

گزارش پروژه: پیش‌بینی بارش

پرهام طالبیان و بهزاد اسمی

۲۸ اردیبهشت ۱۴۰۳

۱ مقدمه

در این پروژه، هدف ما از تحلیل داده و ساخت مدل، پیش‌بینی وقوع بارش بر اساس شرایط آب و هوایی است. بارش یکی از مهم‌ترین وقایع آب و هوایی است که تأثیر زیادی بر زندگی مردم و روند اقتصادی دارد. بررسی دقیق شرایط آب و هوایی و پیش‌بینی بارش می‌تواند به مردم و مدیران آب و هوا در اتخاذ تصمیمات مناسب کمک کند.

۲ تحلیل داده

در این مرحله، داده‌های مربوط به شرایط آب و هوایی را از منابع مختلف جمع‌آوری کرده‌ایم، از جمله داده‌های مشاهده شده از ایستگاه‌های هواشناسی. سپس، ابتدا داده‌ها را پیش‌پردازش کرده و سپس تحلیل آماری انجام داده‌ایم. این تحلیل شامل مطالعه توزیع‌های مختلف داده‌ها، بررسی روندها و الگوهای زمانی، و تعیین ویژگی‌های مهم برای مدل‌سازی است.

درصد null بودن در هر ویژگی را محاسبه کردیم. و ستون‌هایی که بیشتر از بیست درصد null دارند را حذف و در ردیف‌هایی که null دارد نیز حذف کردیم.

دوباره ایندکس بندی کردیم.

با استفاده از Label encoding داده‌های اسمی را به عدد تبدیل کردیم.

و Atmosphere pressure را نرمال کردیم تا داده‌های ما در بازه کمتری باشد. سپس همبستگی بین داده‌ها را روی هیت مپ پلات کردیم. در قسمت بعد داده‌های نویز و پرت را حذف کردیم. سپس دوباره ایندکس‌ها را ریست کردیم

۳ ساخت مدل

پس از تحلیل داده، سه مدل مختلف برای پیش‌بینی بارش پیاده‌سازی شدند:

- مدل (Support Vector Machine) SVM

- مدل (K-Nearest Neighbors) KNN

- مدل درخت تصمیم

سپس، با استفاده از داده‌های آموزشی، هر یک از این مدل‌ها آموزش داده شدند و پارامترهای بهینه آن‌ها تنظیم شد.

۴ ارزیابی مدل

در این مرحله، عملکرد هر یک از مدل‌ها با استفاده از معیارهای دقت، صحت، و $F1$ -score ارزیابی شد. نتایج ارزیابی برای هر مدل گزارش شده است و مدلی که بهترین عملکرد را داشته است مشخص شده است.

۵ نتیجه‌گیری

در این پروژه، با تحلیل دقیق داده‌های مربوط به شرایط آب و هوایی و ساختن مدل‌های پیش‌بینی، به ارزیابی عملکرد مدل‌های مختلف برای پیش‌بینی بارش پرداختیم. نتایج نشان داد که مدل SVM بهترین عملکرد را داشته و می‌تواند به‌طور قابل توجهی دقت پیش‌بینی را افزایش دهد. این پروژه می‌تواند به متخصصان هواشناسی و مدیران در اتخاذ تصمیمات مربوط به مدیریت منابع آب و هوا کمک شایانی کند.