

استفاده از گزارش شهروندان در پیش‌بینی سیلاب شهری با رویکرد یادگیری ماشین و داده تاریخی بارش (مطالعه مورد: شهر قم)

۱. فرشته السادات موسوی پور

دانشجوی دکتری، پردیس بین‌المللی کیش دانشگاه تهران، کیش، ایران

۲. محمدحسین نیک سخن (نویسنده مسئول)

استاد، دانشکده محیط زیست، دانشگاه تهران، تهران ایران^۱

۳. جابر سلطانی

دانشیار، دانشکده‌گان کشاورزی و منابع طبیعی، دانشکده فناوری کشاورزی (ابوریحان)، دانشگاه تهران، تهران، ایران

چکیده

پژوهش حاضر با هدف توسعه روشی علمی و کاربردی برای پیش‌بینی سیلاب‌های شهری، از گزارش‌های مردمی و داده‌های ماهواره‌ای بارش بهره گرفته است. محدوده شهر با استفاده از یک شبکه مربعی با فواصل ۳۰۰۰ متر تقسیم و نقاط برخورد خطوط شبکه به عنوان ایستگاه برداشت بارندگی انتخاب شدند. روزهای تر، خشک و عادی بر اساس گزارش‌های شهروندان و مشاهده بارش، دسته‌بندی گردیده و مدل‌های یادگیری ماشینی شامل ماشین بردار پشتیبان، رگرسیون لجستیک و جنگل تصادفی آموزش داده شدند. نتایج نشان داد که مدل جنگل تصادفی با دقت ۹۸ درصد و امتیاز F1، ۶۶ درصد در شناسایی رویدادهای سیلاب، برترین مدل می‌باشد. با توجه به اقلیم نیمه‌بیابانی منطقه مورد مطالعه (شهر قم) و خسارات غیر جانی سیلاب، تاکید ارزیابی مدل‌ها بر دقت در شناسایی درست موارد سیلاب جهت افزایش اطمینان و عدم ایجاد تشویش در جامعه به دلیل هشدارهای نادرست است. بر اساس تحلیل مهم‌ترین ویژگی‌ها، ایستگاه‌های ۱ و ۲۳ به عنوان مکان‌های کلیدی برای نصب تجهیزات بارش شناخته شدند. سپس، نمودار آستانه بارش بحرانی برای پیش‌بینی وقوع سیلاب ترسیم شد و در اعتبارسنجی با داده‌های واقعی، آستانه بحرانی سیلاب با سطح احتمال ۹۰ درصد روشی قابل درک برای پیش‌بینی سیلاب شناخته شد. بارش‌های بالای آستانه منجر به وقوع سیلاب شهری گشته در حالیکه مقادیر پایین نمودار، موارد بی خطر محسوب می‌شوند. این رویکرد، ابزار مؤثری در سیستم‌های هشدار سریع است و می‌تواند در مدیریت بحران شهری، از خسارات مالی و جانی جلوگیری کند.

کلیدواژه‌ها: بارندگی، گزارش‌های شهروندان، شهر قم، سیلاب شهری و یادگیری ماشین.

Keywords: Rainfall, Citizen reports, Qom, Urban flooding and Machine learning.

¹ niksookhan@ut.ac.ir

۱. مقدمه

سیلاب شهری زمانی رخ می دهد که بارش های حدی منجر به ناتوانی شبکه زهکشی شده و اضافه ظرفیت آب بصورت جریان های سطحی یا آب گرفتگی در خیابان ها بروز پیدا می کند (Beecham & Chowdhury, 2012). این پدیده ناشی از تغییرات اقلیمی و اقدامات گوناگون انسانی - شامل تغییر کاربری زمین، افزایش جمعیت، رشد سریع و برنامه ریزی نشده شهرنشینی، گسترش سطوح نفوذ ناپذیر و زیرساخت های ناکافی - از رایج ترین مخاطراتی است که امروزه با آن روبرو هستیم (Liu et al., 2025). علاوه براین، پیشبینی های بسیاری از برنامه های شبیه سازی اقلیمی در آینده، حاکی از افزایش بارش های حدی در اکثر نقاط جهان به ویژه در فصل بهار تا پایان سده ۲۱ می باشد (O'Shea et al., 2024). بنابراین، یافتن راهکاری برای پیشبینی، پیشگیری و کاهش خسارات ناشی از این پدیده، تبدیل به یک چالش در سطح کلان در زمینه مدیریت آب شهری گشته است (Hill et al., 2023).

تاکنون، تحقیقات متعددی با رویکرد های مختلف به پیشبینی سیلاب و آب گرفتگی در مناطق شهری پرداخته اند. برخی از این روش ها بر استفاده از مدل های هیدرولیکی و یا ترکیب مدل های هیدرولیکی و هیدرولوژیکی تکیه دارند. این فرایند دربرگیرنده شبیه سازی یک رویداد واقعی جهت اعتبارسنجی نتایج محاسباتی مدل های عددی می باشد. سپس، در مرحله بعدی این مدل ها تحت سناریوهای گوناگون، بارها اجرا می شوند تا نتایج پیشبینی حاصل از هر سناریو مشخص گردد (Hassani et al., 2024). در بحث شبیه سازی و پیشبینی سیلاب شهری، انتخاب مدل مناسب و تعیین پارامترهای آن نقش کلیدی و مهمی در صحت نتایج خروجی دارند. داده های در دسترس، اهداف مطالعه و ویژگی های منحصر به فرد هر منطقه یا نمونه موردی از عواملی هستند که بر انتخاب مدل بهینه تاثیر بسزایی دارند.

روش های یادگیری ماشین، پیشبینی های با دقت بالا، بدون نیاز به کشف و درک روابط پیچیده میان پارامترهای موثر در رخداد یک پدیده ارایه می دهند. ابزار یادگیری ماشین، توانایی شناسایی احتمال رخداد سیلاب شهری در آینده، تنها با تکیه بر داده های تاریخی و بدون نیاز به فهم و کشف روابط و فرایند های پیچیده اداره کننده این پدیده را دارد (Tang et al., 2022). افزایش پذیرش عمومی و پیشرفت های محاسبات نرم از جمله روش های یادگیری ماشین، مانند ماشین بردار پشتیبان^۲ (SVMs) (Yan et al., 2018)، جنگل تصادفی^۳ (Random Forests) (Feng et al., 2015; Lee et al., 2017)، شبکه های عصبی گروهی^۴ (ENNs) (Berkhahn et al., 2019)، رگرسیون لجستیک^۵ (LR) (Kim et al., 2020; Lee et al., 2017) و تقویت گرادیان شدید^۶ (XGBoost) (Yuan et al., 2024) در سال های اخیر چشمگیر بوده است.

² Support Vector Machines (SVMs)

³ Random Forests (RFs)

⁴ Ensemble Neural Networks (ENNs)

⁵ Logistic Regression (LR)

⁶ eXtreme Gradient Boosting (XGBoost)

مطالعات متعددی برای دستیابی به درک جامع و کاملی از پیشبینی پدیده های فیزیکی بوسیله جفت کردن داده های تاریخی گزارش های شهروندان یا اطلاعات ثبت شده توسط پایشگرها (مانند سامانه های اجتماعی، تویتر، پهپادهای شناساگر، موبایل و ...) با سایر عوامل موثر بر رویداد یک واقعه با بهره گیری از یادگیری ماشین انجام شده است (Safaei-Moghadam et al., 2024).

در این پژوهش، احتمال کارایی جفت کردن گزارش های شهروندان با داده های تاریخی بارش بوسیله یادگیری ماشین برای دسته بندی⁷ و پیشبینی رخداد سیلاب شهری به عنوان یک روش نوآورانه در جهت ایجاد یک سیستم هشدار مورد بررسی قرار می گیرد. در ادامه، پس از دسته بندی رخداد های سیلاب شهری، یک نمودار آستانه سیلاب، ارایه می گردد که مقادیر بارش های بالاتر از آن منجر به وقوع آبگرفتگی در سطح شهر گشته و مقادیر زیر آن به عنوان سطح امن شناخته می شوند. گسترش یک رویکرد علمی که در نبود اطلاعات گسترده و پیچیده مربوط به شناخت یک پدیده، منجر به استخراج یک روش ساده و قابل فهم (نمودار آستانه سیلاب) برای مدیران شهری جهت تجهیز و آمادگی برای وقوع سیلاب شهری است، نوآوری دیگر تحقیق حاضر می باشد.

