

IOT ƏSASLI AĞILLI KİTABXANA MÜHİTLƏRİNDƏ SENSOR ƏHƏMİYYƏTİNİN İNTERPRETASIYASI VƏ MAŞIN ÖYRƏNMƏSİ MODELƏRİNİN HESABLAMA SƏMƏRƏLİLİYİNİN MÜQAYİSƏLİ TƏHLİLİ

Sərkən MƏMMƏDOV 

Naxçıvan Dövlət Universiteti, Naxçıvan, Azərbaycan

*Yazışılan müəllif: serkan.m@ndu.edu.az

NƏŞR TARİXİ:

Qəbul edilmə tarixi:
06.06.2026

Nəşr edilmə tarixi:
25.06.2026

Açar sözlər:

IoT, Maşın
öyrənməsi, Sensor
analizi, Feature
importance,
Permutation
importance

XÜLASƏ

Bu tədqiqat IoT əsaslı ağıllı kitabxana mühitlərində ətraf mühit sensorlarının maşın öyrənməsi modellərinin qərarvermə prosesinə təsirinin interpretasiyasını və hesablama səmərəliliyinin qiymətləndirilməsini hədəfləyir. Düzce Universiteti Kitabxanasında toplanmış çoxölçülü sensor məlumatları əsasında yeddi fərqli maşın öyrənməsi alqoritmi KNN, Random Forest, Decision Tree, SVM, XGBoost, Logistic Regression və Naive Bayes müqayisəli şəkildə tətbiq edilmişdir. Tədqiqatın əsas elmi töhfəsi sensor əhəmiyyətinin qiymətləndirilməsində iki yanaşmanın, yəni Random Forest Feature Importance və Permutation Importance metodlarının paralel istifadəsidir. Bu yanaşma sensor dəyişənlərinin model daxili qərar mexanizmi ilə real performans təsirini daha dəqiq analiz etməyə imkan vermişdir. Nəticələr göstərir ki, Feature Importance yanaşması temperatur və insan sıxlığı kimi dəyişənləri daha dominant faktorlar kimi qiymətləndirərkən, Permutation Importance analizi işıq intensivliyi və CO₂ əsaslı dəyişənlərin model performansına daha yüksək təsir göstərdiyini ortaya qoyur. Bu fərqlilik sensor interpretasiyasında yalnız model daxili ölçülərin kifayət etmədiyini və daha etibarlı qiymətləndirmə metodlarına ehtiyac olduğunu göstərir. Bundan əlavə, hesablama səmərəliliyi baxımından KNN və Decision Tree modelləri aşağı resurs tələbi və yüksək icra sürəti ilə real vaxt IoT tətbiqləri üçün daha uyğun nəticələr nümayiş etdirmişdir. Ümumilikdə tədqiqat göstərir ki, IoT əsaslı sistemlərdə optimal model seçimi yalnız proqnoz dəqiqliyi deyil, həm də sensor interpretasiyası və hesablama səmərəliliyi birlikdə nəzərə alınaraq aparılmalıdır.

GİRİŞ

Müasir universitet kitabxanaları yalnız informasiya resurslarına çıxış təmin edən məkanlar deyil, eyni zamanda tələbə və akademik heyətin intensiv tədqiqat, öyrənmə və elmi fəaliyyətlərini həyata keçirdiyi əsas akademik mühitlər kimi çıxış edir. Bu mühitlərdə temperatur, səs səviyyəsi, işıqlandırma, CO₂ konsentrasiyası və məkan doluluğu kimi ətraf mühit faktorları istifadəçilərin idrak proseslərinə və akademik məhsuldarlığına birbaşa təsir göstərir (Hoşten & Dalbay, 2018). Tədqiqatlar göstərir ki, bu parametrlərin optimal intervalda tənzimlənməsi akademik fəaliyyətin keyfiyyətini artırır (Bischo & Lahrz, 2008), əksinə, qeyri-adekvat mühit şəraiti isə diqqət dağınıqlığı və məhsuldarlıq itkisinə səbəb olur (Tabuenca et al., 2021).

Əşyaların İnterneti (IoT) texnologiyaları bu parametrlərin real vaxt rejimində izlənilməsini mümkün etmiş, maşın öyrənməsi (ML) əsaslı qərar dəstəyi sistemlərinin inkişafı üçün mühüm zəmin yaratmışdır (Umenga Weerasuriya et al., 2023). Bununla belə, mövcud tədqiqatların əksəriyyəti yalnız alqoritmlərin proqnozlaşdırma dəqiqliyinin müqayisəsinə fokuslanmış (Salamone et al., 2020), sensor əhəmiyyətinin sistemli interpretasiyası (Marzouk & Atef, 2022) və hesablama səmərəliliyi kimi kritik aspektlər isə kifayət qədər araşdırılmamışdır. Xüsusilə resurs məhdud IoT cihazlarında real vaxt rejimində işləyən sistemlər üçün yalnız dəqiqliyi deyil, eyni zamanda hesablama xərclərini nəzərə alan kompleks qiymətləndirmə yanaşmalarına ehtiyac getdikcə artmaqdadır.

Əvvəlki tədqiqatımızda Düzce Universiteti Kitabxanasında IoT sensorları vasitəsilə toplanan çoxölçülü ətraf mühit məlumatları əsasında yeddi ML alqoritmi müqayisə edilmiş və KNN modeli 99.04% F1-score ilə ən yüksək nəticəni nümayiş etdirmişdir (Mammadov & Kucukkulahli, 2025). Lakin həmin işdə sensor dəyişənlərinin model qərarlarına töhfəsi və hesablama resurs istifadəsi sistemli şəkildə analiz edilməmişdir. Hazırkı tədqiqat bu boşluğu aradan qaldırmaq məqsədilə Feature

Importance və Permutation Importance metodlarının paralel tətbiqi ilə sensor interpretasiyasını və hesablama səmərəliliyini kompleks şəkildə qiymətləndirməyi hədəfləyir.

Ətraf mühit amillərinin tədris və öyrənmə proseslərinə təsirini araşdıran çoxsaylı tədqiqatlar göstərir ki, səs-küy akademik performansla mənfi təsir göstərən əsas faktorlardan biridir. Yüksək səs səviyyəsi tələbələrin diqqətini yayındırır, informasiya emal prosesini çətinləşdirir və nəticə etibarilə ümumi akademik göstəricilərin zəifləməsinə səbəb olur (Khritish, 2023). Eyni zamanda, qapalı məkanlarda karbon dioksit (CO_2) konsentrasiyasının yüksəlməsi idrak funksiyalarının zəifləməsi, diqqət davamlılığının azalması və yorğunluğun artması ilə əlaqələndirilmişdir (Twardella et al., 2012).

