ارزیابی توانایی مدل درختی استدلالی در پیش‌بینی احتمال وقوع بارش روزانه

فاطمه میکائیلی[۱]، سعید صمدیان فرد[۲]

۱- دانشجوی کارشناسی ارشد، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

۲- استادیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

تاریخ دریافت: ۹۹/۰۶

تاریخ پذیرش: ۹۹/۰۷

چکیده

با توجه به واقع شدن ایران در اقلیم خشک و نیمه‌خشک و توزیع ناهمگن بارندگی، پیش‌بینی وقوع بارش از اهمیت ویژه‌ای برخوردار است. از این رو، امروزه محققین با استفاده از روش‌های نوین در پی شناخت و پیش‌بینی دقیق آن هستند. بنابراین، هدف از پژوهش حاضر، بررسی توانایی‌های مدل درخت استدلالی (LMT) در پیش‌بینی وقوع بارش روزانه ایستگاه پارس‌آباد با استفاده از داده‌های هواشناسی ۱ تا ۳ روز قبل است. برای این منظور، داده‌های هواشناسی دوره ۲۰۰۴-۲۰۱۹ میلادی جمع‌آوری گردید و سه سناریو ترکیبی از پارامترهای هواشناسی برای واسنجی و صحت‌سنجی روش مورد مطالعه مد نظر قرار گرفت. نتایج نشان داد که دقت پیش‌بینی بهترین سناریو با استفاده از داده‌های ۲ روز قبل حدود ۷۹ درصد بود، اما با استفاده از داده‌های ۱ و ۳ روز قبل، بارش روزانه با دقت ۸۰ درصد پیش‌بینی شد. در نهایت، با بررسی معیارهای ارزیابی، سناریو شماره یک با پارامترهای ورودی حداقل، حداکثر و متوسط رطوبت نسبی (درصد)، دما (درجه سانتی گراد)، مجموع ساعات آفتابی (ساعت) و سرعت باد (متر بر ثانیه) به عنوان دقیق‌ترین سناریو برای پیش‌بینی بارش روزانه تعیین گردید.

واژه‌های کلیدی: بارش روزانه، درخت تصمیم، درصد موارد پیش‌بینی صحیح، مدل درخت استدلالی

مقدمه

امروزه محدودیت منابع آبی در بخش‌های مختلف موجب بروز مشکلات عمده‌ای شده است. از طرفی دیگر، تاثیر عوامل محسوس و نامحسوس در وقوع بارندگی آنچنان بالاست که این روند را از یک نظام قانونمند روشن به سوی یک ساختار پیچیده سوق می‌دهد (ناصری،۱۳۸۲). بنابراین، پیش‌بینی وقوع بارندگی برای اهداف مختلفی از قبیل تخمین سیلاب، خشکسالی، آبخیزداری، گردشگری، کشاورزی و برنامه‌ریزی آبیاری اهمیت فراوانی دارد (Freiwan & Cigizoglu, 2005). با اطلاع از زمان وقوع و مقدار بارش می‌توان برنامه‌ریزی مناسبی برای انجام بهینه فعالیت‌ها در زمینه‌های مختلف انجام داد. امروزه سیستم‌های هوشمند مبتنی بر کاوش داده‌ها در مدل‌سازی فرآیندهای هیدرولوژیکی مورد توجه قرار گرفته‌اند. به عبارت دیگر، این سیستم‌ها به عنوان روش‌های معتبر و شناخته‌شده در مدل‌سازی فرآیندهای پیچیده غیرخطی قادرند با استفاده از داده‌های هواشناسی، تخمین‌های دقیق‌تری ارائه دهند (Bhattacharya & Solomatine, 2005). در سال‌های اخیر، انواع سیستم‌های هوشمند در بخش‌های مختلف مهندسی آب، به ویژه در زمینه پیش‌بینی پارامترهای هواشناسی و هیدرولوژیکی مورد استفاده قرار گرفته‌اند.

[2] نویسنده مسئول، ایمیل: s.samadian@tabrizu.ac.ir

[Downloaded from jircsa.ir on 2026-02-21]

[DOR: 20.1001.1.24235970.1399.8.2.5.2]