۲. مواد و روش ها

مطابق شکل ۱، مرحله نخست شامل ایجاد شبکه ایستگاه اندازه گیری داده های باران، برداشت داده ها با استفاده از ماهواره GSMaP⁸، استخراج و دسته بندی گزارش های مردمی است و آموزش سه مدل یادگیری ماشین شامل SVC، LR و RF در مرحله بعدی قرار می گیرد. در پایان با توجه به موثرترین ایستگاه ها، اقدام به تهیه نمودار آستانه سیلاب می شود.

⁷ Classification

⁸ Global Surface Water Mapper (GSMaP)

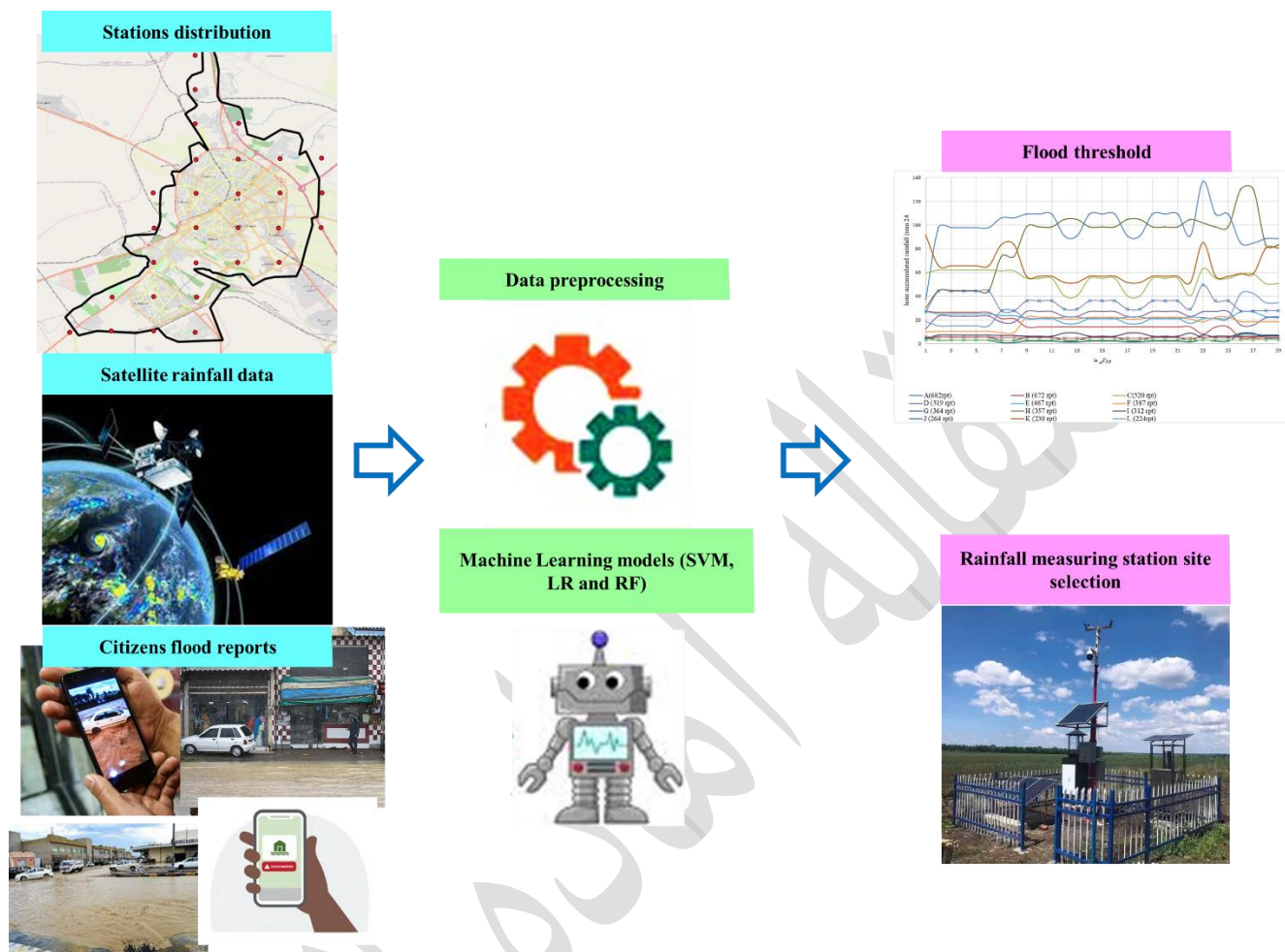


Figure 1. The main steps of the study.

۱.۲ معرفی منطقه مورد مطالعه

مطالعه حاضر بر شهر قم که در عرض و طول جغرافیایی از ۳۴.۵۷ تا ۳۴.۷۸ درجه شمالی و ۵۰.۷۸ تا ۵۰.۹۸ درجه شرقی در بخش مرکزی ایران قرار دارد، متمرکز است (شکل ۲). شهر قم مساحتی برابر ۱۱،۶۱۳ کیلومتر مربع را در بر می گیرد. رقوم ارتفاعی استان قم از ۸۰۰ تا ۳،۳۰۰ متر بالاتر از سطح دریا می باشد و این تغییرات زیاد باعث گوناگونی شرایط آب هوایی از نیمه کوهستانی تا بیابانی گردیده است. اما بیش از ۷۵ درصد مساحت استان شرایط نیمه بیابانی دارد (Panahi et al., 2022). جمعیت شهر بالغ بر ۱/۳۸۰/۰۰۰ نفر در سال ۱۴۰۰ گزارش شده است. آمارها حاکی از رشد بیش از دو برابری جمعیت شهر در بازه زمانی ۳۰ سال (از ۱۳۵۸ تا ۱۳۹۰) می باشد. اگرچه میزان بارش سالیانه در این شهر ۱۵۰ میلیمتر است، در فروردین ۱۳۹۸ (آوریل ۲۰۱۹) تنها در دو ایستگاه بارانسنجی مقادیر ۱۳۷ و ۱۳۹ میلیمتر بارش ثبت شده که اندکی کمتر از کل بارش سالیانه می باشد.

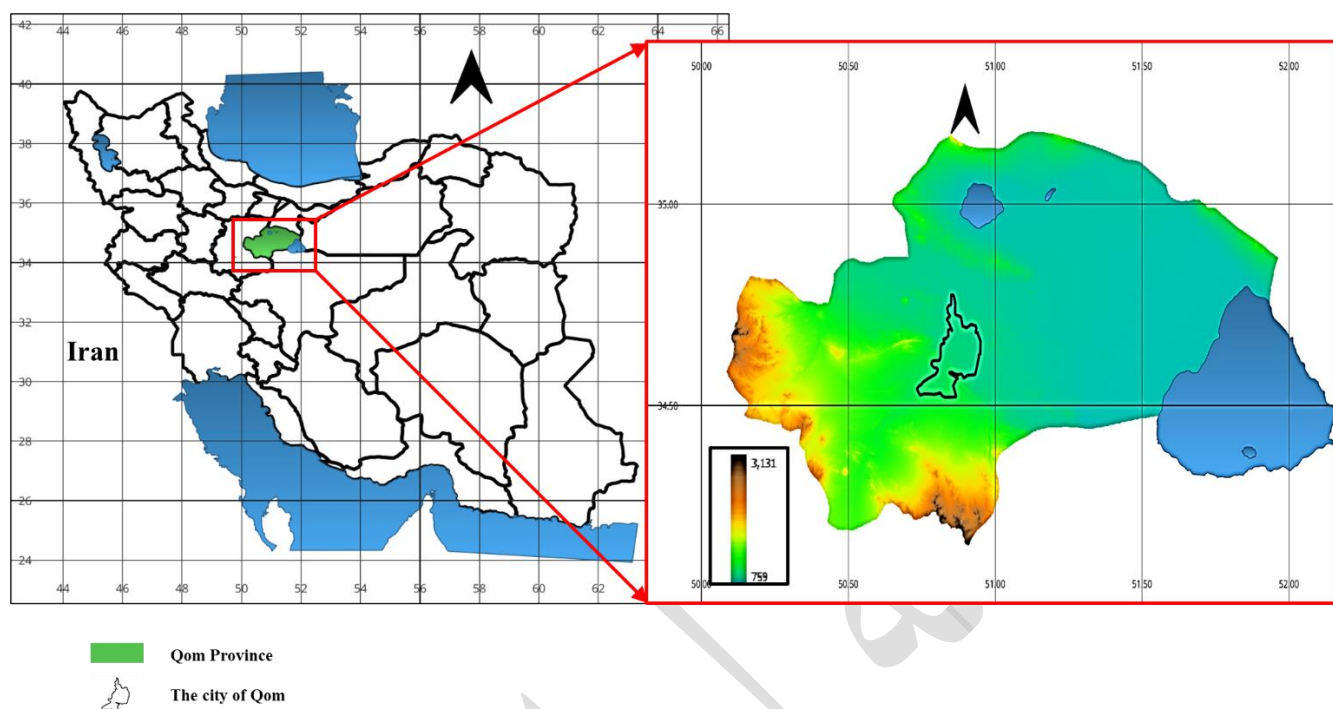


Figure 2. The situation of the city of Qom (study area) in Qom province and central part of Iran.

