



K-NN ALGORITMI BILAN KASALLIK TASHXISI

MUALLIF: MUXTORALIYEV MIRODILJON

EMAIL: [mmirodiljan@gmail.com]

TA'LIM MUASSASASI: FARG'ONA DAVLAT TEXNIKA UNIVERSITETI

AXBOROT TIZIMLARI VA TEXNOLOGIYALARI YO'NALISHI, 3-KURS

TALABASI

Annotatsiya

Ushbu maqolada K-Nearest Neighbors (K-NN) algoritmi yordamida kasallik tashxisini aniqlash jarayoni keng va chuqur yoritilgan. Tadqiqotda K-NN algoritmining nazariy asoslari, tibbiyotda qo'llanilishi, klinik ma'lumotlarni tayyorlash, xususiyatlarni tanlash, modelni qurish va baholash bosqichlari batafsil tushuntiriladi. K-NN algoritmi yurak kasalliklari, diabet, saraton va boshqa kasalliklar tashxisida qanday ishlatilishi misollar bilan ko'rsatilgan. Maqolada shuningdek, modelning aniqligi, F1-score, precision, recall, ROC-AUC kabi baholash metrikalari tahlil qilinadi. Tadqiqot natijalari shuni ko'rsatadiki, K-NN algoritmi sodda, samarali va intuitiv bo'lib, klinik qarorlar qabul qilishda foydali vosita hisoblanadi. Ushbu maqola tibbiyot sohasida sun'iy intellekt va mashinali o'qitish texnologiyalarini qo'llashga qiziqqan talabalar va mutaxassislar uchun amaliy va nazariy manba hisoblanadi.

Kalit so'zlar: K-NN, kasallik tashxisi, mashinali o'qitish, klinik ma'lumotlar, xususiyatlar muhandisligi, model baholash, precision, recall, ROC-AUC.

Annotation

This article provides a comprehensive overview of using the K-Nearest Neighbors (K-NN) algorithm for disease diagnosis. The study explains K-NN's theoretical foundations, application in medical diagnostics, data preprocessing, feature selection, model construction, and evaluation. K-NN applications include diagnosis of heart disease, diabetes, cancer, and other medical conditions.



Evaluation metrics such as accuracy, F1-score, precision, recall, and ROC-AUC are analyzed. The results show that K-NN is simple, effective, and intuitive, making it a valuable tool in clinical decision-making. This article serves as a practical and theoretical resource for students and professionals in healthcare interested in artificial intelligence and machine learning.

Keywords: K-NN, disease diagnosis, machine learning, clinical data, feature engineering, model evaluation, precision, recall, ROC-AUC.

Аннотация

В данной статье подробно рассматривается применение алгоритма K-Nearest Neighbors (K-NN) для диагностики заболеваний. Исследование описывает теоретические основы K-NN, его применение в медицинской диагностике, подготовку данных, выбор признаков, построение модели и оценку. Примеры применения K-NN включают диагностику сердечных заболеваний, диабета, рака и других заболеваний. Также анализируются метрики оценки модели: точность (accuracy), F1-score, precision, recall и ROC-AUC. Результаты исследования показывают, что K-NN является простым, эффективным и интуитивно понятным алгоритмом, полезным для принятия клинических решений. Статья является практическим и теоретическим источником для студентов и специалистов, интересующихся применением искусственного интеллекта и машинного обучения в медицине.

Ключевые слова: облачная платформа, виртуализация, вычислительные ресурсы, сеть, безопасность, управление инфраструктурой

KIRISH

Kasallik tashxisi tibbiyot sohasining eng muhim jarayonlaridan biri hisoblanadi. To'g'ri tashxis bemorning sog'lig'i, davolash strategiyasi va klinik natijalar uchun hal qiluvchi ahamiyatga ega. So'nggi yillarda mashinali o'qitish algoritmlari, sun'iy intellekt va katta ma'lumotlarni tahlil qilish (big data) tibbiyotda



diagnostika jarayonlarini tezkor, aniq va samarali qilish imkonini beradi. K-NN algoritmi intuitiv va sodda yondashuvga ega bo'lib, klinik ma'lumotlar asosida yangi bemorning tashxisini aniqlashda qo'llaniladi.

K-NN algoritmi har bir yangi bemor ma'lumotini mavjud ma'lumotlar to'plamidagi eng yaqin qo'shnilar (k-yaqin) bilan taqqoslaydi va ko'pchilik toifaga binoan tashxisni belgilaydi. Bu algoritm, ayniqsa, bemorlarning simptomlari va laboratoriya natijalari oldindan ma'lum bo'lgan holatlarda samarali ishlaydi.

ASOSIY QISM

Ma'lumotlarni tayyorlash jarayoni: Bemorlarning klinik ma'lumotlari yig'iladi, yetishmayotgan yoki noto'g'ri qiymatlar tozalanadi. Kategorik xususiyatlar kodlanadi, raqamli qiymatlar normalizatsiyadan o'tkaziladi. Trening va test to'plamlari ajratiladi.

Xususiyatlar muhandisligi: Bemorning yoshi, jinsi, qon bosimi, qand miqdori, yurak ritmi, lipid profili kabi parametrlar tanlanadi. Ba'zi holatlarda xususiyatlar kombinatsiyasi yoki derivativ parametrlar ham yaratiladi, bu esa model aniqligini oshiradi.

Model qurish: K-NN algoritmi har bir test bemor uchun k-yaqin qo'shni bemorlarni aniqlaydi. Masofa o'lchovlari sifatida Euclidean, Manhattan yoki Minkowski ishlatiladi. K qiymati optimallashtiriladi, cross-validation orqali eng yaxshi natija aniqlanadi.

Modelni baholash: Model accuracy, F1-score, precision, recall va ROC-AUC metrikalari bilan baholanadi. Yurak kasalliklari bo'yicha K-NN modeli 85–90% aniqlik bilan ishlashi mumkin.

Amaliy qo'llanilishi: K-NN algoritmi diabet tashxisida bemorning qon shakar miqdori, BMI va boshqa xususiyatlar asosida “normal” yoki “yuqori risk” guruhiga ajratadi. Saraton tashxisida esa hujayra biomarkerlari va klinik simptomlar tahlil qilinadi.



XULOSA

K-NN algoritmi kasallik tashxisida samarali va intuitiv vosita hisoblanadi. To'g'ri xususiyatlar tanlangan, ma'lumotlar tozalangan va K parametri optimal belgilangan taqdirda algoritm yuqori aniqlik ko'rsatadi. K-NN algoritmi klinik qarorlarni qo'llab-quvvatlash, bemorlarni tezkor tashxis qilish va profilaktik choralarni belgilashda foydali. Kelajakda K-NN algoritmi va boshqa mashinali o'qitish metodlari birgalikda ishlatilsa, diagnostika jarayoni yanada samarali va aniq bo'lishi kutilmoqda. Ushbu maqola tibbiyot sohasida AI va mashinali o'qitish texnologiyalarini qo'llashga qiziqqan talabalar va mutaxassislar uchun nazariy va amaliy manba bo'lib xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR

1. Géron, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, 2019.
2. Raschka, S., Mirjalili, V. Python Machine Learning. Packt Publishing, 2019.
3. Brownlee, J. Mastering Machine Learning Algorithms. Machine Learning Mastery, 2020.
4. Altman, N., Krzywinski, M. Points of Significance: K-Nearest Neighbors. Nature Methods, 2017.
5. Aggarwal, C.C. Data Mining: The Textbook. Springer, 2015.