Maeda و همکاران (2001) بارش ژاپن را به کمک شبکه‌های عصبی مصنوعی پیش‌بینی کردند و نتایج حاصله را با پیش‌بینی های ساعتی سازمان هواشناسی ژاپن مقایسه کردند. نتایج آنها نشان داد که شبکه‌های عصبی مصنوعی ابزار مناسبی برای پیش‌بینی بارش می‌باشند. Mekanik و همکاران (2013) با بررسی پیش‌بینی بارش فصل بهار در حوضه ویکتوریای استرالیا و با استفاده از روش‌های شبکه عصبی و رگرسیون چندگانه نشان دادند که روش عصبی مصنوعی، عملکرد بهتری در پیش‌بینی بارش دارد. George و همکاران (2016) با استفاده از یک مدل دوبعدی خطی، مقدار بارش روزانه را براساس داده‌های هواشناسی همان روز پیش‌بینی کردند. Dash و همکاران (2018) با بررسی پیش‌بینی بارش در ایالت کرالا در هند با استفاده از روش‌های هوش مصنوعی نتیجه گرفتند که رویکردهای پیشنهادی هوش مصنوعی توانایی پیش‌بینی بارش‌های موسمی تابستانه و پس از آن را با دقت ۸۰ درصد دارا هستند. خلیلی و همکاران (۱۳۸۹) با پیش‌بینی بارش روزانه در مشهد با استفاده از شبکه عصبی مصنوعی نشان دادند که شبکه‌های عصبی مصنوعی قادر به کشف روابط غیرخطی حاکم بر سیستم پیچیده بارش می‌باشند. خسروی و شکیبا (۱۳۸۹) به منظور مدیریت سیل به پیش‌بینی بارش با استفاده از شبکه‌های مصنوعی (ANN) پرداخته و با تحلیل نتایج خروجی مدل شبکه عصبی، نشان دادند که این مدل توانایی بهتر و دقت بالایی برای پیش‌بینی بارش نسبت به روش‌های آماری معمول داشته و با افزایش فاکتورهای ورودی، شبکه دقت بالاتری را در پیش‌بینی ارائه می‌دهد. سلگی و همکاران (۱۳۹۷) به بررسی پیش‌بینی بارش ماهانه در ایستگاه وراینه در شهرستان نهاوند با استفاده از مدل‌های برنامه‌ریزی بیان‌ژن و ماشین بردار پشتیبان پرداختند و نتیجه گرفتند که مدل برنامه‌ریزی بیان‌ژن برای مدل‌سازی و پیش‌بینی بارش ماهانه مناسب می‌باشد. مهتابی و همکاران (۱۳۹۷) با بررسی پیش‌بینی بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل در شهر اصفهان و با استفاده از روش‌های مدل‌های هوشمند بردار پشتیبان، K-نزدیکترین همسایگی، شبکه عصبی مصنوعی و درخت تصمیم نشان دادند که در هر چهار روش دقت پیش‌بینی بهترین سناریوها با استفاده از داده‌های ۶ و ۷ روز قبل کمتر از ۷۵ درصد بوده، اما با استفاده از داده‌های روزهای ۱ تا ۵ روز قبل، بارش روزانه با دقت بیش از ۸۰ درصد پیش‌بینی شد و عملکرد درخت تصمیم بهتر از سه روش دیگر بود.

مدل درخت استدلالی (LMT) یکی از روش‌های طبقه‌بندی نوین بوده که تا حدی که نویسندگان مقاله اطلاع دارند، هیچ گونه تحقیقی تا به حال در مورد استفاده از LMT در پیش‌بینی بارش روزانه گزارش نشده است. ولی در مقالات بین المللی، استفاده از این مدل در پیش بینی کورسپتور HIV (Shoombuatong et al., 2012)، پیش‌بینی مکانی رانش زمین (Chen et al., 2017) و ... گزارش گردیده است.

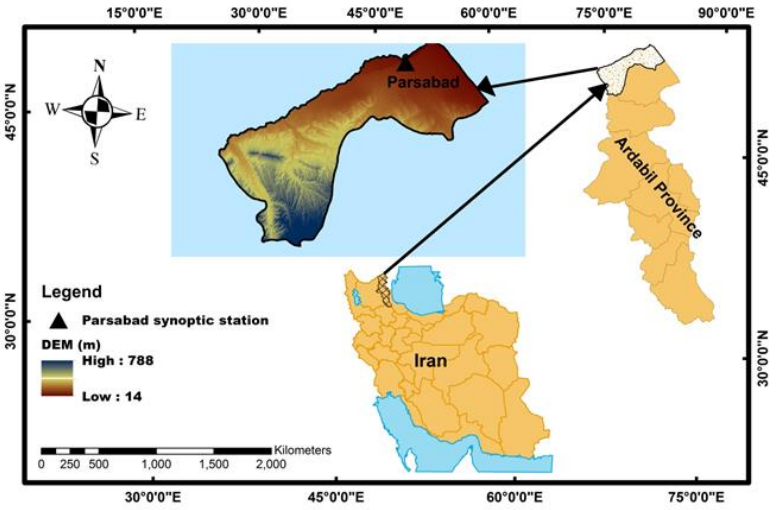
با توجه به مطالعات انجام‌شده در زمینه پیش‌بینی بارش، می‌توان به اهمیت برآورد این متغیر در میان سایر پارامترهای هواشناسی پی برد. بنابراین، پیش‌بینی وقوع بارش به ویژه در مقیاس روزانه با استفاده از داده‌های هواشناسی روزهای قبل در مدیریت سیل، خشکسالی، فعالیت‌های کشاورزی و... موثر خواهد بود. هدف پژوهش حاضر، پیش‌بینی وقوع بارش روزانه به کمک مدل درختی استدلالی (LMT) است. برای این منظور از داده‌های روزانه هواشناسی شهرستان پارس‌آباد در استان اردبیل استفاده شد و وقوع یا عدم وقوع بارش روزانه با استفاده از مدل درختی استدلالی و به کارگیری داده‌های هواشناسی روزهای قبل بررسی گردید.