۲.۲ تعیین شبکه برداشت داده های بارش

مطابق شکل ۳، محدوده شهر قم بوسیله شبکه مربعی با فواصل ۳۰۰۰ متر تقسیم بندی شده و ۲۹ نقطه تقاطع خطوط شبکه به عنوان ایستگاه های برداشت داده های بارندگی انتخاب شدند. انتخاب فواصل کمتر یا بیشتر منجر به بارش های یکسان در ایستگاه ها یا کاهش تعداد ویژگی های^۹ یادگیری ماشین می گردد.

^۹ Features

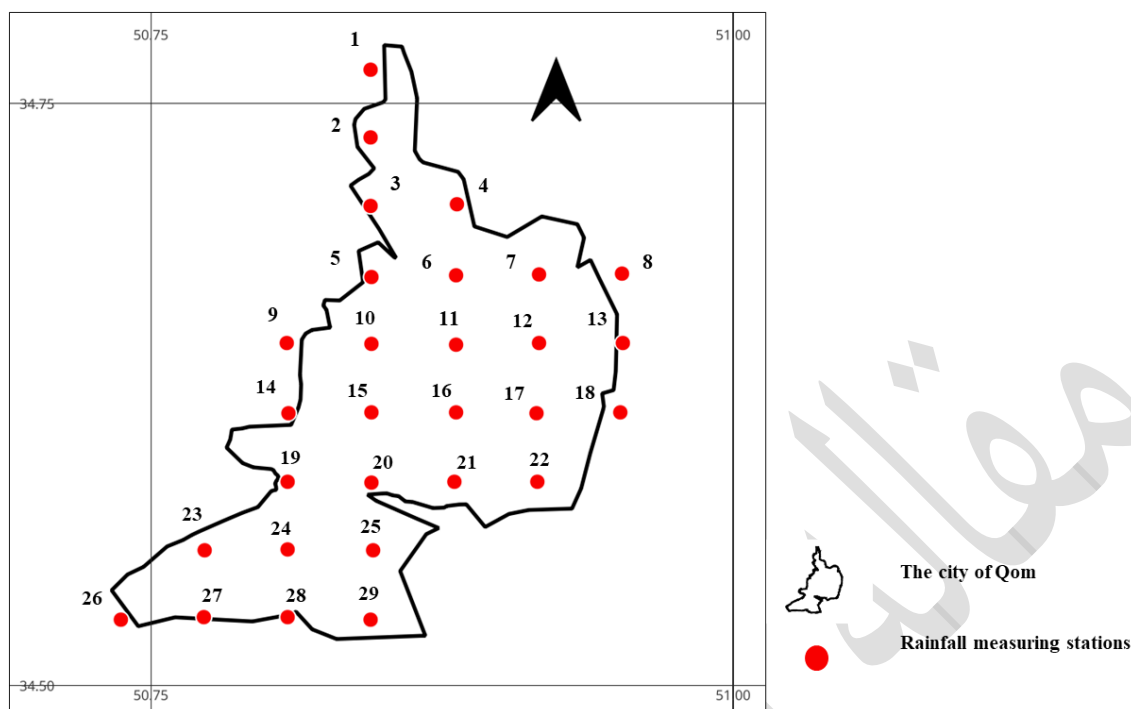


Figure 3. Rainfall measuring stations.

۲.۳ برداشت داده های بارش

مقادیر بارش ۲۴ ساعته در ۲۹ ایستگاه مشخص شده در شکل ۳ از ماهواره GSMMap در بازه زمانی ۱۶ مرداد ۱۳۹۷ (هفتم آگوست ۲۰۱۸) تا ۱۶ مهر ۱۴۰۳ (هفتم اکتبر ۲۰۲۴) برداشت شد. دقت بالای مکانی و زمانی داده های بارندگی این ماهواره، در مطالعات متعددی بویژه در زمینه شبیه سازی مناطق متراکم شهری مورد ارزیابی و مقایسه با اندازه گیری های میدانی قرار گرفته و نتایج قابل قبولی ارائه داده است (Priyambodoho et al., 2021).

۲.۴ مشاهدات و گزارش های شهروندان

در سال ۹۷ (۲۰۱۸) شهرداری قم یک سرویس ارتباط مردمی به نام سامانه ۱۳۷ راه اندازی کرد. بوسیله این سامانه، شهروندان می توانند مشکلات خود را در زمینه های تجهیزات و امکانات شهری مانند خرابی راه ها، روناب سطحی، جمع آوری پسماند، شبکه زهکش شهری و ... با مسئولان محلی در میان گذاشته و فرایند رفع مشکل را پیگیری نمایند. این گزارش ها بوسیله تلفن جمع آوری شده و شامل اطلاعاتی مانند شماره تلفن شهروند، تاریخ و محل حادثه، یک توضیح کوتاه درباره مشکل و اقدامات صورت گرفته توسط واحد مربوطه می باشد. از میان تمام موضوعات گزارش شده، مطالب مربوط به مشاهده سیلاب شهری جدا گردیده و سایر موارد کنار گذاشته شد. برای این منظور، ۲۰۲۵۴ روز از تاریخ ۱۶ مرداد ۱۳۹۷ تا ۱۶ مهر ۱۴۰۳ تمامی گزارش های ثبت شده در سامانه ۱۳۷ مورد بررسی قرار گرفت.

۱۳۷۰ روز، به عنوان روزهای غیر بارانی خشک با توجه به عدم بارندگی در هیچ یک از ۲۹ ایستگاه مشخص شده شناخته شد. در روزهای خشک، تعداد متوسط گزارش های مرتبط با آبگرفتگی در شهر که توسط شهروندان اعلام شده بود برابر ۵.۶ بدست آمد. این عدد، معیار تمایز روزهای عادی یا شرایط معمولی شهر و روزهای سیلابی در نظر گرفته شد. از میان ۸۸۴ روز بارانی تر باقیمانده، آنهایی که تعداد گزارش آبگرفتگی شهری بیشتر یا مساوی ۶ دارند، به عنوان روزهای مشاهده سیلاب شهری شناخته شدند. به این ترتیب، برای دسته بندی

یادگیری ماشین، روزهای تر دارای مقادیر بیشتر یا مساوی ۶ گزارش مشاهده سیلاب شهری، در دسته ۱ و روزهای تر دارای کمتر از ۶ گزارش، در دسته ۰ قرار گرفتند.

از آنجا که بیشتر بارندگی ها (بیش از ۹۷ درصد) در روزهای تر در بازه زمانی مهر تا خرداد رخ داده، آمار این بازه زمانی به عنوان داده های اصلی مطالعه لحاظ گشته و اطلاعات ماه های تیر، مرداد و شهریور حذف شد. با این کار، ۲۲۱ عدد از دسته ۰ (غیر سیلاب) کم شده و ناترازی^{۱۰} تعداد دسته ۰ و ۱ کاهش پیدا کرد. در نهایت، از ۸۸۴ داده در روزهای تر، ۱۶۷ عدد در دسته ۱ و ۴۸۴ رخداد در دسته ۰ قرار گرفتند در حالیکه در حالت نخست، شمارگان دسته ۱ به ۰ برابر ۱۷۹ به ۷۰۵ بود.

۲.۵ یادگیری ماشین

به منظور تحقق اهداف پژوهش، ۳ مدل یادگیری ماشین شامل ماشین بردار پشتیبان، رگرسیون لجستیکی و جنگل تصادفی آموزش داده شد. تمامی رویدادها به دو دسته کلاس ۱ و ۰ تقسیم شدند که به ترتیب نماینده وقوع سیلاب شهری و عدم وقوع آن است. برای ارزیابی و مقایسه کارایی مدل ها، از ماتریس درهم ریختگی^{۱۱}، دقت^{۱۲}، حساسیت^{۱۳} و امتیاز F1^{۱۴} استفاده گردید. برای پیشگیری از رخداد بیش‌برازش^{۱۵} از روش اعتبارسنجی متقاطع k-fold^{۱۶} بهره گرفته شد.

