

TASODIFIY O'RMON (RANDOM FOREST) YORDAMIDA OB-HAVONI BASHORAT QILISH

Muallif; Jo'raho'ziyev Avazbek

Annotatsiya:

Ushbu maqolada tasodifiy o'rmon (Random Forest) algoritmidan foydalangan holda ob-havoni bashorat qilish jarayoni batafsil yoritilgan. Maqolada Random Forest algoritmining ishlash tamoyillari, ma'lumotlarni tayyorlash va qayta ishlash bosqichlari, shuningdek, prognoz aniqligini oshirish usullari tahlil qilingan. Tadqiqot jarayonida meteorologik parametrlar, masalan, harorat, namlik, shamol tezligi, atmosfera bosimi va yog'ingarchilik miqdori kabi ko'rsatkichlar asosida ma'lumotlar to'plami tuzilgan. Ushbu algoritmnining ansambl yondashuvi va ovoz berish mexanizmi orqali prognozlarning barqarorligi va aniqligi oshiriladi. Tadqiqot natijalari Random Forest algoritmining ob-havo bashoratida yuqori samaradorlikka ega ekanligini, murakkab meteorologik munosabatlarni aniqlashda kuchli vosita ekanligini tasdiqlaydi.

Kalit so'zlar: tasodifiy o'rmon, ob-havo bashorati, mashinali o'qitish, klassifikatsiya, ma'lumotlar to'plami, ansambl modellari, prognoz aniqligi.

Аннотация:

В данной статье рассматривается процесс прогнозирования погоды с использованием алгоритма случайного леса (Random Forest). Анализируются принципы работы алгоритма, этапы подготовки и обработки данных, а также методы повышения точности прогнозирования. В исследовании используются метеорологические параметры, такие как температура, влажность, скорость ветра, атмосферное давление и количество осадков. Благодаря ансамблевому подходу и механизму голосования Random Forest достигается высокая стабильность и точность прогнозов. Результаты исследования подтверждают, что алгоритм случайного леса является эффективным инструментом для

прогнозирования погоды и способен выявлять сложные зависимости между метеорологическими показателями.

Ключевые слова: случайный лес, прогноз погоды, машинное обучение, классификация, набор данных, ансамблевые модели, точность прогноза.

Annotation:

This article discusses the process of weather forecasting using the Random Forest algorithm. It analyzes the working principles of the algorithm, stages of data preparation and processing, and methods to enhance prediction accuracy. The study involves meteorological parameters such as temperature, humidity, wind speed, atmospheric pressure, and precipitation levels. The ensemble approach and voting mechanism of Random Forest ensure high stability and reliability of forecasts. The results demonstrate that the Random Forest algorithm is an effective tool for weather prediction and capable of identifying complex relationships among meteorological variables.

Keywords: random forest, weather forecasting, machine learning, classification, dataset, ensemble models, prediction accuracy.

1.

KIRISH

So'nggi yillarda mashinali o'qitish texnologiyalari turli sohalarda, jumladan, meteorologiyada ham keng qo'llanilmoqda. Ob-havoni aniqlik bilan bashorat qilish inson faoliyatining muhim qismi hisoblanadi, chunki transport, qishloq xo'jaligi, energetika va favqulodda vaziyatlarni boshqarish kabi sohalarda to'g'ri prognozlar ish samaradorligini oshiradi va xavfsizlikni ta'minlaydi. Shu sababli, meteorologik ma'lumotlarni tahlil qilish va aniqlik bilan bashorat qilish uchun ilg'or mashinali o'qitish usullari tobora keng qo'llanilmoqda.

Random Forest algoritmi — bu aynan shunday vazifalarni bajarishda yuqori natijalarni taqdim etuvchi samarali usullardan biridir. U ansambl yondashuviga asoslangan bo'lib, bir nechta qaror daraxtlarining natijasini birlashtirish orqali yakuniy prognozni hosil qiladi. Har bir daraxt alohida ma'lumot to'plamida o'qitiladi va o'z mustaqil qarorini beradi. Yakuniy natija esa barcha daraxtlarning ovoz berish

mexanizmi yordamida aniqlanadi. Shu tarzda, bitta daraxtning ekstremal yoki noto'g'ri prognozi ansambl natijasiga minimal ta'sir ko'rsatadi.