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

مواد و روش‌ها

منطقه مورد مطالعه

در این مطالعه از داده‌های هواشناسی ایستگاه سینوپتیک پارس‌آباد استفاده شد. پارس‌آباد دومین شهر بزرگ استان اردبیل در شمال غربی ایران است. پارس‌آباد منطقه‌ای است که در جلگه‌ای هموار با آب و هوایی معتدل تا گرم قرار دارد. ایستگاه سینوپتیک پارس‌آباد درارتفاعی معادل ۷۲/۶ متر از سطح دریا در موقعیت جغرافیایی '۳۶ °۳۹ عرض جغرافیایی و '۴۶ °۴۷ طول شرقی واقع شده است (شکل ۱).



شکل (۱): موقعیت جغرافیایی ایستگاه سینوپتیک پارس‌آباد در استان اردبیل

در پژوهش حاضر، از داده‌های روزانه هواشناسی ۱۵ سال آماری مربوط به سال‌های ۲۰۱۹-۲۰۰۴ میلادی استفاده شده‌است. داده‌های روزانه شامل بارش (P)، متوسط رطوبت نسبی (RHmean) ، حداقل رطوبت نسبی (RHmin)، حداکثر رطوبت نسبی (RHmax)، متوسط دمای هوا (Tmean)، دمای حداقل (Tmin)، دمای حداکثر (Tmax)، مجموع ساعات آفتابی (Sun) و متوسط سرعت باد (Wind) هستند. لازم به ذکر است که داده‌های هواشناسی مورد استفاده بر اساس یافته‌های محققین قبلی (مانند مهتابی و همکاران، ۱۳۹۷ و ...) و در دسترس بودن آن‌ها انتخاب شدند. براساس داده‌های بارش، روزهای سال در دو کلاس روزهای وقوع بارش (Y) و روزهای عدم وقوع بارش (N) جای گرفتند و داده‌های هواشناسی در قالب ۳ سری داده با ۱ تا ۳ روز تاخیر برای پیش‌بینی وقوع و یا عدم وقوع بارش روزانه بررسی گردیدند.

درخت تصمیم

درخت تصمیم یکی از روش‌های داده‌کاوی و از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی می‌باشد که برخلاف شبکه‌های عصبی به تولید قانون می‌پردازد. درخت تصمیم پیش‌بینی خود را در قالب یکسری قوانین توضیح می‌دهد، درحالی که شبکه عصبی تنها پیش‌بینی خود را بیان می‌کند و چگونگی آن در خود شبکه پنهان باقی می‌ماند. علاوه بر آن در درخت تصمیم‌گیری بر خلاف شبکه‌های عصبی، می‌توان از داده‌های غیرعددی و کیفی

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

نیز استفاده نمود (دستورانی و همکاران ۱۳۹۱). فرآیند ایجاد یک درخت تصمیم شامل سه مرحله است: تقسیم گره‌ها، تعیین گره‌های انتهایی، تخصیص طبقه‌ها به گره‌های انتهایی. یک درخت تصمیم، ترکیبی از یک ریشه یا گره اصلی (دربرگیرنده تمام داده‌ها)، یک سری گره داخلی (شاخه‌ها) و یک سری گره‌های انتهایی (برگ‌ها) است. در درخت تصمیمی که عمل طبقه‌بندی را انجام می‌دهد، برگ‌ها بیانگر کلاس‌ها هستند و در هر یک از گره‌های دیگر (گره‌های غیربرگ) با توجه به یک یا چند صفت تصمیم‌گیری صورت می‌گیرد.

مدل درختی استدلالی

مدل درختی استدلالی[۲](LMT) یک مدل طبقه‌بندی است که ترکیبی از روش‌های یادگیری درخت تصمیم‌گیری و رگرسیون لوجستیک می‌باشد (Tien Bui et al., 2016). در این روش، از معیار information gain برای تقسیم استفاده می‌شود. الگوریتم LogisticBoost برای تولید مدل در هر گره از درخت بکار برده می‌شود و درخت با استفاده از الگوریتم CART هرس می‌شود (Breiman et al., 1984). LMT برای جلوگیری از اتصال بیش از حد داده‌های آموزشی از اعتبارسنجی متقابل در یافتن تعداد تکرار استفاده می‌کند. همچنین در این روش، الگوریتم استدلالی متناسب با حداقل مربعات برای هر کلاس مورد استفاده قرار می‌گیرد.

$$L_M(X) = \sum_{i=1}^{n} \beta_i X_i + \beta_0 \qquad (۱)$$

که β_i ضریب i امین جز بردار x و n تعداد عامل ها می‌باشد.