- ماشین بردار پشتیبان، الگوریتم یادگیری توانایی است که اغلب در دسته بندی داده ها به روش دودویی ۱ و ۰ (مثلاً سیلاب یا غیر سیلاب) به کار می رود. این الگوریتم با توجه به هسته^{۱۷} مورد استفاده دارای انواع خطی^{۱۸}، چند جمله‌ای^{۱۹}، سیگموئید^{۲۰} و تابع پایه شعاعی^{۲۱} برای افزایش قدرت دسته بندی داده های پیچیده می باشد (Rifath et al., 2024).

- رگرسیون لجستیکی، یک روش آماری است که برای بررسی رابطه بین متغیرهای پاسخ دسته‌ای، چه اسمی یا ترتیبی، و یک یا چند متغیر توضیحی کمی یا دسته‌ای استفاده می‌شود. این تحلیل به عنوان یک روش پایه برای دسته بندی عمل می‌کند و نشان می‌دهد چگونه یک متغیر پاسخ دودویی (Y)، که ممکن است نشان‌دهنده نتایج مثل موفقیت یعنی مقدار Y برابر ۱ و شکست به ازای Y برابر ۰، با متغیرهای پیشینی X باشد. نتیجه دسته‌ای است و معمولاً دو مقدار دارد، مانند ۰/۱ که نشان‌دهنده رویداد سیل یا عدم وقوع آن است (Tella et al., 2023).

- جنگل تصادفی، یک الگوریتم یادگیری ماشین گروهی نظارتی است که چندین درخت تصمیم غیرهمبسته را ترکیب می‌کند. این درخت‌ها با ساختن از مجموعه‌های آموزش مختلف که به صورت تصادفی از روی مجموعه داده اصلی نمونه‌گیری شده‌اند، غیرهمبسته می‌شوند. این فرآیند از تکنیک بسته‌بندی^{۲۲} استفاده می‌کند تا تنوع بین درخت‌ها و دقت مدل افزایش یابد (Golabi et al., 2020).

¹⁰ Imbalance

¹¹ Confusion Matrix

¹² Precision

¹³ Recall

¹⁴ F1-score

¹⁵ Overfitting

¹⁶ K-fold cross validation

¹⁷ Kernel

¹⁸ Linear

¹⁹ Polynomial

²⁰ Sigmoid

²¹ Radial Basis Function (RBF)

²² Bagging

- به دلیل محدودیت‌های تعدادی در داده‌های تاریخی این مطالعه، از روش اعتبارسنجی متقاطع k-fold استفاده شده است، زیرا این روش عملکرد قوی‌تر و پایدارتری نسبت به تقسیم داده‌ها به ۳۰٪ برای آزمایش و ۷۰٪ برای آموزش فراهم می‌کند (Tian et al., 2019). در روش فوق که برای ارزیابی مدل و پیشگیری از بیش‌برازش می‌باشد (Montesinos López et al., 2022)؛ پس از ارزیابی مقادیر مختلف foldها، دو مقدار ۵ و ۷ بسته به مدل یادگیری به کار گرفته شد. بیش‌برازش که به موفقیت صددرصدی کاذب ماشین در دسته بندی داده‌ها اشاره دارد؛ در حقیقت شرايطی است که در آن ماشین چیزی را یاد نگرفته بلکه با دسته بندی همه داده‌ها (هر داده در یک دسته) و تنها از بر کردن آن‌ها به موفقیت نادرست در پیشبینی رسیده است. علاوه بر این، برای تنظیم^{۲۳} پارامترهای مدل‌ها از روش جستجوی شبکه‌ای^{۲۴} استفاده گردید. این روش در یک فرایند آزمون و خطا ترکیبات ممکن از پارامترهای تعریف شده مدل را تولید می‌نماید. سپس، هر ترکیب را آزمایش کرده؛ مجموعه‌ای از بهترین پارامترها که دارای عملکرد بهتری با معیار مشخص (در اینجا صحت^{۲۵}) دارند را مشخص می‌کند.
- ماتریس درهم ریختگی، یک جدول دو بعدی است که برای ارزیابی عملکرد دسته بندی‌های چندکلاسه استفاده می‌شود. این جدول شامل چهار مؤلفه اصلی است: منفی‌های درست (TN)^{۲۶}، مثبت‌های درست (TP)^{۲۷}، منفی‌های نادرست (FN)^{۲۸}، و مثبت‌های نادرست (FP)^{۲۹}. در این زمینه، FP نشان‌دهنده تعداد رویدادهایی است که به نادرستی به عنوان سیل شناخته شده‌اند، در حالی که FN نشان‌دهنده تعداد رخدادهایی است که بصورت نادرست به عنوان غیرسیل دسته بندی شده‌اند. برعکس، TN و TP به ترتیب تعداد رویدادهایی است که به درستی به عنوان رویداد غیرسیل و سیل شناسایی شده‌اند (Heydarian et al., 2022).
- دقت، یکی از معیارهای استاندارد سنجش در یادگیری ماشین است که نشان می‌دهد چند عدد از رویدادهای سیلی که مدل پیش‌بینی کرده است، با سیل‌های واقعی مطابقت دارند (Ghosh et al., 2022). دقت بالا نشان دهنده کم بودن نرخ مثبت‌های نادرست مدل است. این معیار را می‌توان به صورت زیر محاسبه کرد:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (۱)$$

- حساسیت، نسبت رویدادهای واقعی سیل که به درستی توسط مدل پیش‌بینی شده‌اند، است (Tanim et al., 2022). حساسیت بالا نشان‌دهنده این است که نرخ منفی‌های نادرست مدل کم است. این مقدار را می‌توان با استفاده از رابطه زیر محاسبه کرد:

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (۲)$$

- امتیاز F1، میانگین هارمونیک دقت و حساسیت است (Ghosh et al., 2022). امتیاز F1 بالا نشان‌دهنده تعادل مناسب بین دقت و حساسیت است. فرمول امتیاز F1 در معادله ۳ نشان داده شده است.

$$\text{F1-score} = \text{TP} / (\text{TP} + (\text{FP} + \text{FN}) / 2) \quad (۳)$$

²³ Tuning

²⁴ GridsearchCV

²⁵ Accuracy

²⁶ True Negative

²⁷ True Positive

²⁸ False Negative

²⁹ False Positive

۳. نتایج

۱.۳ نتیجه عملکرد مدل‌های یادگیری ماشین

نتایج و معیارهای ارزیابی دقت، حساسیت و امتیاز F1 آموزش سه مدل یادگیری ماشین شامل SVC، LR و RF در جدول ۱ نمایش داده شده است. مدل‌های SVM با هسته‌های مختلف — خطی، چندجمله‌ای، سیگموئید و RBF — تعداد رویدادهای درست سیلاب (TP) را به ترتیب ۳۵، ۲۲، ۴۶ و ۳۳ از مجموع ۱۶۷ رویداد واقعی سیلاب پیش‌بینی کردند. قابلیت این مدل‌ها در تشخیص درست رویدادهای غیرسیلاب (TN) بسیار بالا بود و در بازه ۴۵۳ تا ۴۸۲ از ۴۸۴ مورد واقعی غیرسیلاب قرار داشتند. دقت مدل‌ها در پیش‌بینی سیلاب و غیرسیلاب به ترتیب در بازه‌های ۶۰ تا ۹۲ درصد و ۷۷ تا ۷۹ درصد قرار دارد. امتیاز F1 برای شناسایی سیلاب کمتر از ۴۰ درصد است، در حالی که برای تشخیص غیرسیلاب حدود ۸۷ درصد است.

مطابق نتایج مدل LR، تعداد نمونه‌های درست مثبت (TP) برابر با ۴۳ و نمونه‌های نادرست منفی (FN) برابر با ۱۲۴ است، همچنین تعداد نمونه‌های منفی درست (TN) برابر با ۴۷۰ و نمونه‌های نادرست مثبت (FP) برابر با ۱۴ است. این اعداد منجر به دقت ۷۵ درصد و حساسیت ۲۶ درصد برای موارد سیلاب شده است. در مقابل، برای رویدادهای غیرسیلاب، دقت ۷۹ درصد و حساسیت ۹۷ درصد محاسبه شده است. مقایسه امتیاز F1 مدل LR با مدل‌های SVM تفاوت معنی‌داری بین آن‌ها نشان نمی‌دهد.

