DAM AND TUNNEL CONFERENCE & EXPO همایش و نمایشگاه سد و تونل
ششمین همایش و نمایشگاه سد و تونل ایران
The 6th Dam and Tunnel Conference & Exhibition
16 - 17 February 2021
۲۸-۲۹ بهمن ماه ۱۳۹۹

بهبود عملکرد روش KNN در پیش بینی جریان ورودی به سد با استفاده از روشهای هوش مصنوعی

احسان ابراهیمی ، مجتبی شوریان

چکیده

پیش‌بینی جریان آب یکی از مهمترین بخش‌های مدیریت منابع آب می‌باشد. پیش‌بینی‌های دقیق در بلند مدت برای برنامه‌ریزی تامین و ذخیره سازی آب و در کوتاه مدت برای پیش‌بینی جریان‌های شدید و استفاده آن در سیستم‌های هشدار سیل حائز اهمیت می‌باشد. مدل‌های داده مبنا به عنوان روش‌هایی نسبتاً ساده اما قدرتمند به طور گسترده برای پیش‌بینی جریان آب استفاده می‌شوند. . روش K-همسایه نزدیک (K-NN-)یک روش یادگیری غیر پارامتری موثر است که در حل مسائل مختلف مورد استفاده قرار می‌گیرد. در این پژوهش، یک روش جدید برای انتخاب همسایه‌ها به نام K-همسایه نزدیک پویا (D-K-K-N-)معرفی شده که با استفاده از یک مدل S-V-M فواصل بهینه‌ای یافته شده و همسایگانی که در این فاصله بهینه وجود دارند برای هر مورد پیش‌بینی استفاده می‌شوند. عملکرد روش پیشنهادی با به کارگیری ۲ سال داده ورودی روزانه سد قشلاق در غرب ایران مورد ارزیابی قرار گرفته است. نتایج نشان می‌دهد روش پیشنهادی دقت پیش‌بینی را با کاهش خطای کلی (RMSE)به میزان ۶٪، بهبود می‌دهد که این بهبود در پیش‌بینی موارد حدی به ۸/۷ ٪ می‌رسد.

۱- مقدمه

یکی از مهمترین مراحل در برنامه‌ریزی منابع آبی، پیش‌بینی دقیق جریان ورودی به یک مخزن می‌باشد تا مدیریت مخزن برای مصارف آینده به صورت کارآمد انجام بگیرد. در مطالعات و پژوهش‌های مختلف از مدل‌های فیزیکی و داده مبنا برای پیش‌بینی جریان مورد استفاده قرار می‌گیرند. مدل‌های فیزیکی معمولا وقتگیر، پیچیده و دشوار هستند و پیاده‌سازی آن‌ها راحت نمی باشد. برای حل این مشکلات مدل‌های داده مبنا به صورت فراگیر مورد استفاده قرار گرفته‌اند. با اینکه مدلهای داده مبنا پارامترهای فیزیک و هیدرولوژیکی را شبیه‌سازی نمی‌کنند، اما نتایج آن‌ها تا حد مناسبی دقیق می‌باشد تا به برنامه‌ریزی و مدیریت منابع آب کمک قابل توجهی کنند. از مدل‌های پر استفاده برای پیش‌بینی جریان می‌توان به MLR ، روش‌های غیر پارامتری مانند KNN ، و پارامتری مانند ANN، ANFIS و SVMاشاره کرد. روش KNN یکی از روش‌های پر استفاده است که در مسائل کلاس‌بندی ، خوشه‌بندی و رگرسیون مورد استفاده قرار می‌گیرد.

۲- روش توسعه یافته KNN با همسایگی پویا (DKNN)

قدرت اصلی K-N-N-این واقعیت است که از یک روش ساده پیروی می‌کند و می‌تواند برای مشکلات خطی و غیرخطی به کار رود. اما نتایج این روش به شدت وابسته به انتخاب تعداد بهینه همسایه‌ها می‌باشد و افزایش خطا در هنگام پیش‌بینی مقادیر حدی به شدت افزایش پیدا می‌کند. این موضوع به دلیل کمبود این مقادیر در دادهای زمانی مورد استفاده است و کالیبره کردن مدل برای یافتن مقدار مناسب K ممکن است باعث شامل شدن همسایه‌هایی با عملکرد نا مناسب بر روی روند پیش‌بینی شوند.

روش پیشنهادی جدید یک روش پویای جدید است که همسایگان را در یک فاصله بهینه از متغیرهای پیش‌بینی انتخاب کرده و بنابراین تعداد همسایه‌های انتخاب شده ثابت نبوده و وابسته به متغیر پیش‌بینی مورد نظر است. این روش جدید دقت بهتری نسبت به K-N-N-کلاسیک داشته و از همه مهم‌تر اجازه استفاده از همسایه‌های بیشتر در محاسبات برای استخراج ویژگی‌های مورد نظر از داده‌ها را دارا می‌باشد.