معیارهای ارزیابی مدل

در جدول (۱) معیارهای آماری مختلفی که در مطالعه حاضر به عنوان ارزیابی عملکرد مدل استفاده گردیده، نشان داده شده است (Rahmati et al., 2019 ; Kavitha et al., 2016). همچنین، پارامترهای جذر میانگین مربعات خطا[۳] و میانگین خطای مطلق[۴]، به منظور ارزیابی کلی خطای مدل درختی مورد مطالعه، استفاده گردید (روابط ۲ و ۳).

جدول (۱): معیارهای آماری مختلف به عنوان ارزیاب عملکرد مدلLMT

آماره	تعریف	معادله
مثبت‌های واقعی	نسبت مثبت واقعی که به درستی به عنوان فعال معرفی می‌شود.	TPR = TP/(TP+FN)
مثبت‌های کاذب	نسبت فعال نادرست پیش بینی شده به تعداد واقعی غیر فعال	FPR = FP/(FP+TN)
دقت	درصد نمونه های طبقه بندی شده صحیح که اثربخشی کلی طبقه بندی کننده را نشان می‌دهد.	CCI = (TP+TN)/(TP+TN+FP+FN)
ارزش پیش بینی مثبت یا دقیق	نسب نمونه هایی که واقعا از یک کلاس هستند به کل نمونه های طبقه بندی شده به عنوان آن کلاس	Precision = TP/(TP+FP)
فراخوانی	نسبت مثبت مثبت که مثبت پیش بینی شده است.	Recall = TP/(TP+FN)

[2] Logistic Model Tree
[3] Root Mean Squared Error
[4] Mean Absolute Error

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

F-Measure	اندازه گیری ترکیبی برای دقت و یادآوری یا میانگین هارمونیکی	F-Measure = (2 * Precision * Recall)/(Precision + Recall)
ROC	ناحیه ROC	Area under the curve (TPR vs. FPR)
آماره کاپا	معیار تصحیح تصادفی بین طبقه بندی و طبقات واقعی	Kappa = (Total CCI – Random CCI/(1- Random CCI)
ضریب همبستگی متیوز	تصادفی بودن مدل	MCC = (TP*TN-FN*FP)/((TP+FN)(TP+FP)(TN+FN)(TN+FP))^0.5
PRC	ناحیه PRC	Area under the curve (Precision vs. recall)

این معیارها شامل TP Rate[۵]، FP Rate[۶]، Precision، Recall، F-Measure، ROC Area[۷]، PRC Area[۸] می‌باشند.

نتایج و بحث

در مطالعه حاضر، امکان‌سنجی استفاده از پارامترهای هواشناسی رطوبت نسبی، دما، سرعت باد، مجموع ساعات آفتابی در ترکیب‌های مختلف به منظور پیش‌بینی وقوع بارش روزانه مد نظر قرار گرفت. به منظور پیش‌بینی وقوع بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل، سناریوهای مختلف پارامترهای ورودی در روش LMT بررسی گردید. سناریوها در قالب ۳ سری داده با ۱ تا ۳ روز تاخیر در نظر گرفته شدند (جدول ۲). در بررسی عملکرد انواع سناریوها با ترکیب‌های مختلف، با توجه به تعداد زیاد سناریوها، تنها نتایج سناریوهایی ارائه شد که درصد موارد صحیح آن‌ها نزدیک به ۸۰ درصد بود (جدول۱). سپس محاسبات واسنجی بر روی دو سوم از داده‌ها و محاسبات صحت‌سنجی بر روی یک سوم داده‌های باقیمانده با استفاده از نرم‌افزار Weka انجام گردید. این نرم‌افزار حاوی تعداد زیادی از تکنیک‌های یادگیری ماشین و داده‌کاوی است که امکان مقایسه‌ی تکنیک‌های یادگیری ماشین را فراهم می‌سازد (Written & Frank, 2000).

جدول (۲): سناریوهای تعریف شده در مدل LMT

شماره سناریو	پارامترهای ورودی	تاخیر	خروجی	پیش‌بینی شرایط بارش
۱	T_{max}, T_{min}, T_{mean}, $RH_{max,}$, RH_{min}, RH_{mean}, SS, U	t-1	P_t	۱ روز بعد
۲	T_{max}, T_{min}, T_{mean}, $RH_{max,}$, RH_{min}, RH_{mean}, SS, U	t-1, t-2	P_t	۲ روز بعد
۳	T_{max}, T_{min}, T_{mean}, $RH_{max,}$, RH_{min}, RH_{mean}, SS, U	t-1, t-2, t-3	P_t	۳ روز بعد

[5] True Positive Rate
[6] False Positive Rate
[7] Receiver Operating Characteristics
[8] Precision-Recall Curves

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

در جدول‌های (۳) و (۴)، جزئیات معیارهای ارزیابی روش LMT در پیش‌بینی وقوع و یا عدم وقوع بارش ۱ روز بعد، ۲روز بعد و ۳ روز بعد براساس سری داده‌های هواشناسی روزانه (تاخیر ۱ روزه، تاخیر ۲ روزه و تاخیر ۳ روزه) ارائه شده‌است.