در میان سه مدل یادگیری ماشین، مدل RF بیشترین تعداد شناسایی مثبت درست (۸۳)، منفی درست (۴۸۲) و شاخص‌های عملکرد را به دست آورد، با دقت ۹۸ درصد، حساسیت ۵۰ درصد و امتیاز F1 برابر با ۶۶ درصد برای موارد سیلاب.

Table 1. The results of 3 ML models including confusion matrix, precision, recall and F1-score.

		Actual Flood (1)	Actual Non-Flood (0)	Precision	Recall	F1-score
SVM (Linear)	Predicted Flood (1)	TP= 35 35 flood events are truly identified as floods	FP= 9 9 non-flood events are incorrectly predicted as floods	0.8	0.21	0.33
	Predicted Non-Flood (0)	FN= 132 132 flood events are incorrectly identified as non-floods	TN= 475 475 of non-flood events are truly identified as non-floods	0.78	0.98	0.87
SVM (Poly-Nominal)	Predicted Flood (1)	TP= 22 22 flood events are truly identified as floods	FP= 2 2 non-flood events are incorrectly predicted as floods	0.92	0.13	0.23
	Predicted Non-Flood (0)	FN= 145 145 flood events are incorrectly identified as non-floods	TN= 482 482 of non-flood events are truly identified as non-floods	0.77	1	0.87

SVM (Sigmoid)	Predicted Flood (1)	TP= 46 46 flood events are truly identified as floods	FP= 31 31 non-flood events are incorrectly predicted as floods	0.6	0.28	0.38
	Predicted Non-Flood (0)	FN= 121 121 flood events are incorrectly identified as non-floods	TN= 453 453 of non-flood events are truly identified as non-floods	0.79	0.94	0.86
SVM (RBF)	Predicted Flood (1)	TP=33 33 flood events are truly identified as floods	FP= 13 13 non-flood events were incorrectly predicted as floods	0.72	0.2	0.31
	Predicted Non-Flood (0)	FN= 134 134 flood events were incorrectly identified as non-floods	TN= 471 471 of non-flood events are truly identified as non-floods	0.78	0.97	0.87
LR	Predicted Flood (1)	TP= 43 43 flood events are truly identified as floods	FP= 14 14 non-flood events are incorrectly predicted as floods	0.75	0.26	0.38
	Predicted Non-Flood (0)	FN= 124 124 flood events are incorrectly identified as non-floods	TN= 470 470 of non-flood events are truly identified as non-floods	0.79	0.97	0.87
RF	Predicted Flood (1)	TP=83 83 flood events are truly identified as floods	FP= 2 2 non-flood events are incorrectly predicted as flood	0.98	0.5	0.66
	Predicted Non-Flood (0)	FN= 84 84 flood events are incorrectly identified as non-floods	TN= 482 482 of non-flood events are truly identified as non-floods	0.85	1	0.92

برای اطمینان از صحت کارایی مدل‌ها و عدم وقوع بیش‌برازش از روش اعتبارسنجی متقاطع k-fold (تعداد ۵ و ۷) به همراه تنظیم نمودن ابرپارامترهای هر مدل با روش جستجوی شبکه‌ای استفاده شده است. برای بررسی این موضوع از روش اجرای هر مدل بصورت جداگانه روی داده‌های آموزشی و آزمایشی استفاده گردید (Bishop & Nasrabadi, 2006; Kohavi, 1995) که نتایج آن برای مدل برگزیده جنگل تصادفی در نمودار زیر نمایش داده شده است. اختلاف ناچیز (حدود ۰.۰۵) نتایج معیارهای ارزیابی در آموزش و آزمون حاکی از تعمیم‌پذیری و پایداری قابل قبول مدل می‌باشد (Dietterich, 1998; Kohavi, 1995).

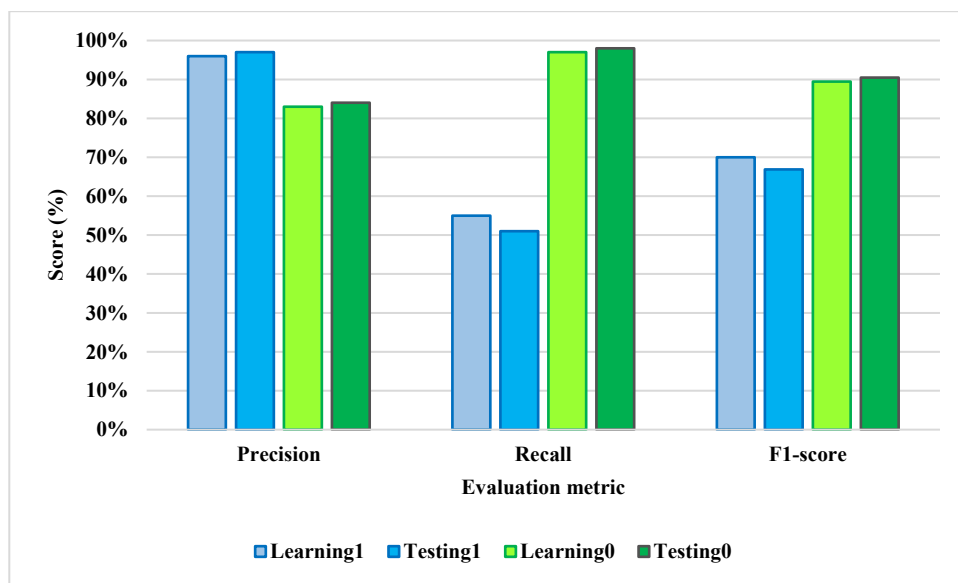


Figure 6. Comparison of RF model performance metrics on training and testing datasets (0 = Non-Flood, 1 = Flood).

۲.۳ مهمترین ویژگی ها

با وجود تفاوت در عملکرد مدل ها، تمام مدل ها به جز مدل SVM با هسته سیگموئید، به طور متفق ایستگاه های ۱ و ۲۳ را به عنوان مهم ترین ویژگی ها در یادگیری ماشین برای پیش بینی سیلاب شهری تشخیص دادند. نتایج مربوط به ضریب اهمیت هریک از ایستگاه ها (ویژگی ها) در جدول زیر نمایش داده شده است و بیشترین ضریب اهمیت مستخرج از هر مدل به رنگ سبز مشخص می باشد.

Table 3. The results of feature importance for three ML models.

Feature	SVC Linear	SVC Polynomial	SVC Sigmoid	SVC RBF	LR	RF
	Importance					
1	0.6908	0.0040	-0.0007	0.0040	1.2305	0.0676
2	0.1721	0.0001	0.0000	0.0034	0.2982	0.0225
3	0.1721	0.0001	0.0000	0.0010	0.2982	0.0215
4	0.1721	0.0001	0.0000	0.0010	0.2982	0.0240
5	0.1721	0.0001	0.0000	0.0035	0.2982	0.0224
6	0.1721	0.0001	0.0000	0.0034	0.2982	0.0227
7	0.1426	0.0003	0.0006	0.0000	0.1233	0.0397
8	0.1426	0.0003	0.0006	0.0000	0.1233	0.0366
9	0.0175	0.0015	0.0023	0.0035	0.0076	0.0304
10	0.0175	0.0015	0.0023	0.0035	0.0076	0.0360
11	0.0175	0.0015	0.0023	0.0035	0.0076	0.0312
12	0.0480	0.0014	-0.0007	0.0035	0.0812	0.0361

13	0.0480	0.0014	-0.0007	0.0035	0.0812	0.0326
14	0.0175	0.0015	0.0023	0.0034	0.0076	0.0265
15	0.0175	0.0015	0.0023	0.0035	0.0076	0.0345
16	0.0175	0.0015	0.0023	0.0034	0.0076	0.0273
17	0.0480	0.0014	-0.0007	0.0000	0.0812	0.0412
18	0.0480	0.0014	-0.0007	0.0000	0.0812	0.0299
19	0.0175	0.0015	0.0023	0.0035	0.0076	0.0267
20	0.0175	0.0015	0.0023	0.0035	0.0076	0.0318
21	0.0175	0.0015	0.0023	0.0025	0.0076	0.0306
22	0.0480	0.0014	-0.0007	0.0035	0.0812	0.0309
23	0.6502	0.0016	0.0015	0.0034	0.9553	0.0741
24	0.0175	0.0015	0.0023	0.0000	0.0076	0.0327
25	0.0175	0.0015	0.0023	0.0035	0.0076	0.0368
26	0.1602	0.0002	0.0025	0.0035	0.2705	0.0420
27	0.1602	0.0002	0.0025	0.0035	0.2705	0.0422
28	0.0186	0.0001	0.0020	0.0034	0.3411	0.0330
29	0.0186	0.0001	0.0020	0.0034	0.3411	0.0364