روند کالیبراسیون روش جدید به شرح زیر است:

۱- انتخاب یک مجموعه داده کالیبراسیون

۲- با استفاده از روش بهینه‌سازی P-S-O، فاصله‌ای بهینه یافت می‌شود که همسایه‌های انتخاب شده در آن فاصله، حداقل خطای پیش‌بینی را برای هر کدام متغیرهای پیش‌بینی در داده‌های کالیبراسیون را بدهد. متغیر هدف در روش PSO، حداقل کردن خطای RMSE تعیین شده است.

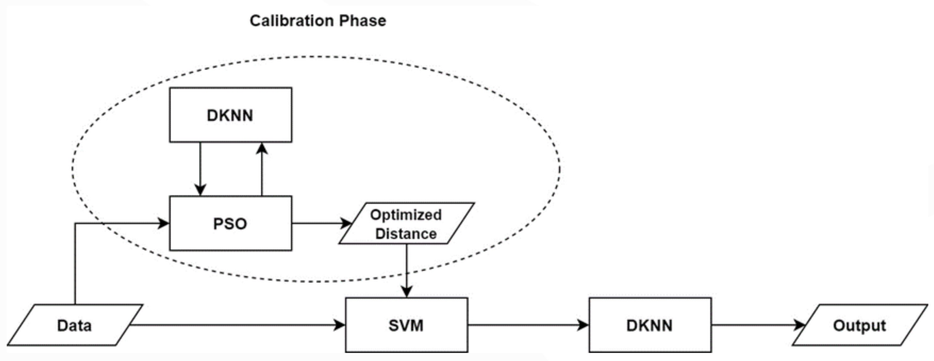
۳- از متغیرهای پیش‌بینی و فواصل بهینه یافته شده توسط P-S-O به عنوان ورودی و هدف (به ترتیب) برای ساخت یک مدل SVM استفاده می‌شود.

برای استفاده از مدل جدید DKNN برای موارد جدید پیش‌بینی به صورت زیر عمل می‌شود:

۱- ابتدا متغیرهای پیش‌بینی به مدل S-V-M ساخته شده داده شده و یک فاصله بهینه به عنوان خروجی مدل SVM به دست می‌آید.

۲- از این فاصله بهینه برای پیدا کردن همسایه‌هایی که در این فاصله محاسبه شده قرار می‌گیرند، استفاده می‌شود.

۳- ادامه محاسبات برای پیدا کردن مقدار پیش‌بینی، به همان روند KNN کلاسیک پیش می‌رود.



۳- نتایج

در جدول مقادیر شاخص‌های عملکردی مدل‌های کلاسیک و مدل جدید D-K-N-N-گزارش شده است. با توجه به جدول ، مدل جدید D-K-N-N-نسبت به مدل K-N-N-کلاسیک ۶ درصد بهبود داشته که عملکرد مناسب این مدل پیشنهادی جدید را نشان می‌دهد. در پیش‌بینی موارد حدی بالاتر از ۱۵ متر مکعب بر ثانیه که شاخص مهمی برای در عملکرد مدل‌های پیش‌بینی جریان آب است مدل پیشنهادی جدید عملکرد بهتری داشته و بهبود دقت به ۸/۷ درصد افزایش یافته که نشان دهنده نقطه قوت این روش برای داده‌های حدی می‌باشد.

Model	Nash-Sutcliffe	Total RMSE	Extreme Values RMSE	Low Values RMSE	Correlation Coefficient
DKNN	0.939	1.260	2.864	0.152	0.970
Classic KNN	0.931	1.340	3.105	0.140	0.969

٤- جمع‌بندی

با توجه به نتایج، استفاده از مدل جدید پیشنهادی D-K-N-N-نه تنها دقت پیش‌بینی را بالاتر می‌برد، بلکه با پویا شدن تعداد همسایگان در هر مورد از پیش‌بینی می‌توان اطلاعات و ویژگی‌های بیشتری از داده ها استخراج کرد. همچنین با حساس بودن تعداد همسایگان برای پیش‌بینی موارد حدی، پویایی دقت پیش‌بینی این موارد را بالاتر برده و این مدل عملکرد مناسبی برای استفاده در سیستم‌های هشدار سیل دارا می‌باشد. پیش‌بینی دست پایین مدل کلاسیک در موارد حدی خطرناک بوده، و می‌تواند باعث شکست سد و خسارات غیر قابل جبران شود.