İşıqlandırma səviyyəsi də tədris mühitinin keyfiyyətini müəyyən edən mühüm amillərdən biri kimi çıxış edir. Tədqiqatlar göstərir ki, xüsusilə təbii işıqlandırma tələbələrin diqqət səviyyəsinə, ümumi rifahına və akademik məhsuldarlığına müsbət təsir göstərir (Viola et al., 2008). Bundan əlavə, temperatur və rütubət kimi iqlim göstəricilərinin optimal hədlərdən kənara çıxması da öyrənmə səmərəliliyini əhəmiyyətli dərəcədə azaldır (Akanmu et al., 2021). Bununla belə, mövcud elmi ədəbiyyatın əksəriyyəti bu ətraf mühit faktorlarını ayrı-ayrılıqda nəzərdən keçirmiş, onların qarşılıqlı təsirini və birgə dinamikasını sistemli şəkildə analiz etməmişdir (Sadick et al., 2020). Bu isə ətraf mühit keyfiyyətinin kompleks və inteqrasiya olunmuş şəkildə qiymətləndirilməsi üçün vahid analitik çərçivənin zəruriliyini ortaya qoyur.

Son illərdə IoT (Internet of Things) texnologiyalarının kitabxana mühitlərində tətbiqi sahəsində aparılan tədqiqatların sayı nəzərəcarpacaq dərəcədə artmışdır. Mövcud araşdırmalar göstərir ki, temperatur, rütubət, karbon dioksit (CO_2) və səs səviyyəsi kimi ətraf mühit parametrlərinin IoT sensorları vasitəsilə real vaxt rejimində monitorinqi istifadəçi məhsuldarlığının artırılması baxımından effektiv yanaşma hesab olunur (Hong et al., 2022).

Kampus kitabxanaları üzərində aparılmış tədqiqatlarda yüksək temperatur, artmış CO_2 konsentrasiyası və qeyri-kafi işıqlandırma kimi ekoloji problemlər müəyyən edilmiş, bu çatışmazlıqların aradan qaldırılması üçün həm memarlıq, həm də texniki optimallaşdırma yanaşmaları təklif olunmuşdur (Umenga Weerasuriya et al., 2023). Bu yanaşmalar kitabxana mühitlərinin daha komfortlu və məhsuldar öyrənmə mühitinə çevrilməsində mühüm rol oynayır.

Digər tərəfdən, IoT texnologiyalarının kitabxana idarəetməsinə inteqrasiyası ilə bağlı tədqiqatlar göstərir ki, işçi heyətinin effektiv idarə olunması və motivasiya səviyyəsi bu sistemlərin uğurlu tətbiqi üçün kritik amillərdən biridir (Khan et al., 2022). Bundan əlavə, süni intellekt əsaslı hava keyfiyyəti qiymətləndirmə sistemləri mikrokontrollerlər və sensor şəbəkələri vasitəsilə real vaxt məlumatlarını emal edərək kitabxana idarəetməsinə dəstəkləyən qərarvermə mexanizmləri kimi təqdim edilmişdir (Marzouk & Atef, 2022).

Bununla yanaşı, fog və bulud hesablamalarının inteqrasiyası böyük həcmli IoT məlumatlarının emalı və saxlanması üçün genişlənə bilən (Maashi et al., 2024), eyni zamanda aşağı gecikmə təmin edən effektiv bir infrastruktur modeli kimi qiymətləndirilmişdir (Mahapatra et al., 2024).

Bununla belə, mövcud tədqiqatların əksəriyyəti əsasən ətraf mühit monitorinqinə fokuslanmış, toplanan sensor məlumatlarının hansı maşın öyrənməsi alqoritmləri vasitəsilə daha effektiv şəkildə təhlil olunduğu və hansı sensorların model qərarvermə prosesinə daha çox təsir etdiyi sisteməlik şəkildə araşdırılmamışdır. Bu boşluq növbəti bölmədə təqdim olunan alqoritmlərin müqayisəli təhlilinə əsas verir.

Ətraf mühit keyfiyyətinin təsnifatı və proqnozlaşdırılmasında maşın öyrənməsi alqoritmlərinin tətbiqi son illərdə sürətlə inkişaf edən tədqiqat istiqamətlərindən birinə çevrilmişdir. Bu sahədə aparılan araşdırmalar göstərir ki, müxtəlif alqoritmlərin performansları həm istifadə olunan verilənlər bazasının xüsusiyyətlərindən, həm də tətbiq edilən qiymətləndirmə meyarlarından əhəmiyyətli dərəcədə asılıdır.

Ətraf mühit keyfiyyətinin təsnifatına yönəlmiş tədqiqatlarda müxtəlif ML alqoritmləri sınaqdan keçirilmiş, ansambl əsaslı metodların çoxölçülü sensor məlumatlarında daha yüksək proqnozlaşdırma dəqiqliyi nümayiş etdirdiyi müəyyən edilmişdir. Eyni zamanda, mikrokontrollerlər vasitəsilə toplanan real vaxt sensor məlumatlarının süni intellekt modelləri ilə birləşdirilməsi

kitabxana və kampus mühitlərinin idarəedilməsində daha stabil qərar dəstəyi sistemlərinin qurulmasına imkan vermişdir (Marzouk & Atef, 2022).

Bununla belə, mövcud tədqiqatların əksəriyyətində alqoritmlərin müqayisəsi əsasən dəqiqlik metrikaları olan F1-score, precision və recall üzərindən aparılmış, lakin hesablama səmərəliliyi, o cümlədən öyrənmə müddəti, yaddaş istifadəsi və test mərhələsinin icra vaxtı kimi resurs əsaslı göstəricilər çox vaxt nəzərə alınmamışdır (Zareb et al., 2021). Halbuki IoT əsaslı real vaxt sistemlərində, xüsusilə Raspberry Pi kimi məhdud hesablama resurslarına malik cihazlarda tətbiq olunan həllərdə bu göstəricilər alqoritm seçimi baxımından kritik əhəmiyyət daşıyır.

Digər tərəfdən, sensor əhəmiyyəti analizi yəni hər bir sensorun model qərarvermə prosesinə verdiyi töhfənin kəmiyyətə qiymətləndirilməsi mövcud ədəbiyyatda nisbətən az araşdırılmış istiqamətlərdən biri olaraq qalmaqdadır (Shah et al., 2020). Bu boşluq, ətraf mühit keyfiyyətinin daha dərin və izaholunan şəkildə modelləşdirilməsi üçün əlavə tədqiqat ehtiyacını ortaya qoyur.