۳.۳ نمودار آستانه سیلاب

با توجه به اینکه مدل RF بیشترین امتیاز را در زمینه‌های دقت و امتیاز F1 در بین تمام مدل‌های مقایسه شده در پیش‌بینی رویدادهای سیلاب و غیرسیلاب کسب کرده است، منحنی آستانه‌ی بارش بحرانی از روی این مدل استخراج شده و در شکل ۵ نشان داده شده است. با در نظر گرفتن احتمالات ۹۵ و ۹۰ درصد، مقدار بارش‌های بالای آستانه، احتمالاً منجر به سیلاب‌های شهری می‌شود، خصوصاً در ایستگاه‌های ۱ و ۲۳ که به عنوان مهم‌ترین ویژگی‌ها شناسایی شده‌اند. در مقابل، نقاط زیر منحنی احتمال کمتری برای وقوع سیل دارند. آستانه‌های بحرانی بارش برای مهم‌ترین ویژگی‌ها (ایستگاه‌های ۱ و ۲۳) در حدود ۵/۸۷ میلی‌متر و ۵/۴۲ میلی‌متر در سطح احتمالی ۹۵ درصد و در حدود ۳.۲۲ میلی‌متر و ۳.۱۷ میلی‌متر در سطح احتمالی ۹۰ درصد اطمینان از درستی تشخیص سیلاب شهری است.

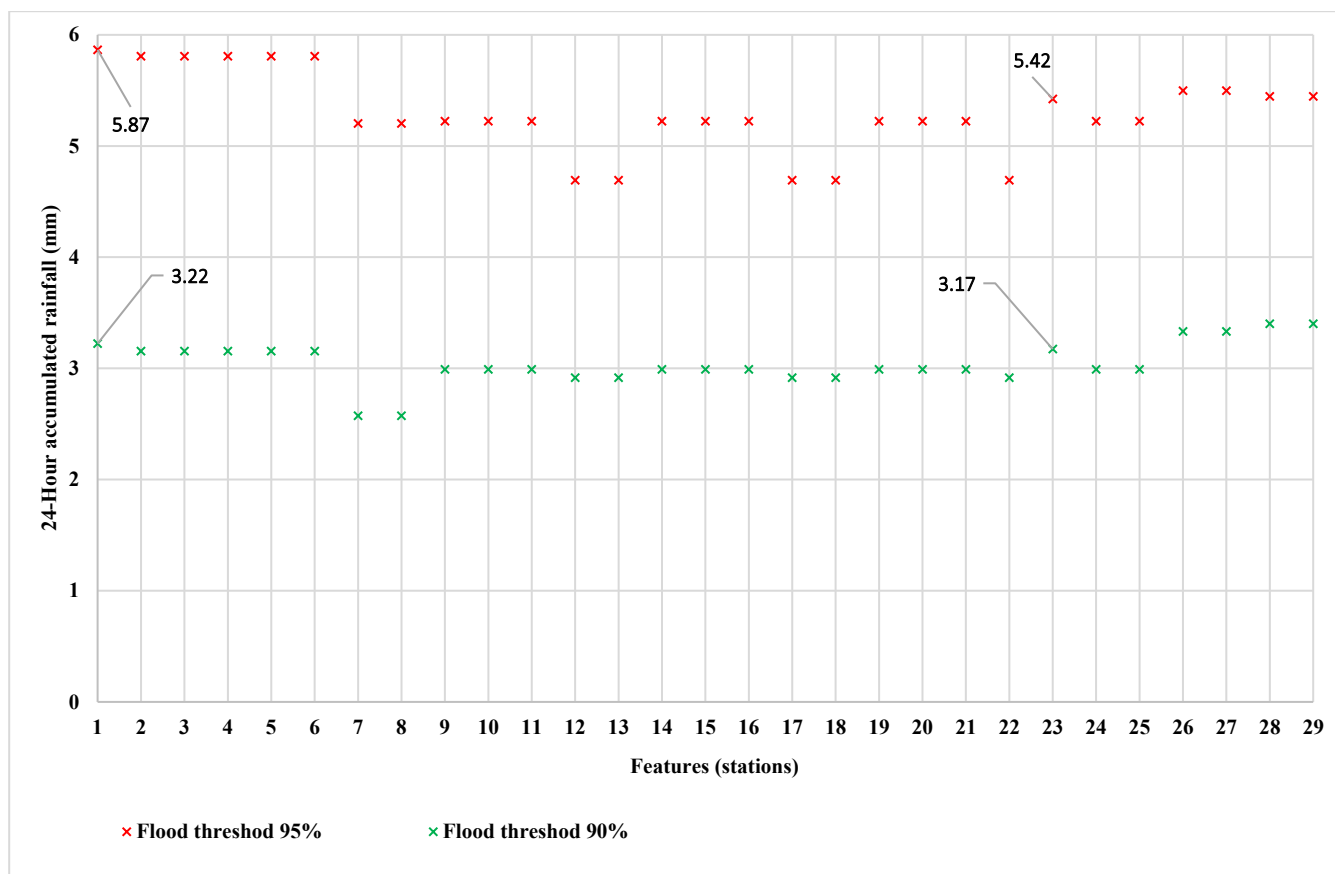


Figure 5. Flood threshold diagrams at probability levels of 95 and 90 percent.

۴. بحث

۱.۴ بررسی عملکرد مدل های یادگیری ماشین

همانطور که پیشتر ذکر شد، قم در منطقه‌ای خشک و بیابانی در قسمت مرکزی ایران قرار دارد. بنابراین، سیلاب‌های ناگهانی اغلب منجر به جریان‌های سطحی، آبرگرفتنی خیابانی، اختلال در حمل و نقل، خسارت به اموال و در موارد نادر، خطرات جانی می‌شوند. نسبت هشدارهای نادرست (FP) بالا می‌تواند اعتماد عمومی را تضعیف کرده، تمایل مردم به پاسخگویی به سیستم‌های هشدار دهنده را کاهش دهد (Barnes et al., 2007; Nikolopoulos et al., 2014; Staley et al., 2013). دقیقاً به همین منظور، مفهوم دقت مطرح می‌شود که نشان می‌دهد چند درصد از پیش‌بینی‌های سیلاب، واقعاً درست بوده‌اند. برعکس، زمانی که هدف اصلی شناسایی تمامی رخداد‌های سیلاب است و در شرایطی که حتی عدم پیش‌بینی یک مورد، مخاطرات جدی در پی دارد، معیار حساسیت اهمیت پیدا می‌کند تا میزان توانایی مدل در تشخیص تمامی سیلاب‌ها ارزیابی شود.

در این مطالعه، نقش دقت بسیار مهم است و همراه با امتیاز F1، کارایی مدل‌های یادگیری ماشین را نشان می‌دهد. مدل RF بیشترین نمره‌ها را هم در تشخیص سیلاب و هم در عدم وقوع آن دارد، با دقت ۹۸ درصد، حساسیت ۵۰ درصد و نمره F1 برابر با ۶۶ درصد برای رویدادهای سیلاب و دقت، حساسیت و نمره F1 به ترتیب ۸۵، ۱۰۰ و ۹۲ درصد در موارد عدم سیلاب.

۲.۴ ارزیابی نمودار آستانه سیلاب

اطلاعات بارش باران روزانه مرتبط با دوازده رویداد با بیشترین مجموع گزارش‌ها در پایگاه داده‌ها (بیش از ۲۰۰ گزارش در روز) مورد استفاده قرار گرفت تا کارایی منحنی‌های آستانه بحرانی سیلاب در سطوح احتمال ۱۰۰، ۹۵ و ۹۰ درصد ارزیابی شود. مطابق شکل ۶، در آستانه احتمال ۱۰۰ درصد، تنها چهار رویداد (A، H، C و K) در تمامی ایستگاه‌ها بالای نقاط آستانه سیلاب قرار می‌گیرند که نشان‌دهنده وقوع قطعی سیلاب ناشی از بارش است. آستانه ۹۰ درصد تمام رویدادهای سیلاب را با موفقیت شناسایی می‌کند، بدون اینکه سیلابی را از دست بدهد، به‌ویژه در نقاط ۱ و ۲۳. بنابراین، نصب باران‌سنج در این مکان‌ها می‌تواند به تحلیل الگوهای سیلاب و بهبود سیستم‌های هشدار سریع در شهر کمک کند.