جدول (۳): نتایج معیارهای مختلف الگوریتم LMT

PRC Area	ROC Area	MCC	F-Measure	Recall	Precision	FP Rate	TP Rate	Class	LMT
۰/۳۰۴	۰/۶۱۸	۰/۲۲۵	۰/۲۴۰	۰/۱۵۱	۰/۵۹۰	۰/۰۲۷	۰/۱۵۱	Y	Lag1
۰/۸۳۷	۰/۶۱۸	۰/۲۲۵	۰/۸۸۹	۰/۹۷۳	۰/۸۱۸	۰/۸۴۹	۰/۹۷۳	N	
۰/۲۷۳	۰/۵۸۴	۰/۱۸۸	۰/۲۴۴	۰/۱۶۳	۰/۴۸۲	۰/۰۴۵	۰/۱۶۳	Y	Lag2
۰/۸۲۷	۰/۵۸۴	۰/۱۸۸	۰/۸۸۱	۰/۹۵۵	۰/۸۱۷	۰/۸۳۷	۰/۹۵۵	N	
۰/۳۰۱	۰/۵۹۹	۰/۲۱۰	۰/۱۹۳	۰/۱۱۴	۰/۶۳۸	۰/۰۱۶	۰/۱۱۴	Y	Lag3
۰/۸۳۰	۰/۵۹۹	۰/۲۱۰	۰/۸۹۰	۰/۹۸۴	۰/۸۱۳	۰/۸۸۶	۰/۹۸۴	N	

جدول (۴): میانگین وزنی معیارهای مختلف الگوریتم LMT

PRC Area	ROC Area	MCC	F-Measure	Recall	Precision	FP Rate	TP Rate	LMT
۰/۷۲۹	۰/۶۱۸	۰/۲۲۵	۰/۷۵۷	۰/۸۰۶	۰/۷۷۲	۰/۶۸۲	۰/۸۰۶	Lag1
۰/۷۱۴	۰/۵۸۴	۰/۱۸۸	۰/۷۵۱	۰/۷۹۴	۰/۷۴۹	۰/۶۷۶	۰/۷۹۴	Lag2
۰/۷۲۳	۰/۵۹۹	۰/۲۱۰	۰/۷۴۹	۰/۸۰۷	۰/۷۷۸	۰/۷۰۹	۰/۸۰۷	Lag3

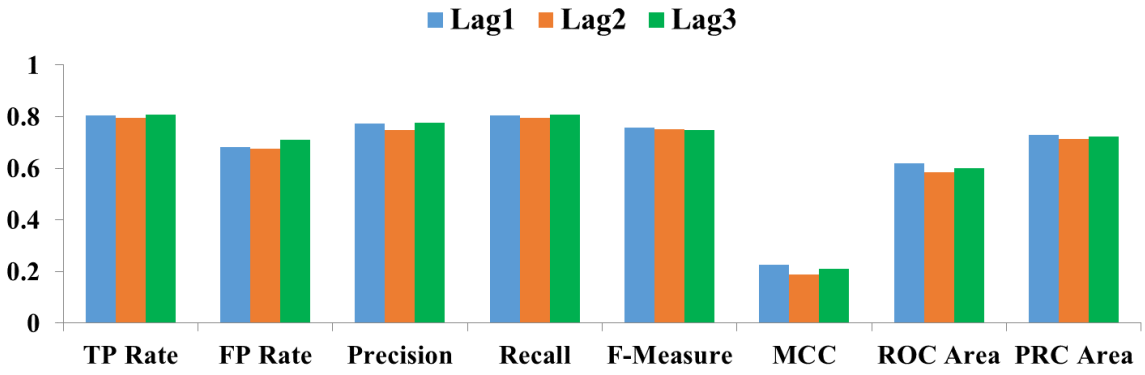
معیار Recall کارایی الگوریتم را با توجه به تعداد رخداد آن طبقه نشان می‌دهد. حال آنکه معیار Precision اساسا مبتنی بر دقت پیش‌بینی الگوریتم می‌باشد و بیانگر این موضوع است که به چه میزان می‌توان به خروجی‌های الگوریتم اعتماد کرد. معیار F-Measure ترکیب معیارهای فراخوانی و دقت را نشان می‌دهد. نمودار ROC روشی برای بررسی کارایی الگوریتم‌ها بوده و هرچه این مقدار بزرگتر باشد، کارایی نهایی الگوریتم مطلوب‌تر ارزیابی می‌شود. با توجه به نتایج به دست آمده، پیش‌بینی وقوع بارش با تاخیر ۱ روزه داده‌ها، با دقت ۷۷/۲ درصد، معیار Recall برابر با ۰/۸۰۶ و ROC برابر با ۰/۶۱۸ دقت بیشتری نسبت به پیش‌بینی وقوع بارش در شرایط تاخیر ۲ و ۳ روز داده‌ها، با دقت ۷۴/۹ و ۷۷/۸ درصد، معیار Recall برابر با ۰/۷۹۴ و ۰/۸۰۷ و ROC برابر با ۰/۵۸۴ و ۰/۵۹۹ داشته است. علاوه بر این، بالا بودن F-Measure در پیش‌بینی وقوع بارش در شرایط ۱ روز بعد خود گواه بر بالا