Bu yondashuvning eng muhim afzalliklaridan biri modelning haddan tashqari moslashib ketishi (overfitting) ehtimolini sezilarli darajada kamaytirishi hisoblanadi. Shuningdek, Random Forest murakkab meteorologik jarayonlar va nolinear munosabatlarni o'rganishga qodir, bu esa qisqa va o'rta muddatli ob-havo prognozlarini yuqori aniqlik bilan chiqarishga yordam beradi. Algoritmning ansambl tuzilmasi va ovoz berish mexanizmi prognozlarning barqarorligini oshiradi, shuningdek, ma'lumotlardagi ekstremal qiymatlar yoki shovqinli natijalar ta'sirini kamaytiradi.

Bundan tashqari, Random Forest algoritmi meteorologik ko'rsatkichlarning prognozga ta'sirini baholash imkonini beradi, ya'ni qaysi parametrlar prognoz natijasiga ko'proq ta'sir qilayotganini aniqlash mumkin. Bu esa modelning ilmiy talqini va amaliy qo'llanilishini osonlashtiradi. Shu bilan birga, algoritm katta hajmdagi ma'lumotlar bilan ishlay oladi, bu esa yirik meteorologik ma'lumotlar to'plamlarida prognoz aniqligini oshirish imkonini beradi.

Natijada, Random Forest algoritmi ob-havoni bashorat qilishda yuqori samaradorlikka ega bo'lib, murakkab meteorologik o'zgarishlarni aniqlik bilan prognoz qilish imkonini beradi va ilmiy hamda amaliy jihatdan ishonchli vosita sifatida namoyon bo'ladi.

Random Forest shuningdek, ma'lumotlardagi murakkab, nolinear bog'lanishlarni aniqlashga ham qodir bo'lib, bu meteorologik ma'lumotlar bilan ishlashda katta afzallik beradi. Natijada, algoritm qisqa va o'rta muddatli ob-havo prognozlarida yuqori aniqlikni ta'minlaydi.

2.

ASOSIY

QISM

Random Forest algoritmining mohiyati

Random Forest — bu bir nechta qaror daraxtlaridan tashkil topgan ansambl mashinali o'qitish usuli bo'lib, klassifikatsiya va regressiya masalalarida yuqori aniqlik

bilan ishlaydi. Har bir daraxt alohida ma'lumot to'plamida o'qitiladi va yakuniy qaror barcha daraxtlarning natijasini birlashtiruvchi ovoz berish mexanizmi orqali aniqlanadi. Shu tarzda, bitta daraxtning xatolari ansambl natijasiga kam ta'sir qiladi va model barqarorroq bo'ladi.

Random Forest murakkab va nolinear bog'lanishlarni o'rganishga qodir bo'lib, bu ko'p o'zgaruvchili ma'lumotlar bilan ishlashda, jumladan meteorologik parametrlarni prognoz qilishda muhim afzallik beradi. Shuningdek, algoritmi har bir atributning prognozga ta'sirini baholash imkoniyatini beradi, bu esa model talqinini osonlashtiradi. Ob-havoni bashorat qilish jarayoni

Ob-havoni bashorat qilish jarayoni bir nechta bosqichdan iborat. Avvalo, harorat, atmosfera bosimi, nisbiy namlik, shamol tezligi va yog'ingarchilik kabi meteorologik parametrlar bo'yicha ma'lumotlar yig'iladi. Keyin bu ma'lumotlar tozalanadi, yo'qotilgan qiymatlar to'ldiriladi va modelni o'qitishga tayyorlash jarayoni amalga oshiriladi.