Mövcud ədəbiyyatın ümumi təhlili göstərir ki, həm proqnoz dəqiqliyini, həm hesablama səmərəliliyini, həm də sensor əhəmiyyətini vahid analitik çərçivədə birləşdirən kompleks müqayisəli tədqiqatlara ciddi ehtiyac vardır. Hazırkı tədqiqat məhz bu boşluğu aradan qaldırmaq məqsədilə həyata keçirilmişdir.

DATASET VƏ MƏLUMATLARIN HAZIRLANMASI

Məlumatın Toplanması. Tədqiqatda istifadə edilən məlumatlar Düzce Universiteti Kitabxanasının əsas oxu zalında toplanmışdır (Mammadov & Kucukkulahli, 2025). Səs səviyyəsi, işıq intensivliyi, temperatur, rütubət, CO₂, eCO₂, TVOC və anlıq istifadəçi sayı kimi 8 ətraf mühit parametri IoT sensorları vasitəsilə real vaxt rejimində qeydə alınmış, istifadəçi rəyləri isə QR kod əsaslı sorğu vasitəsilə toplanaraq sensor məlumatları ilə zaman damğasına görə sinxronlaşdırılmışdır (Mammadov & Kucukkulahli, 2025).

Məlumat toplusu 8 sensor dəyişənini, istifadəçi rəylərinin kodlaşdırılmış formalarını və hər qeydin aid olduğu zaman damğasını əhatə edir (Mammadov & Kucukkulahli, 2025). Ümumi ətraf mühit keyfiyyəti qiymətləndirməsini əks etdirən hədəf dəyişəni 5 sinifdən ibarətdir: çox pis, pis, orta, yaxşı və çox yaxşı. Sinif balansını təmin etmək məqsədilə SMOTE metodu tətbiq edilmiş və nəticədə 9404 sətirdən ibarət balanslaşdırılmış dataset əldə edilmişdir (Mammadov & Kucukkulahli, 2025). Məlumat toplusunun strukturu Cədvəl 1-də təqdim edilmişdir. Cədvəldən göründüyü kimi, sensor göstəriciləri ilə istifadəçi qiymətləndirmələri paralel şəkildə qeydə alınmış, hər sətir müəyyən bir zaman anında oxu zalındakı ətraf mühit vəziyyətini əks etdirir.

Cədvəl 1. Məlumat toplusundan nümunə qeydlər (N=9).

Səs (dB)	Səs İstif.	İşıq (lux)	İşıq İstif.	Temp. (°C)	Rütubət (%)	Temp. İstif.	CO ₂ (ppm)	eCO ₂ (ppm)	TVOC (ppb)	Hava İstif.	Dolul uq (nəf.)	Dolul uq İstif.	Ümumi Qiymət
130.49	4	2677	4	25.6	40.85	4	293	670	131	2	43	4	4
126.15	3	2679	4	25.76	40.78	3	286	652	126	4	42	3	4
117.56	4	2678	4	25.94	40.48	4	313	652	129	3	40	3	3
103.72	4	2690	4	25.8	41.15	4	272	642	110	2	50	4	4
105.57	4	2706	5	25.87	41.05	4	325	583	90	3	50	3	4

Səs (dB)	Sə s İst if.	İşı q (lux)	İşı q İst if.	Te mp. (°C)	Rütu bət (%)	Te mp. İstif .	CO ₂ (ppm)	eC O ₂ (ppm)	TV OC (ppb)	Ha va İstif.	Dolul uq (nəf.)	Dolul uq İstif.	Ümu mi Qiy mət
97.1	4	27 42	3	25.6 6	40.84	4	246	515	56	2	48	3	4
98.2 4	5	26 91	4	25.5 6	40.83	5	209	478	39	4	48	4	4
94.1 1	5	26 95	4	25.2 9	41.17	4	258	413	0	4	45	4	4
53.8 7	4	27 66	4	25.7 7	39.52	4	223	400	0	4	48	4	4

Məlumatların ön emalı və balanslaşdırılması. Xam məlumatlar modelləşdirməyə hazırlanmadan əvvəl bir sıra ön emal mərhələlərindən keçirilmişdir (Mammadov & Kucukkulahli, 2025). Kateqorial istifadəçi rəyləri — "çox yetərli", "yetərli", "qərarısam", "yetərsiz" və "heç yetərli deyil" - Label Encoding metodu ilə uyğun olaraq 5, 4, 3, 2, 1 rəqəmsal dəyərlərinə çevrilmişdir. Məlumat toplusunda sinif balansızlığı aşkar edildiyi üçün az təmsil olunan siniflərin nümunə sayını artırmaq məqsədilə SMOTE (Synthetic Minority Over-sampling Technique) metodu tətbiq edilmişdir (Elreedy et al., 2023). Bu mərhələdən sonra sensor məlumatları ilə istifadəçi rəyləri zaman damğasına əsasən sinxronlaşdırılmış və model üçün vahid strukturlu məlumat toplusu formalaşdırılmışdır.

TƏDQIQAT METODOLOGİYASI

Bu tədqiqatda yeddi müxtəlif maşın öyrənməsi alqoritmi tətbiq edilmişdir: KNN, Random Forest, Decision Tree, SVM, XGBoost, Logistic Regression və Naive Bayes. Hər bir alqoritm eyni məlumat toplusu üzərində sınaqdan keçirilmiş, hiperparametrlərin optimallaşdırılmasında GridSearchCV metodu, modellərin qiymətləndirilməsində isə 10 qatlı çarpaz yoxlama (cross-validation) tətbiq edilmişdir (Mammadov & Kucukkulahli, 2025).

Bu tədqiqatda sensor əhəmiyyətinin müəyyən edilməsi üçün iki müxtəlif metod paralel şəkildə tətbiq edilmişdir. Birinci metod olan Random Forest Feature Importance, hər bir sensorun model qərarlarına töhfəsini alqoritmin daxili Gini indeksinə əsasən hesablayır. İkinci metod olan Permutation Importance isə hər bir sensor dəyişəninə dəyərlərini ardıcıl olaraq qarışdıraraq modelin F1-score göstəricisindəki azalmanı ölçür, bu azalma nə qədər böyükdürsə, həmin sensor bir o qədər kritik hesab edilir. İki metodun paralel tətbiqi sensor əhəmiyyətinin daha etibarlı və hərtərəfli şəkildə qiymətləndirilməsinə imkan yaradır.