[Downloaded from jircsa.ir on 2026-02-21]

[DOR: 20.1001.1.24235970.1399.8.2.5.2]

بودن دقت آن است. بنابراین، با توجه به نتایج جدول‌های (۳) و (۴)، پیش‌بینی وقوع بارش با تاخیر ۱ روزه داده‌ها، دارای بهترین عملکرد و پیش‌بینی با تاخیر۲ روزه داده‌ها، دارای عملکرد ضعیف بوده است.

شکل (۲)، نمودار جزئیات دقت میانگین وزنی معیارهای الگوریتم LMT، برای هر سه حالت (تاخیرهای ۱ تا ۳ روزه)، را به‌صورت نمودار ستونی نشان می‌دهد.



شکل (۲): نمودار میانگین وزنی معیارهای الگوریتم LMT

میزان صحت الگوها از طریق آماره CCI[9] (درصد موارد پیش‌بینی صحیح) اندازه‌گیری می‌شود و با توجه به نتایج مندرج در جدول شماره ۴، بالاترین میزان صحت الگوهای مورد استفاده در زمینه پیش‌بینی وقوع بارش را تاخیر ۱ روزه و کمترین آن را تاخیر ۲ روزه داراست. همچنین، آماره Kappa یک معیار تصحیح شده بر مبنای تصادف برای تطابق بین طبقه‌بندی و کلاس‌های صحیح می‌باشد که مقدار بالای صفر این آماره نشان دهنده این است که طبقه‌بندی کننده موردنظر بهتر از حالت تصادفی کار می‌کند. هرچه این میزان بیشتر باشد، نشان‌دهنده عملکرد بهتر الگوریتم استفاده شده‌است. در نهایت، نتایج عملکرد شاخص‌ها برای مدل مورد استفاده در جدول (۵) ارائه شده است.

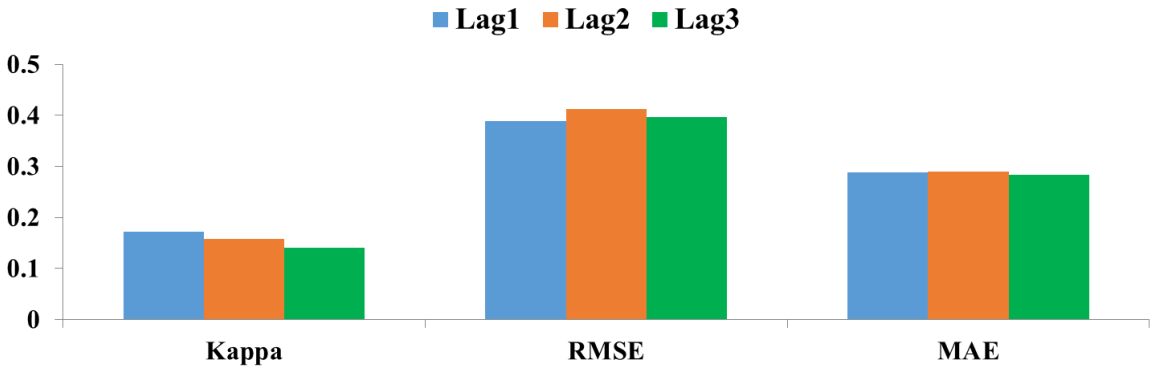
جدول (۵): نتایج عملکرد شاخص‌ها برای مدل LMT

Lag3	Lag2	Lag1	Index (LMT)
۸۰/۶۰۰	۷۹/۴۲۵	۸۰/۶۱۳	CCI
۰/۱۴۰۳	۰/۱۵۷	۰/۱۷۱۷	Kappa
۰/۳۹۶۷	۰/۴۱۱۳	۰/۳۸۹۱	RMSE
۰/۲۸۳۴	۰/۲۹	۰/۲۸۸۵	MAE

[9] Correctly Classified Instances

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

علاوه بر این، شکل (۳) نمودار ستونی عملکرد شاخص‌های Kappa، RMSE و MAE برای تاخیرهای مورد مطالعه با استفاده از مدل LMT را نشان می‌دهد.