Random Forest algoritmi tayyorlangan ma'lumotlarga asoslanib, yangi kun yoki soat uchun ob-havo parametrlarini prognoz qiladi. Har bir daraxt alohida prognoz chiqaradi va barcha daraxtlarning natijalari birlashtiriladi. Shu tarzda, algoritmi murakkab meteorologik munosabatlarni aniqlash va xatolarni kamaytirish imkoniyatiga ega bo'ladi. Model samaradorligini baholash uchun bir nechta mezonlar qo'llaniladi. Aniqlik — modelning to'g'ri prognoz bergan holatlarning umumiy natijaga nisbati. Aniqlik darajasi — prognoz qilingan natijalar orasida to'g'ri natijalar ulushi. Xatolik darajasi — prognoz va real natija o'rtasidagi farqni ko'rsatadi.

Tahlil natijalari shuni ko'rsatadiki, Random Forest algoritmi ob-havo bashoratida boshqa oddiy modellarga nisbatan ancha yuqori aniqlik beradi. Bu yuqori samaradorlik algoritmining ansambl yondashuvi, daraxtlar soni va ovoz berish mexanizmi bilan bog'liq. Shu bilan birga, Random Forest algoritmi qisqa muddatli prognozlar uchun ayniqsa samarali bo'lib, murakkab ob-havo sharoitlaridagi o'zgarishlarni ham aniq bashorat qila oladi.

3. XULOSA VA TAVSIYALAR

Random Forest algoritmi ob-havoni bashorat qilishda samarali vosita hisoblanadi. U katta hajmdagi ma'lumotlar bilan ishlash, xatolikni kamaytirish va modelning umumiy barqarorligini ta'minlash imkonini beradi. Kelgusida neyron tarmoqlar bilan integratsiyalash orqali aniqlikni yanada oshirish mumkin. Random Forest modeli bir nechta qaror daraxtlaridan tashkil topgan bo'lib, har bir daraxt alohida ma'lumot namunalarida o'qitildi. Modelni optimallashtirish maqsadida daraxtlar soni, chuqurligi, tanlanadigan belgilar soni kabi giperparametrlar moslashtirildi. Algoritmnining asosiy tamoyili shundan iboratki, bir nechta mustaqil daraxtlar natijasi birlashtiriladi va ovoz berish (voting) mexanizmi yordamida yakuniy prognoz aniqlanadi. Shu tarzda, bitta daraxtning ekstremal xatolari umumiy natijaga kam ta'sir qiladi. Random Forest modeli bir nechta qaror daraxtlaridan tashkil topgan bo'lib, har bir daraxt alohida ma'lumot namunalarida o'qitiladi. Bu yondashuv har bir daraxtni mustaqil ravishda tayyorlashga imkon beradi, natijada ansamblning umumiy natijasi individual daraxt xatolaridan kamroq ta'sirlanadi. Modelni optimallashtirish maqsadida daraxtlar soni, chuqurligi, tanlanadigan atributlar soni kabi giperparametrlar moslashtiriladi. Daraxtlar sonining ko'payishi model barqarorligini oshiradi, ammo hisoblash resurslarini ko'proq talab qiladi. Daraxtning chuqurligi esa har bir daraxtning murakkab qarorlarni o'rganish qobiliyatini belgilaydi, shuningdek ortiqcha moslashuv ehtimolini ham nazorat qiladi. Tanlanadigan atributlar soni esa har bir bo'linishda ko'rib chiqiladigan ma'lumotlar sonini belgilaydi va modelning umumiy samaradorligiga ta'sir qiladi.

Algoritmnining asosiy tamoyili shundan iboratki, bir nechta mustaqil daraxtlar natijasi birlashtiriladi va ovoz berish (voting) mexanizmi yordamida yakuniy prognoz aniqlanadi. Shu tarzda, bitta daraxtning ekstremal xatolari umumiy natijaga kam ta'sir qiladi, bu esa prognozlarning aniqligi va barqarorligini oshiradi. Shu bilan birga, Random Forest algoritmi parametrlarning ahamiyatini baholashga ham imkon beradi,

ya'ni qaysi meteorologik ko'rsatkichlar prognozga ko'proq ta'sir qilayotganini aniqlash mumkin.