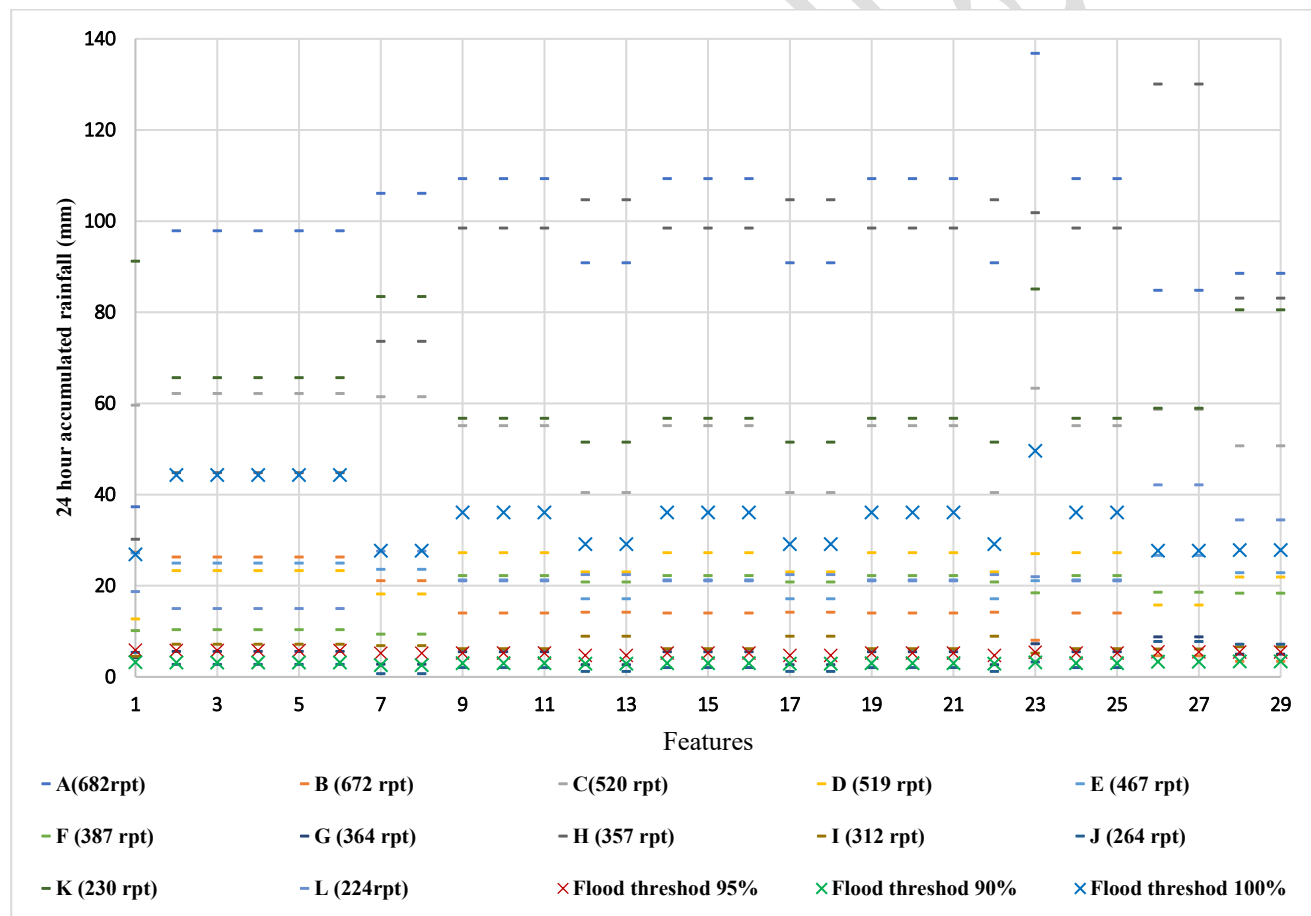


Figure 6. Flood threshold evaluation at probability levels of 100, 95, and 90 percent based on 12 real events (A to L) with more than 200 reports (rpt).

۳.۴ محدودیت‌های پژوهش

با اینکه گزارش‌های شهروندان می‌توانند جایگزین ارزشمندی برای مدل‌های هیدرولوژیکی و هیدرولیکی پیچیده باشند، همچنان ابهاماتی وجود دارد، مانند تأخیر زمانی در دریافت گزارش‌ها. ممکن است بارندگی در شب رخ داده و گزارش‌ها در روز بعد ثبت شوند یا اینکه مشکلات ناشی از بارندگی پس از گذشت زمان ایجاد شده باشند. علاوه بر این، به‌روزرسانی‌ها و ارتقای دوره‌ای شبکه زهکشی شهری آب نیز در نظر گرفته نشده است. تراکم جمعیتی نیز بر حجم گزارش‌ها تأثیر دارند؛ مناطق پر جمعیت معمولاً گزارش‌های بیشتری دارند.

۵. نتیجه گیری و پیشنهادها

گزارش‌های سیلاب ثبت شده توسط شهروندان، داده‌های بارش و مدل‌های یادگیری ماشین، نقش قابل اعتمادی در جایگزینی مدل‌های هیدرولیکی و هیدرولوژیکی در پیش‌بینی سیلاب‌های شهری در قم، ایران را دارند. این نتیجه با یافته‌های تحقیقات دیگر مبنی بر تایید کارایی تلفیق اطلاعات میدانی توده مردم و روش‌های عددی همسو است (Ke et al., 2020; Wasko et al., 2021).

نتایج در توافق با سایر مطالعات (Yang & Cervone, 2019) نشان می‌دهد که مدل جنگل تصادفی بیشترین انطباق در پیش‌بینی سیلاب شهری را ارائه داده است؛ با معیارهایی مانند دقت ۹۸ درصد و امتیاز F1، ۶۶ درصد برای رویدادهای سیلاب و دقت و امتیاز F1 به ترتیب ۸۵ و ۹۲ درصد برای رویدادهای غیرسیلاب. با توجه به آب و هوای بیابانی منطقه مورد مطالعه و درجات غیرمهلک آسیب سیلاب شهری، توجه پژوهش بر معیار سنجش دقت، برای کارایی مدل است تا از شناسایی درست رویدادهای سیلاب اطمینان حاصل شده و تعداد پیش‌بینی‌های نادرست سیلاب کاهیده شود از این دیدگاه در تحقیقات دیگر به مساله پرداخته نشده است.

تهیه روش‌های ساده کاربرپسند به عنوان سیستم هشدار سیلاب بویژه در شهرهای ساحلی و پرباران در مطالعات دیگر با استفاده از اطلاعات آماری و نرم‌افزاری صورت گرفته است (Martina et al., 2006; Tian et al., 2019). در این مطالعه به عنوان نوآوری با تمرکز بر مهم‌ترین ویژگی‌ها و استفاده از یادگیری ماشین، آستانه بارش بحرانی منجر به وقوع سیلاب در سطوح احتمالاتی ۱۰۰، ۹۵ و ۹۰ درصد با مشاهدات میدانی رویدادها ارزیابی شده و مشخص گشت سطح احتمال ۹۰ درصد به عنوان مؤثرترین آستانه سیلاب می‌باشد و بارش‌های متناظر با وقوع سیلاب بالای آن قرار میگیرند. مهم‌تر اینکه، با تمرکز بر روی ایستگاه‌های ۱ و ۲۳ به عنوان مکان‌های مهم‌ترین ویژگی‌ها برای پیش‌بینی سیلاب‌های ناشی از بارش، می‌توان به جانمایی برای راه اندازی ایستگاه‌های بارانسنجی دست یافت. یافته‌ها بر نیاز به مکانیزم‌های بهبود یافته گزارش‌دهی در زمان و مختصات جغرافیایی واقعی مانند برنامه‌های موبایل برای افزایش قابلیت اطمینان مدل‌ها تأکید می‌کنند. این پیشنهاد در برخی مطالعات دیگر نیز تأکید شده است (Hill et al., 2023; Safaei-Moghadam et al., 2024).