شکل (۳): نمودار عملکرد شاخص‌ها در مدل LMT

با توجه به نتایج جدول (۵) و شکل (۳) می‌توان چنین نتیجه گرفت که پیش‌بینی وقوع بارش با تاخیر ۱ روزه داده‌ها، با شاخص عملکرد CCI برابر با ۸۰/۶۱۳، Kappa برابر با ۰/۱۷۱۷، RMSE برابر با ۰/۳۸۹۱ و MAE برابر با ۰/۲۸۸۵ عملکرد به مراتب بهتری نسبت به وقوع بارش با تاخیر ۲ و ۳ روزه داده‌ها داشته است. این مورد نشان دهنده آن است که شرایط آب و هوایی یک روز قبل نقش موثری در پیش‌بینی وقوع بارش روزانه داشته و می‌تواند برای اهداف پیش‌بینی وقوع بارش با اطمینان بالاتری مورد استفاده قرار گیرد. نتیجه گرفته شده مطابقت معنایی کاملی با نتایج ارائه شده توسط مهتابی و همکاران (۱۳۹۷) دارد که بیان نمودند که دقت پیش‌بینی مدل‌های مورد مطالعه با استفاده از داده‌های روزهای نزدیک (۱ تا ۵ روز قبل) بیشتر از حالتی بود که از داده‌های ۶ و ۷ روز قبل استفاده کرده بودند.

نتیجه گیری

پیش‌بینی بارش و شرایط وقوع آن از اهمیت ویژه‌ای برخوردار است، لذا در این مطالعه سعی گردید تا به پیش‌بینی شرایط وقوع بارش روزانه در ایستگاه پارس‌آباد به کمک مدل درختی استدلالی و با استفاده از داده‌های روزهای قبل (۱ تا ۳ روز) پرداخته شود. مقایسه نتایج تاخیرهای ۱ تا ۳ روزه نشان داد که با استفاده از داده‌های هواشناسی روزانه می‌توان وقوع بارش را با دقت مناسب و قابل قبولی (۸۰ درصد) پیش‌بینی نمود. دقت پیش‌بینی وقوع بارش با استفاده از داده‌های هواشناسی با تاخیر ۲ روزه نسبت به تاخیرهای ۱ و ۳ روزه پایین بود. در نهایت، می‌توان چنین نتیجه‌گیری نمود که روش LMT توانایی مناسبی در پیش‌بینی بارش داشته و می‌تواند با اطمینان قابل قبولی در پیش‌بینی وقوع و یا عدوم وقوع بارش با استفاده از داده‌های هواشناسی روزهای قبل به کار گرفته شود.

منابع

۱. خسروی، م. و ه. شکیبا (۱۳۸۹). پیش‌بینی بارش با استفاده از شبکه‌های مصنوعی به منظور مدیریت سیل (مورد منطقه ایرانشهر)، مجموعه مقالات چهارمین کنگره بین‌المللی جغرافیدانان جهان اسلام (ICIWG 2010)، ۲۵ و ۲۷ فروردین ماه، زاهدان.

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

۲. خلیلی، ن.، س.ر. خداشناس، ک. داوری و م. موسوی بایگانی (۱۳۸۹). پیش‌بینی بارش روزانه با استفاده از شبکه‌های عصبی مصنوعی (مطالعه موردی: ایستگاه سینوپتیک مشهد)، پژوهش‌های آبخیزداری (پژوهش و سازندگی)، ۸۹، ۱۵-۸.
۳. دستورانی، م. ت.، ا. حبیبی‌پور، م. ر. اختصاصی، ع. طالبی و ج. محجوبی (۱۳۹۱). بررسی کارایی مدل درخت تصمیم در پیش‌بینی بارش (مطالعه موردی ایستگاه سینوپتیک یزد)، تحقیقات منابع آب، ۸(۳)، ۲۷-۱۴.
۴. سلگی، ا.، ح. زارعی، م. شهنی دارابی و ص. علیدادی ده کهنه (۱۳۹۷). پیش‌بینی بارش ماهانه با استفاده از مدل برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان، نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۸(۵۰)، ۱۰۳-۹۱.
۵. مهتابی، ق.، ف. تاران و س. مظفری (۱۳۹۷). پیش‌بینی بارش روزانه با استفاده از داده‌های هواشناسی روزهای قبل (مطالعه موردی: شهر اصفهان)، فصل‌نامه جغرافیایی طبیعی ۱۱۰(۳۹)، ۱۱۴-۹۹.
۶. ناصری، م. (۱۳۸۲). پیش‌بینی بارندگی در دامنه زمان و مکان با استفاده از شبکه عصبی مصنوعی و الگوریتم ژنتیک، پایان‌نامه دوره کارشناسی ارشد، دانشگاه صنعتی اصفهان، گروه عمران.