Bu tədqiqatda alqoritmlərin müqayisəsi iki əsas istiqamət üzrə həyata keçirilmişdir. Birinci istiqamətdə təsnifat performansını F1-score, precision və recall metrikaları vasitəsilə qiymətləndirilmişdir. İkinci istiqamətdə isə hesablama səmərəliliyi öyrənmə vaxtı, sınaq vaxtı və yaddaş istifadəsi ölçülmüşdür. Bu iki istiqamətin birlikdə qiymətləndirilməsi IoT əsaslı real vaxt sistemləri üçün ən optimal alqoritmin müəyyən edilməsinə imkan yaradır.

EKSPERİMENTAL NƏTİCƏLƏR

Tədqiqat çərçivəsində tətbiq edilən yeddi müxtəlif maşın öyrənməsi alqoritminin performansını F1-score, precision və recall metrikaları əsasında müqayisəli şəkildə qiymətləndirilmişdir. Əldə olunan nəticələr göstərir ki, modellərin performansını arasında nəzərəcərpacaq fərqlər mövcuddur və bu fərqliliklər əsasən alqoritmlərin struktur xüsusiyyətləri və verilənlərdəki qeyri-xətti münasibətlərlə izah olunur.

Cədvəl 2. Maşın öyrənməsi modellərinin performans müqayisəsi.

Alqoritm	F1 Score ↑	Precision ↑	Recall ↑
KNN	0.9904	0.9906	0.9901
SVM	0.945	0.951	0.941
Random Forest	0.9902	0.9905	0.9906
Decision Tree	0.984	0.984	0.984
XGBoost	0.741	0.750	0.743
Naive Bayes	0.429	0.434	0.450
Logistic Regression	0.493	0.491	0.515

Yeddi maşın öyrənməsi modelinin F1-score, precision və recall metrikaları üzrə müqayisəli performans nəticələri Cədvəl 2-də təqdim edilmişdir. Cədvəldən göründüyü kimi, modellər arasında nəzərəcarpacaq performans fərqləri mövcuddur. Bu fərqliliklər əsasən alqoritmlərin struktur xüsusiyyətləri və verilənlərdəki qeyri-xətti münasibətlərlə izah olunur.

Ən yüksək ümumi performans KNN alqoritmı tərəfindən nümayiş etdirilmişdir. Bu model 0.9904 F1-score, 0.9906 precision və 0.9901 recall göstəriciləri ilə bütün modellər arasında ən balanslı və yüksək nəticəni əldə etmişdir. KNN-in üstün performansı verilənlərdə nümunələr arasında güclü lokal oxşarlıq strukturunun mövcudluğunu təsdiqləyir və məsafə əsaslı metodların bu tip çoxölçülü sensor məlumatlarında effektiv işlədiyini göstərir.

İkinci ən yüksək nəticə Random Forest alqoritmı tərəfindən əldə olunmuşdur (F1-score = 0.9902). Random Forest-un yüksək performansı onun ansambl əsaslı strukturu ilə əlaqədardır; bu yanaşma çoxsaylı qərar ağaclarının nəticələrini birləşdirərək həm variasiyanı azaldır, həm də overfitting riskini minimallaşdırır.

Decision Tree modeli 0.984 F1-score ilə yüksək, lakin KNN və Random Forest ilə müqayisədə bir qədər aşağı performans nümayiş etdirmişdir. Bu fərq tək ağac strukturunun stabil baxımından məhdudiyyətləri ilə izah olunur.

SVM modeli 0.945 F1-score ilə orta-yüksək səviyyəli nəticə göstərmişdir. XGBoost isə 0.741 F1-score ilə orta səviyyəli performans nümayiş etdirmiş, gradient boosting yanaşmasının bu konkret dataset üçün tam potensialını reallaşdırmadığını göstərmişdir.

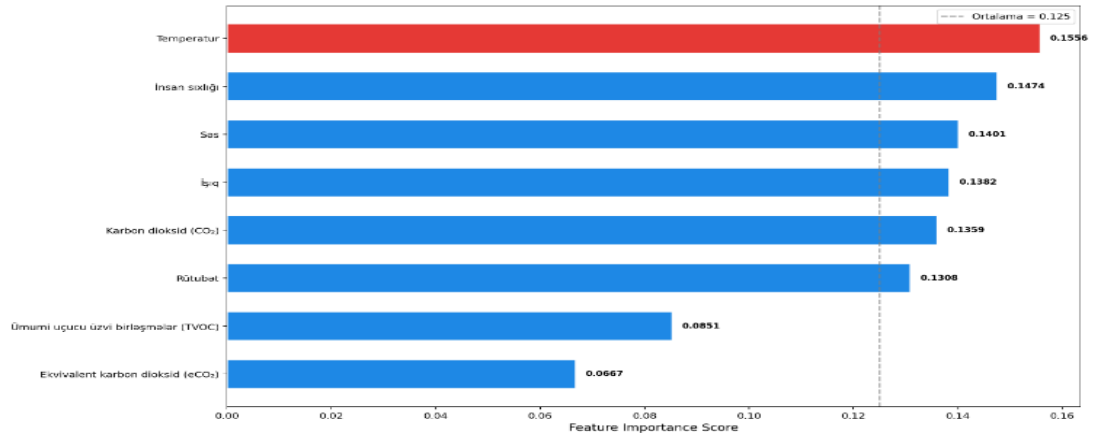
Bunun əksinə olaraq, Logistic Regression (0.493), Naive Bayes (0.429) modelləri nisbətən zəif performans nümayiş etdirmişdir. Bu nəticə verilənlərdə xətti ayrılmanın zəif olduğunu və dəyişənlər arasında mürəkkəb qeyri-xətti münasibətlərin mövcudluğunu göstərir. Xüsusilə Naive Bayes modelinin aşağı nəticəsi atributlar arasında müstəqillik fərziyyəsinin bu dataset üçün keçərli olmadığını təsdiqləyir.

Ümumilikdə əldə olunan nəticələr göstərir ki, məsafə əsaslı və ansambl əsaslı modellər bu tip IoT əsaslı çoxölçülü ətraf mühit təsnifatı problemlərində daha yüksək performans nümayiş etdirir.