Natijada, Random Forest algoritmi ob-havoni bashorat qilishda yuqori aniqlik va barqarorlikni ta'minlovchi kuchli vosita sifatida namoyon bo'ladi, va kelajakda uni boshqa ilg'or mashinali o'qitish yondashuvlari bilan integratsiyalash orqali yanada samarali qilish mumkin.

Taklif va tavsiyalar:

Ob-havo bashorati jarayonida ma'lumotlarni tozalash va to'ldirish bosqichiga alohida e'tibor qaratish zarur. Meteorologik ma'lumotlar ko'pincha yetishmaydigan qiymatlar yoki shovqinli o'lchovlar bilan ifodalanadi. Agar bu qiymatlar e'tiborsiz qoldirilsa, model noto'g'ri prognozlar chiqarishi mumkin. Shu sababli, ma'lumotlarni tozalash jarayonida noto'g'ri yoki yetishmayotgan ma'lumotlarni aniqlash, ularni oqilona to'ldirish va kerak bo'lsa, statistik yoki interpolatsiya usullari yordamida to'g'rilash muhim hisoblanadi. Tozalangan va yaxshilangan ma'lumotlar modeli barqarorroq bo'lib, prognozlar aniqligi sezilarli darajada oshadi. Shuningdek, model parametrlarini optimallashtirish jarayonida kross-valitatsiya usulidan foydalanish tavsiya etiladi. Kross-valitatsiya yordamida model turli ma'lumotlar to'plamlarida sinovdan o'tkaziladi va uning umumiy samaradorligi baholanadi. Bu yondashuv modelning ortiqcha moslashuvini kamaytirishga, turli sharoitlarda barqaror natijalar berishga imkon yaratadi. Shu bilan birga, kross-valitatsiya algoritmini optimallashtirishda giperparametrlarni aniqlash va modelning ishlashini yaxshilash uchun samarali vosita hisoblanadi. Ob-havo parametrlarini kengaytirish, masalan, geografik omillarni qo'shish ham model aniqligini oshirishga yordam beradi. Geografik joylashuv, balandlik, yer sirtining xususiyatlari va hududiy iqlim o'zgarishlarini hisobga olish orqali model murakkab meteorologik jarayonlarni yanada aniqroq prognoz qilishi mumkin. Shuningdek, boshqa qo'shimcha parametrlar, masalan, suv sirtining harorati, o'simlik qoplamining turi yoki shahar va qishloq hududining tabiiy sharoitlari kabi ko'rsatkichlar prognoz sifatini oshirishga xizmat qiladi. Bundan tashqari, modelni

muntazam yangilab turish va yangi ma'lumotlar bilan qayta o'qitish tavsiya etiladi. Bu yondashuv meteorologik o'zgarishlarni doimiy kuzatish va modelning real sharoitlarga moslashuvchanligini ta'minlashga yordam beradi. Shu tarzda, Random Forest algoritmi nafaqat mavjud ma'lumotlar asosida, balki kengaytirilgan va yangilangan ma'lumotlar yordamida ham yuqori aniqlik bilan ob-havoni bashorat qilish imkoniyatiga ega bo'ladi.

Natijada, ma'lumotlarni tozalash, parametrlarni optimallashtirish, geografik va boshqa qo'shimcha omillarni hisobga olish hamda modelni doimiy yangilash ob-havo prognozlarining sifatini oshirish va ilmiy hamda amaliy jihatdan ishonchli natijalar olish uchun muhim tavsiyalar hisoblanadi.

Foydalanilgan

adabiyotlar:

[1] Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32.

[2] Zhang, X., & Wang, Y. (2019). Weather Prediction Using Machine Learning Techniques. IEEE Access.

[3] Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

Muallif: Avazbek Jo'raqo'ziyev

Talaba, Farg'ona davlat texnika universiteti

Telefon: +998 339931010

Email: jorayevavazbek99@gmail.com

Yo'nalish: Mashinali o'qitish tamoyillari