7. Bhattacharya B. and Solomatine D.P. *(2005). Neuro networks and M5 model trees in modeling water level-discharge relationship.* Neuro computing, Vol. *63, pp 381-396.*
8. Breiman L., Friedman J.H., Olshen R.A. and Stone C.J. *(1984). Classification and regression trees Belmont.* Wadsworth international group,CA.
9. Chen W., Xie X., Wang J., Pradhan B., Hong H., Tien Bui D., Duan Z. and Ma J. *(2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility*. CATENA 151, *pp 147-160.*
10. Dash Y., Mishra S. and Panigrahi B. *(2018). Rainfall prediction for the Kerala state of India using artificial intelligence approaches*, Computers & Electrical Engineering, 70, pp 66-73.
11. Freiwan M. and Cigizoglu H.K. *(2005). Prediction of total monthly rainfallin Jordan using feed forward backpropagation method*. Fresenius Environmental Bulletin, *14(2), pp 142-151.*
12. George J., Letha J. and Jairaj P.G. *(2016). Daily rainfall prediction using generalized linear Bivariate model- A case sudy.* Procedia technology, *24, pp 31-38.*
13. Kavitha A.P., Abdul Jaleel U.C., Abdul Mujeeb V.M. and Muraleedharan K. *(2016). Performance of knowledge-based biological models in higher dimensional chemical space.* Chemometrics and Intelligent Laboratory Systems. *153, pp 58-66.*
14. Maeda N. k., Shunichi I., Kaoru. I., Shigekazu K. and Michihiro A. *(2001). Prediction by a neural network method* , Journal of natural disaster science, Vol *23(1), pp 22-33.*
15. Mekanik F., Imteaz M.A., Gato- Trinida S. and Elmahdi A. *(2013). Multiple regression and artificial neural network for long term rainfall forecasting using large scale climate modes.* Journal of Hydrology , *503, pp 11-21*.
16. Rahmati O., Kornejady A., Samadi M., Deo R.C., Conoscenti C., Lombardo L., Dayal K., Taghizadeh-Mehrjardi R., Pourghasemi H.R., Kumar S. and Bui D.T. *(2019). PMT: New analytical framework for automated evaluation of geo-environmental modelling approaches*. Science of The Total Environment 664, *pp 296-311*.
17. Shoombuatong W., Hongjaisee S., Barin F., Chaijaruwanich J. and Samleerat T. *(2012). HIV-1 CRF01_AE coreceptor usage prediction using kernel methods based logistic model trees.* Computers in Biology and Medicine *42(9), pp 885-889.*
18. Tien Bui D., Tuan T.A., Klempe H., Pradhan B. and Revhaug I. *(2016c). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree.* Landslides13(2), *pp 361-378.*
19. Written L. and Frank E. (2000). *Data mining: Practical machine learning tools and techniques with Java implementation,* Morgan kaufmann publishers.

[Downloaded from jircsa.ir on 2026-02-21]
[DOR: 20.1001.1.24235970.1399.8.2.5.2]

Evaluating the capabilities of Logistic Model Tree in predicting the occurrence probability of daily precipitation

Fatemeh Mikaeili[1], Saeed Samadianfard*[2]

[1]M.Sc. Student, Department of Water Engineering, Faculty of Agriculture, University of Tabriz

[2]Assistant Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz

Received: 2020/08

Accepted: 2020/09

Abstract

Due to the location of Iran in arid and semi-arid regions and the inhomogeneous distribution of precipitation, predicting the occurrence of precipitation is important, therefore, researchers are implementing novel methods to identify and predict this parameter accurately. Thus the purpose of the current study is to investigate the capabilities of Logistic Model Tree (LMT) in predicting the occurrence of daily precipitation at Parsabad station using 1 to 3-day meteorological data. For this purpose, meteorological data for 2004-2016 were collected, and three combined scenarios of meteorological parameters were considered for calibration and validation of the studied method. The results showed that the prediction accuracy of the best-case scenario using the data from 2 days ago was about 79%, however, with the data from 1 and 3 days ago, the daily precipitation was with 80% prediction accuracy. Finally, by investigating the evaluation criteria, scenario 1 with the input parameters of minimum, maximum and average relative humidity (%), temperature (oC), total sunshine hours, and wind speed (m/s) was determined as the most accurate scenario to predict daily precipitation.

Keywords: Correctly Classified Instances, Daily precipitation, Decision tree, LMT.

*Corresponding author email: s.samadian@tabrizu.ac.ir

[Downloaded from jircsa.ir on 2026-02-21]

[DOR: 20.1001.1.24235970.1399.8.2.5.2]