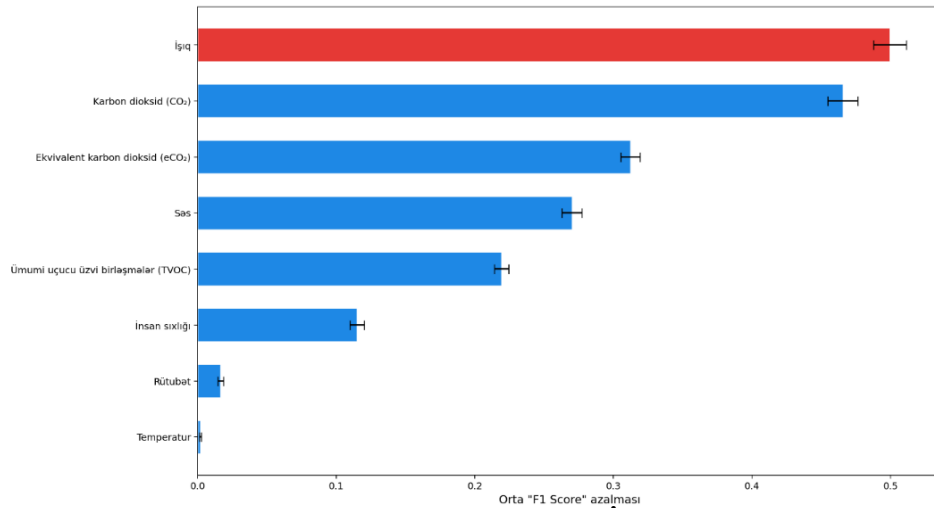
Sensor əhəmiyyəti analizi çərçivəsində Random Forest və K-Nearest Neighbors (KNN) modelləri üçün müvafiq olaraq Gini əsaslı Feature Importance və Permutation Importance metodları tətbiq edilmişdir. Bu yanaşma sensor dəyişənlərinin model qərarlarına təsirini həm daxili model strukturu, həm də performans həssaslığı baxımından qiymətləndirməyə imkan verir.

Random Forest modelinin daxili Feature Importance sıralaması Şəkil 1-də təqdim edilmişdir. Şəkildən göründüyü kimi, sensor dəyişənləri arasında əhəmiyyət baxımından fərqlər mövcuddur və bu fərqlər modelin qərar vermə mexanizmini anlamaq üçün mühüm əsas yaradır. Random Forest modelinə əsaslanan Feature Importance nəticələri göstərir ki, ən yüksək təsirə malik dəyişən temperatur (0.1556) olmuşdur. Bu nəticə modelin qərar mexanizminin əsasən istilik dəyişikliklərinə həssas olduğunu göstərir. Daha sonra insan sıxlığı (0.1474), səs səviyyəsi (0.1401) və işıq intensivliyi (0.1382) yüksək əhəmiyyət göstəricilərinə malik olmuşdur. Bu dəyişənlər məkanın həm fiziki, həm

də sosial dinamikasını əks etdirərək modelin qərar prosesində mühüm rol oynayır. Digər tərəfdən, TVOC (0.0851) və eCO₂ (0.0667) dəyişənləri nisbətən aşağı təsir göstərmişdir.



Şəkil 1. "Random Forest" modeli üzrə xüsusiyyətlərin əhəmiyyəti dərəcəsi (Feature Importance) Sensor göstəricilərinin mühit keyfiyyətinin təyininə təsiri.



Şəkil 2. "KNN" modeli üçün "Permutation Feature Importance" təhlili (Daha yüksək göstərici müvafiq sensorun yüksək əhəmiyyətini ifadə edir).

KNN modeli üçün Permutation Importance analizi nəticələri Şəkil 2-də göstərilmişdir. Şəkildən aydın olur ki, sensor dəyişənlərinin model performansına təsiri Feature Importance metodundan əhəmiyyətli dərəcədə fərqlənir. KNN modeli üçün aparılan Permutation Importance analizi fərqli nəticə strukturu ortaya qoymuşdur. Bu modeldə işıq intensivliyi (~0.50) ən kritik dəyişən kimi müəyyən edilmiş və onun pozulması model performansında ciddi azalmaya səbəb olmuşdur. CO₂ (~0.46) və eCO₂ (~0.31) göstəriciləri də yüksək əhəmiyyət nümayiş etdirmişdir. Bunun əksinə olaraq, temperatur və rütubət dəyişənləri model performansına minimal təsir göstərmişdir. Bu nəticə göstərir ki, KNN modeli daha çox məsafə əsaslı oxşarlıq strukturlarına və vizual/hava keyfiyyəti dəyişənlərinə həssasdır.

Cədvəl 3. Random Forest və Permutation Importance sıralamalarının müqayisəsi.

Sensor	RF Ranking	Permutation Ranking
Temperatur	1	8
İnsan sıxlığı	2	6

Sensor	RF Ranking	Permutation Ranking
Səs	3	4
İşıq	4	1
Karbon dioksidi (CO ₂)	5	2
Rütubət	6	7
Ümumi uçucu üzvi birləşmələr (TVOC)	7	5
Ekvivalent karbon dioksidi (eCO ₂)	8	3

Cədvəl 3, Random Forest modelinin daxili Feature Importance sıralaması ilə Permutation Importance metodundan əldə edilən nəticələrin müqayisəsini təqdim edir. Bu müqayisə sensor dəyişənlərinin model daxilində necə qiymətləndirildiyini və real performans təsiri ilə nə dərəcədə uyğunlaşdığını göstərmək məqsədi daşıyır.

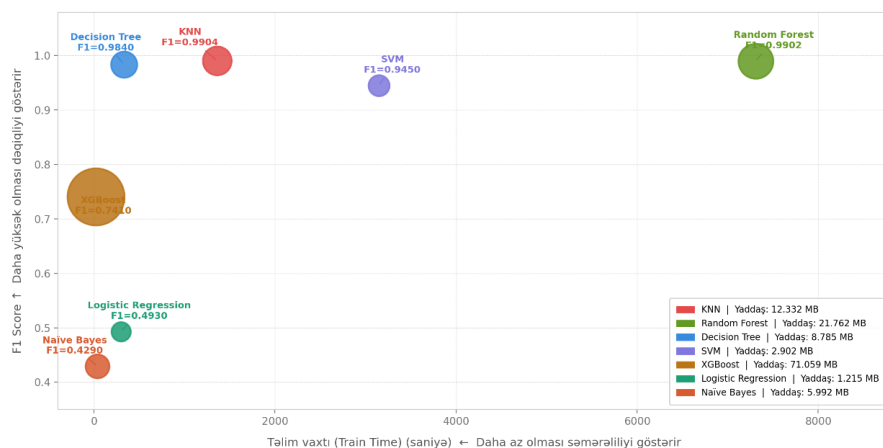
Əldə olunan nəticələr göstərir ki, Random Forest modelinə görə ən vacib dəyişən temperatur (1-ci yer) olmuşdur, halbuki Permutation Importance nəticələrində bu dəyişən ən aşağı sıralarda (8-ci yer) yerləşmişdir. Bu, modelin daxili qərar mexanizmi ilə real performans təsiri arasında ciddi fərq olduğunu göstərir.