واژگان پی‌نوشت

1	Accuracy
2	Bagging
3	Classification
4	Confusion matrix
5	Ensemble Neural Networks (ENNs)
6	eXtreme Gradient Boosting (XGBoost)
7	F1-score
8	False Negative (FN)
9	False Positive (FP)

10	Features
11	Global Surface Water Mapper (GSMap)
12	GridsearchCV
13	Imbalance
14	Kernel
15	K-fold cross validation
16	Linear
17	Logistic Regression (LR)
18	Overfitting
19	Polynomial
20	Precision
21	Radial Basis Function (RBF)
22	Random Forests (RFs)
23	Recall
24	Sigmoid
25	Support Vector Machines (SVMs)
26	True Negative (TN)
27	True Positive (TP)
28	Tuning

- Barnes, L. R., Grunfest, E. C., Hayden, M. H., Schultz, D. M., & Benight, C. (2007). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting*, 22(5), 1140-1147 .
- Beecham, S., & Chowdhury, R. (2012). Effects of changing rainfall patterns on WSUD in Australia. Proceedings of the Institution of Civil Engineers-Water Management ,
- Berkhahn, S., Fuchs, L., & Neuweiler, I. (2019). An ensemble neural network model for real-time prediction of urban floods. *Journal of Hydrology* . ٧٠٤-٧٤٣ , ٢٧٠ ,
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer .
- Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7), 1895-1923 .
- Feng, Q., Liu, J., & Gong, J. (2015). Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China. *Water*, 7(4), 1437-1455 .
- Ghosh, B., Garg, S., & Motagh, M. (2022). Automatic flood detection from sentinel-1 data using deep learning architectures. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 201-208 .
- Golabi, M. R., Radmanesh, F., Akhoond-Ali, A. M., Niksokhan, M. H., & Kisi, O. (2020). Development of an indirect method for modelling the water footprint of electricity using wavelet transform coupled with the random forest model. *Hydrological Sciences Journal*, 65(15), 2521-2534 .
- Hassani, M. R., Niksokhan, M. H., Janbehsarayi, S. F. M., & Nikoo, M. R. (2024). Integrated nonurban-urban flood management using multi-objective optimization of LIDs and detention dams based on game theory approach. *Journal of Cleaner Production*, 142737 .
- Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLMCM: Multi-label confusion matrix. *Ieee Access*, 10, 19083-19095 .
- Hill, B., Liang, Q., Boshier, L., Chen, H., & Nicholson, A. (2023). A systematic review of natural flood management modelling: Approaches, limitations, and potential solutions. *Journal of Flood Risk Management*, 16(3), e12899 .
- Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., Huang, X., Yang, H., & Liu, J. (2020). Urban pluvial flooding prediction by machine learning approaches—a case study of Shenzhen city, China. *Advances in Water Resources*, 145, 103719 .
- Kim, H. I., Han, K. Y., & Lee, J. Y. (2020). Prediction of urban flood extent by LSTM model and logistic regression. *KSCE Journal of Civil and Environmental Engineering Research*, 40(3), 273-283 .
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai* ,
- Lee, S., Kim, J.-C., Jung, H.-S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics, Natural Hazards and Risk*, 8(2), 1185-1203 .
- Liu, Y., Zhang, X., Liu, J., Wang, Y., Jia, H., & Tao, S. (2025). A flood resilience assessment method of green-grey-blue coupled urban drainage system considering backwater effects. *Ecological Indicators*, 170, 113032 .
- Martina, M., Todini, E., & Libralon, A. (2006). A Bayesian decision approach to rainfall thresholds based flood warning. *Hydrology and earth system sciences*, 10(3), 413-426 .
- Montesinos López, O. A., Montesinos López, A & , Crossa, J. (2022). Overfitting, model tuning, and evaluation of prediction performance. In *Multivariate statistical machine learning methods for genomic prediction* (pp. 109-139). Springer .
- Nikolopoulos, E. I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., & Borga, M. (2014). Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence. *Geomorphology*, 221, 286-297 .
- O'Shea, D., Nathan, R., Wasko, C., Ho, M., & Sharma, A. (2024). Evaluation of key flood risk drivers under climate change using a bottom-up approach. *Journal of Hydrology*, 640, 131694 .
- Panahi, D. M., Destouni, G., Kalantari, Z., & Zahabiyou, B. (2022). Distinction of driver contributions to wetland decline and their associated basin hydrology around Iran. *Journal of Hydrology: Regional Studies*, 42, 101126 .
- Priyambodoho, B. A., Kure, S., Yagi, R., & Januriyadi, N. F. (2021). Flood inundation simulations based on GSMaP satellite rainfall data in Jakarta, Indonesia. *Progress in Earth and Planetary Science*, 8(1), 34 .
- Rifath, A. R., Mukhtadir, M. G., Hasan, M., & Islam, M. A. (2024). Flash flood prediction modeling in the hilly regions of Southeastern Bangladesh: A machine learning attempt on present and future climate scenarios. *Environmental Challenges*, 17, 101029 .
- Safaei-Moghadam, A., Hosseinzadeh, A., & Minsker, B. (2024). Predicting real-time roadway pluvial flood risk: A hybrid machine learning approach coupling a graph-based flood spreading model, historical vulnerabilities, and Waze data. *Journal of Hydrology*, 131406 .
- Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., & Laber, J. L. (2013). Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides*, 10, 547-562 .
- Tang, X., Machimura, T., Li, J., Yu, H., & Liu, W. (2022). Evaluating seasonal wildfire susceptibility and wildfire threats to local ecosystems in the largest forested area of China. *Earth's Future*, 10(5), e2021EF002199 .

- Tanim ,A. H., McRae, C. B., Tavakol-Davani, H., & Goharian, E. (2022). Flood detection in urban areas using satellite imagery and machine learning. *Water*, 14(7), 1140 .
- Tella, A., Mustafa, M. R. U., Balogun, A. O., Okolie, C. J., Bello Yamusa, I., & Ibrahim, M .(۲۰۲۲) .Spatial prediction of flood in Kuala Lumpur City of Malaysia using logistic regression. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 363-369 .
- Tian, X., Schleiss, M., Bouwens, C., & van de Giesen, N. (2019). Critical rainfall thresholds for urban pluvial flooding inferred from citizen observations. *Science of the total environment*, 689, 258-268 .
- Wasko, C., Nathan, R., Stein, L., & O'Shea, D. (2021). Evidence of shorter more extreme rainfalls and increased flood variability under climate change. *Journal of Hydrology*, 603, 126994 .
- Yan, J., Jin, J., Chen, F., Yu, G., Yin, H., & Wang, W. (2018). Urban flash flood forecast using support vector machine and numerical simulation. *Journal of Hydroinformatics*, 20(1), 221-231 .
- Yang, L., & Cervone, G. (2019). Analysis of remote sensing imagery for disaster assessment using deep learning: a case study of flooding event. *Soft Computing*, 23(24), 13393-13408 .
- Yuan, H., Wang, M., Zhang, D., Ikram, R. M .A., Su, J., Zhou, S., Wang, Y., Li, J., & Zhang, Q. (2024). Data-driven urban configuration optimization: An XGBoost-based approach for mitigating flood susceptibility and enhancing economic contribution. *Ecological Indicators*, 166, 112247 .

Using citizen reports for urban flood prediction with machine learning approach and rainfall Data (Case Study: Qom, Iran)

This research aims to develop a scientific and practical method for urban flood prediction using citizen reports and satellite rainfall data. The city area is divided into a grid with 3,000-meter intervals, and the intersection points of the grid lines are selected as rainfall measurement stations. Days are classified as wet, dry, and normal based on citizen reports and observed rainfall, and machine learning models including Support Vector Machine, Logistic Regression, and Random Forest are trained on these data. The results show that the Random Forest model has the highest performance, with 98percent precision and an F1 score of 66 percent in correctly identifying flood events. Considering Qom's semi-arid climate and the non-lethal damages caused by floods, emphasis was placed on the precision of models in correctly detecting flood cases to enhance confidence and prevent community anxiety due to false alarms. Stations 1 and 23 are identified as critical locations for rainfall equipment installation based on feature importance analysis. Subsequently, a critical rainfall threshold curve at a 90 percent probability level was plotted for flood prediction, and validated with real data. Rainfall above the threshold generally results in floods, while values below are considered safe. This approach is an effective tool for rapid warning systems and can aid urban disaster management to prevent financial and human losses.

Keywords: Keywords: Citizen reports, Machine learning, Qom, Rainfall, Urban flooding.