results based on EigenRumor algorithm," presented at the Conference on Surveying & Spatial Information, Tehran, 2017. [Online]. Available: <http://fa.seminars.sid.ir/ViewPaper.aspx?ID=90158>.
- [21] J. N. Da Costa, "Novel tool for examination of data completeness based on a comparative study of VGI data and official building datasets," *Geodetski Vestnik*, vol. 60, no. 3, pp. 495-508, 2016.
- [22] R. Martella, E. Clementini, and C. Kray, "Crowdsourcing geographic information with a gamification approach," *Geodetski Vestnik*, vol. 63, no. 2, 2019.
- [23] R. Can, S. Kocaman, and C. Gokceoglu, "A convolutional neural network architecture for auto-detection of landslide photographs to assess citizen science and volunteered geographic information data quality," *ISPRS International Journal of Geo-Information*, vol. 8, no. 7, p. 300, 2019.
- [24] A. Alghanim, M. Jilani, M. Bertolotto, and G. McArdle, "Leveraging Road Characteristics and Contributor Behaviour for Assessing Road Type Quality in OSM," *ISPRS International Journal of Geo-Information*, vol. 10, no. 7, p. 436, 2021. [Online]. Available: <https://www.mdpi.com/2220-9964/10/7/436>.
- [25] W. Shi, Z. Liu, Z. An, and P. Chen, "RegNet: a neural network model for predicting regional desirability with VGI data," *International Journal of Geographical Information Science*, vol. 35, no. 1, pp. 175-192, 2021/01/02 2021, doi: 10.1080/13658816.2020.1768261.
- [26] S. Soltani, S. Sardari, M. S. Pour, and S. Mousavi, *Artificial Neural Networks: Basics, Applications, and Introduction to Easy NN-plus and NeuroSolutions Software*. Tehran: Nass, 2010, p. 216.
- [27] R. J. Shalkof, *Artificial Neural Networks*. Ahvaz: Shahid Chamran University of Ahvaz, 2003.
- [28] L. Bragagnolo, R. V. da Silva, and J. M. V. Grzybowski, "Landslide susceptibility mapping with r.landslide: A free open-source GIS-integrated tool based on Artificial Neural Networks," *Environmental Modelling & Software*, vol. 123, p. 104565, 2020/01/01/ 2020, doi: <https://doi.org/10.1016/j.envsoft.2019.104565>.
- [29] P. Fogliaroni, F. D'Antonio and E. Clementini, "Data trustworthiness and user reputation as indicators of VGI quality," *Geo-spatial Information Science*, vol. 21, no. 3, pp. 213-233, 2018.
- [30] J.C.H. Meier, "An Analysis of Quality for Volunteered Geographic Information," M.S Thesis, Wilfrid Laurier University, Ontario, Canada, 2015.



Quality Evaluation of Volunteered Geographic Information (VGI) by assessing user's reliability, using Artificial Neural Networks (ANN)

Elahe Azariasgari^{1}, Farhad Hosseinali²*

1- Department of Geomatics Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

2- Department of Surveying Engineering, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

Abstract

Nowadays Volunteered Geographic Information (VGI) is an important source of geographic information. Unlike the formal data, precision and accuracy of VGI is not known at the time of data gathering. Thus, several methods have been developed to evaluate the VGI quality. One of them is assessing the VGI quality by evaluating the reliability (trustworthiness) of VGI participants. In this study, by using the background information of VGI participants and Artificial Neural Networks (ANN), the user's reliability in producing spatial information is estimated. To collect the user's background information, a mobile application was designed under the Android operating system. In this program, the map of Tehran was used and some changes were applied to some of its parts. When using this program, the users must answer questions such as gender, age, education, familiarity with GPS or GIS, etc. Then the users should answer the questions about the changes made to the map. All of the answers are compared with the correct ones. Then the percentage of the user's correct answers is calculated. Each user should answer the questions of at least three regions. Finally, this information was collected for 1102 regions. From this data 70% is for the neural network training, 15% for validation and 15% for testing. The ANN which is a feed forward back propagation multilayer perceptron network, was trained by various number of neurons and hidden layers. The best network with mean squared error value 0.19988 was selected. Using the trained ANN, it is possible that in a VGI system, a new user enters his/her background information and the percentage of his/her predicted correct responses be estimated. This percentage may be assumed as one of the criteria of the user's reliability in VGI.

Key words : Volunteered Geographic Information (VGI), Geospatial Information System (GIS), Artificial Neural Network (ANN).

Correspondence Address: Department of Geomatics Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran.
Tel: : 02133722831.
Email : st_e_azariasgari@azad.ac.ir