Digər tərəfdən, işıq intensivliyi dəyişəni Permutation Importance analizində 1-ci sıraya yüksələrək model performansına ən böyük təsir göstərən dəyişən olduğunu sübut etmişdir. Eyni zamanda eCO₂ və CO₂ dəyişənləri də Permutation nəticələrində daha yüksək mövqelərə yüksəlmişdir ki, bu da onların model üçün kritik rol oynadığını göstərir.

Bəzi dəyişənlər isə hər iki metodda daha stabil nəticələr göstərmişdir. Xüsusilə səs və rütubət dəyişənləri həm RF, həm də Permutation sıralamasında orta mövqelərdə yerləşərək nisbi sabitlik nümayiş etdirmişdir.

Ümumilikdə nəticələr göstərir ki, Random Forest-in daxili feature importance yanaşması bəzi hallarda dəyişənlərin real təsir gücünü tam əks etdirməyə bilər. Permutation Importance isə model performansına birbaşa təsirə əsaslandığı üçün daha etibarlı interpretasiya imkanı təqdim edir. Buna görə də sensor əhəmiyyəti analizi aparılarkən hər iki metodun birlikdə istifadə edilməsi daha düzgün nəticələr verir.

Bu bölmədə müxtəlif maşın öyrənməsi alqoritmlərinin performansı yalnız proqnozlaşdırma dəqiqliyi baxımından deyil, həm də hesablama səmərəliliyi təlim müddəti, test müddəti və yaddaş istifadəsi baxımından müqayisə edilmişdir. Bu analiz IoT əsaslı real vaxt sistemlərində model seçiminin optimallaşdırılması üçün kritik əhəmiyyət daşıyır.



Şəkil 3. Alqoritmlərin hesablama səmərəliliyinin və performansının müqayisəsi (Dairənin ölçüsü = Yaddaş istifadəsi)

Şəkil 3-də göstərilən bubble chart nəticələri alqoritmlərin F1-score, təlim müddəti və yaddaş istifadəsi üzrə üçölçülü müqayisəsini təqdim edir. Nəticələr göstərir ki, KNN modeli yüksək F1-score (0.9904) əldə etməklə yanaşı, 1360.9 saniyəlik təlim müddəti və 12.332 MB yaddaş istifadəsi ilə yüksək performanslı modeldir. Decision Tree modeli isə 330.47 saniyə təlim müddəti, 8.785 MB yaddaş istifadəsi və 0.984 F1-score ilə həm sürət, həm də dəqiqlik baxımından ən balanslı nəticəni nümayiş etdirir.

Bunun əksinə olaraq, Random Forest modeli ən yüksək dəqiqliyə (0.9902) malik olsa da, 7309.2 saniyəlik təlim müddəti və 21.762 MB yaddaş istifadəsi ilə ən ağır hesablama yükünü yaradan modeldir. XGBoost modeli 71.059 MB yaddaş istifadəsi ilə yaddaş baxımından ən bahalı model olmasına baxmayaraq, cəmi 0.741 F1-score göstəricisi ilə orta səviyyəli performans nümayiş etdirmişdir. SVM (3146.9 saniyə təlim müddəti, 0.945 F1-score) və Logistic Regression (300.32 saniyə, 0.493 F1-score) modelləri isə resurs xərci ilə performans balansı baxımından nisbətən əlverişsiz nəticələr göstərmişdir.

Cədvəl 4. Alqoritmlərin hesablama səmərəliliyinin (vaxt və yaddaş) müqayisəsi.

Model	Təlim Sürəsi (Saniyə) ↓	Test Sürəsi (Saniyə) ↓	Yaddaş İstifadəsi (MB) ↓
KNN	1360.9	0.071	12.332
SVM	3146.9	6.510	2.902
Random Forest	7309.2	0.422	21.762
Decision Tree	330.470	0.005	8.785
XGBoost	20.602	0.152	71.059
Naïve Bayes	36.160	0.004	5.992
Logistic Regression	300.32	0.155	1.215

Cədvəl 4, alqoritmlərin təlim və test müddəti ilə yanaşı yaddaş istifadəsini daha detallı şəkildə göstərir. Nəticələrə əsasən Naive Bayes modeli ən qısa təlim müddətinə (0.009 s) malik olmaqla ən sürətli modeldir. Decision Tree və Logistic Regression modelləri isə test mərhələsində çox aşağı gecikmə ilə işləyir.

Yaddaş istifadəsi baxımından KNN modeli ən qənaətcil yanaşma təqdim edir (0.387 MB), halbuki Random Forest və XGBoost modelləri daha yüksək yaddaş tələbi ilə seçilir. SVM modeli isə xüsusilə test mərhələsində daha yüksək gecikmə göstərərək real vaxt tətbiqləri üçün nisbətən əlverişsiz nəticə ortaya qoyur. Ümumi nəticələr göstərir ki, alqoritmlər arasında ciddi “accuracy–efficiency trade-off” mövcuddur. Random Forest maksimum dəqiqlik təqdim etsə də, resurs istifadəsi baxımından daha bahalıdır. KNN və Decision Tree isə yüksək performans minimal hesablama xərci ilə birləşdirərək IoT əsaslı sistemlər üçün daha praktik seçimlər kimi çıxış edir.

Bu nəticələr göstərir ki, real vaxt IoT tətbiqlərində model seçimi yalnız dəqiqlik göstəricisinə deyil, eyni zamanda hesablama resurslarına uyğunluğa əsaslanmalıdır.

MÜZAKİRƏ VƏ TƏHLİL

Bu bölmədə əldə olunan eksperimental nəticələr model performans, sensor dəyişənlərinin interpretasiyası və hesablama səmərəliliyi aspektlərindən kompleks şəkildə təhlil edilir və onların IoT əsaslı real vaxt sistemləri kontekstində tətbiq imkanları müzakirə olunur.

Əldə olunan nəticələr göstərir ki, KNN modeli 0.9904 F1-score ilə ümumi proqnozlaşdırma performans baxımından ən yüksək nəticəni əldə edərək ən güclü təsnifatçı kimi çıxış edir. Random

Forest modeli isə 0.9902 F1-score ilə ikinci sırada yerləşmiş, ansambl əsaslı strukturunun gücünü təsdiqləmişdir. Bununla yanaşı, Random Forest-in yüksək təlim vaxtı (7309.2 s) və yaddaş tələbi (21.762 MB) onun resurs məhdud mühitlərdə tətbiqini müəyyən dərəcədə məhdudlaşdırır bilər. Decision Tree modeli isə həm hesablama səmərəliliyi (330.47 s, 8.785 MB), həm də nisbətən yüksək və sabit performans (0.984 F1-score) ilə ən balanslı alternativ kimi çıxış edir.

Sensor əhəmiyyəti analizləri göstərir ki, modellərin daxili Feature Importance mexanizmləri ilə Permutation əsaslı həssaslıq analizləri arasında nəzərəcarpacaq fərqlər mövcuddur. Random Forest modeli temperatur və insan sıxlığı kimi dəyişənləri daha dominant faktorlar kimi qiymətləndirərkən, Permutation Importance nəticələri işıq intensivliyi və qaz konsentrasiyası (CO₂, eCO₂) dəyişənlərinin model performansına daha yüksək təsir göstərdiyini ortaya qoymuşdur. Bu uyğunsuzluq göstərir ki, yalnız model daxili əhəmiyyət ölçülərinə əsaslanmaq sensor təsirinin tam və etibarlı interpretasiyası üçün kifayət deyildir və daha robust analiz yanaşmalarına ehtiyac vardır.

Hesablama səmərəliliyi nəticələri də IoT əsaslı real vaxt tətbiqləri üçün mühüm nəticələr ortaya qoyur. XGBoost ən qısa təlim vaxtına (20.602 s) malik olsa da, 71.059 MB yaddaş istifadəsi və cəmi 0.741 F1-score ilə performans-resurs balansı baxımından optimal seçim deyildir. Decision Tree isə həm sürətli təlim vaxtı, həm aşağı yaddaş istehlakı, həm də yüksək dəqiqliyi ilə real vaxt IoT sistemləri üçün ən praktik model kimi seçilir. SVM və Naïve Bayes modelləri isə bu verilənlər toplusu üzərində həm performans, həm də səmərəlilik baxımından nisbətən zəif nəticələr göstərmişdir.

Ümumilikdə nəticələr göstərir ki, IoT əsaslı ətraf mühit monitoring sistemlərində model seçimi zamanı yalnız proqnozlaşdırma dəqiqliyi deyil, eyni zamanda hesablama səmərəliliyi və interpretasiya ediləbilənlik də nəzərə alınmalıdır. Bu kontekstdə "accuracy–efficiency trade-off" problemi sistem dizaynının əsas determinantlarından biri kimi çıxış edir. Bu tədqiqatın əsas elmi töhfəsi ondan ibarətdir ki, o, yalnız klassifikasiya performansını deyil, eyni zamanda sensor əhəmiyyəti və hesablama səmərəliliyini birlikdə qiymətləndirərək daha kompleks və praktik bir analiz çərçivəsi təqdim edir. Bu yanaşma, xüsusilə ağıllı kampus və IoT əsaslı ətraf mühit optimallaşdırma sistemlərində daha düzgün və dayanıqlı model seçimi üçün mühüm nəzəri və tətbiqi baza formalaşdırır.

NƏTİCƏ

Bu tədqiqat IoT əsaslı ağıllı kitabxana mühitlərində ətraf mühit keyfiyyətinin qiymətləndirilməsi üçün maşın öyrənməsi yanaşmalarını daha dərin və interpretasiya edilə bilən şəkildə analiz etməyi hədəfləmişdir. Aparılan eksperimentlər göstərir ki, KNN modeli 0.9904 F1-score ilə ən yüksək proqnozlaşdırma performansını nümayiş etdirmişdir. Random Forest modeli 0.9902 F1-score ilə ikinci sırada yerləşmiş, ansambl əsaslı strukturunun effektivliyini təsdiqləmişdir. Bununla yanaşı, Decision Tree modeli həm aşağı hesablama xərci (330.47 s, 8.785 MB), həm də yüksək dəqiqliyi (0.984 F1-score) ilə resurs məhdud IoT mühitləri üçün ən balanslı alternativ kimi çıxış edir.

Sensor əhəmiyyəti analizləri nəticəsində müəyyən edilmişdir ki, müxtəlif modellər sensor dəyişənlərinə fərqli prioritetlər verir. Random Forest modeli əsasən temperatur və insan sıxlığı kimi dəyişənləri daha dominant faktorlar kimi qiymətləndirərkən, Permutation Importance analizləri işıq intensivliyi və qaz konsentrasiyası (CO₂, eCO₂) kimi dəyişənlərin model performansına daha kritik təsir göstərdiyini ortaya qoymuşdur. Bu nəticələr sensor interpretasiyasının yalnız model daxili ölçülərlə deyil, performans əsaslı yanaşmalarla da qiymətləndirilməsinin vacibliyini göstərir.

Hesablama səmərəliliyi baxımından əldə olunan nəticələr göstərir ki, IoT əsaslı real vaxt sistemlərində model seçimi yalnız dəqiqlik əsasında aparılmamalıdır. Random Forest yüksək dəqiqlik təmin etsə də 7309.2 saniyəlik təlim vaxtı və 21.762 MB yaddaş istifadəsi ilə ən ağır hesablama yükünü yaradır. XGBoost ən qısa təlim vaxtına (20.602 s) malik olsa da, 71.059 MB yaddaş istifadəsi və aşağı F1-score (0.741) ilə optimal seçim deyildir. Decision Tree modeli isə həm performans, həm də səmərəlilik baxımından ən balanslı yanaşmanı təqdim edir.

Qeyd etmək lazımdır ki, bu tədqiqat əvvəlki işin davamı və genişləndirilmiş analitik mərhələsi kimi nəzərdən keçirilməlidir (Mammadov & Kucukkulahli, 2025). Əvvəlki tədqiqatda eyni dataset üzərində müxtəlif maşın öyrənməsi alqoritmləri tətbiq edilərək yüksək təsnifat dəqiqliyi əldə

olunmuş, lakin əsas diqqət model performansına yönəlmişdir. Sensor dəyişənlərinin təsirinin interpretasiyası və hesablama səmərəliliyi kimi mühüm aspektlər həmin işdə sistemli şəkildə araşdırılmamışdır.

Hazırkı tədqiqat bu boşluqları aradan qaldıraraq yalnız proqnozlaşdırma nəticələrini deyil, eyni zamanda sensor əhəmiyyəti və modelin hesablama davranışını da kompleks şəkildə analiz etmişdir. Bu yanaşma nəticələrin daha interpretasiya edilə bilən, müqayisə edilə bilən və real vaxt IoT sistemlərinə daha uyğun olmasını təmin edir. Ümumilikdə əldə olunan nəticələr göstərir ki, IoT əsaslı ətraf mühit monitorinq sistemlərində model seçimi zamanı "accuracy–efficiency trade-off" və interpretasiya qabiliyyəti birlikdə nəzərə alınmalıdır. Təklif olunan yanaşma gələcək ağıllı kampus və oxşar IoT tətbiqlərində daha optimallaşdırılmış və dayanıqlı qərar dəstəyi sistemlərinin inkişafına töhfə verə bilər.

ƏDƏBİYYATLAR

- Akanmu, W. P., Nunayon, S. S., & Eboson, U. C. (2021). Indoor environmental quality (IEQ) assessment of Nigerian university libraries: A pilot study. *Energy and Built Environment*, 2(3), 302–314. <https://doi.org/10.1016/J.ENBENV.2020.07.004>
- Bischo, W., & Lahrz, T. (2008). Gesundheitliche Bewertung von Kohlendioxid in der Innenraumluft. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2008 51:11, 51(11), 1358–1369. <https://doi.org/10.1007/S00103-008-0707-2>
- Elreedy, D., Amir, ., Atiya, F., Kamalov, . Firuz, Moniz, N., Branco, P., Torgo, L., Japkowicz, N., Wozniak, M., Wang, S., Atiya, A. F., & Kamalov, F. (2023). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning* 2023 113:7, 113(7), 4903–4923. <https://doi.org/10.1007/S10994-022-06296-4>
- Hong, S., Kim, Y., & Yang, E. (2022). Indoor environment and student productivity for individual and collaborative work in learning commons: a case study. *Library Management*, 43(1–2), 15–34. <https://doi.org/10.1108/LM-06-2021-0055>
- Hoşten, G., & Dalbay, N. (2018). Kapalı Ortam Hava Kalitesinin Ofis Ortamlarında İş Sağlığı ve Güvenliği Açısından Değerlendirilmesi. *Aydın Journal of Health*, 4(2), 1–12. <https://dergipark.org.tr/en/pub/asder/article/530250>
- Khan, A. U., Zhang, Z., Chohan, S. R., & Rafique, W. (2022). Factors fostering the success of IoT services in academic libraries: a study built to enhance the library performance. *Library Hi Tech*, 40(6), 1976–1995. <https://doi.org/10.1108/LHT-06-2021-0179>
- Khritish, S. (2023). The Impact of Study Environment on Students' Academic Performance: An Experimental Research Study. <https://doi.org/https://doi.org/10.36227/techrxiv.170250911.12566888/v1>
- Maashi, M., Alabdulkreem, E., Maray, M., Shankar, K., Darem, A. A., Alzahrani, A., & Yaseen, I. (2024). Elevating Survivability in Next-Gen IoT-Fog-Cloud Networks: Scheduling Optimization With the Metaheuristic Mountain Gazelle Algorithm. *IEEE Transactions on Consumer Electronics*, 70(1), 3802–3809. <https://doi.org/10.1109/TCE.2024.3371774>
- Mahapatra, A., Majhi, S. K., Mishra, K., Pradhan, R., Rao, D. C., & Panda, S. K. (2024). An Energy-Aware Task Offloading and Load Balancing for Latency-Sensitive IoT Applications in the Fog-Cloud Continuum. *IEEE Access*, 12, 14334–14349. <https://doi.org/10.1109/ACCESS.2024.3357122>
- Mammadov, S., & Kucukkulahli, E. (2025). A User-Centric Smart Library System: IoT-Driven Environmental Monitoring and ML-Based Optimization with Future Fog-Cloud Architecture. *Applied Sciences* 2025, Vol. 15, Page 3792, 15(7), 3792. <https://doi.org/10.3390/APP15073792>
- Marzouk, M., & Atef, M. (2022a). Assessment of Indoor Air Quality in Academic Buildings Using IoT and Deep Learning. *Sustainability* 2022, Vol. 14, Page 7015, 14(12), 7015. <https://doi.org/10.3390/SU14127015>

- Marzouk, M., & Atef, M. (2022b). Assessment of Indoor Air Quality in Academic Buildings Using IoT and Deep Learning. *Sustainability* 2022, Vol. 14, Page 7015, 14(12), 7015. <https://doi.org/10.3390/SU14127015>
- Sadick, A. M., Kpamma, Z. E., & Agyefi-Mensah, S. (2020). Impact of indoor environmental quality on job satisfaction and self-reported productivity of university employees in a tropical African climate. *Building and Environment*, 181, 107102. <https://doi.org/10.1016/J.BUILDENV.2020.107102>
- Salamone, F., Bellazzi, A., Belussi, L., Damato, G., Danza, L., Dell'aquila, F., Ghellere, M., Megale, V., Meroni, I., & Vitaletti, W. (2020). Evaluation of the Visual Stimuli on Personal Thermal Comfort Perception in Real and Virtual Environments Using Machine Learning Approaches. *Sensors* 2020, Vol. 20, Page 1627, 20(6), 1627. <https://doi.org/10.3390/S20061627>
- Shah, S. K., Tariq, Z., Lee, J., & Lee, Y. (2020). Real-Time Machine Learning for Air Quality and Environmental Noise Detection. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data* 2020, 3506–3515. <https://doi.org/10.1109/BIGDATA50022.2020.9377939>
- Tabuenca, B., Borner, D., & Kalz, M. (2021). Effects of an Ambient Learning Display on Noise Levels and Perceived Learning in a Secondary School. *IEEE Transactions on Learning Technologies*, 14(1), 69–80. <https://doi.org/10.1109/TLT.2021.3056038>
- Twardella, D., Matzen, W., Lahrz, T., Burghardt, R., Spiegel, H., Hendrowarsito, L., Frenzel, A. C., & Fromme, H. (2012). Effect of classroom air quality on students' concentration: Results of a cluster-randomized cross-over experimental study. *Indoor Air*, 22(5), 378–387. <https://doi.org/10.1111/J.1600-0668.2012.00774.X>
- Umenga Weerasuriya, A., Wai Tsang, K., Li, Y., Lee, C.-C., Aflaki, A., Esfandiari, M., & Jarrahi, A. (2023). Multi-Criteria Evaluation of a Library's Indoor Environmental Quality in the Tropics. *Buildings* 2023, Vol. 13, Page 1233, 13(5), 1233. <https://doi.org/10.3390/BUILDINGS13051233>
- Viola, A. U., James, L. M., Schlangen, L. J. M., & Dijk, D. J. (2008). Blue-enriched white light in the workplace improves self-reported alertness, performance and sleep quality. *Scandinavian Journal of Work, Environment and Health*, 34(4), 297–306. <https://doi.org/10.5271/SJWEH.1268>
- Zareb, M., Bakhti, B., Bouzid, Y., Batista, C. E., Ternifi, I., & Abdenour, M. (2021). An intelligent IoT fuzzy based approach for automated indoor air quality monitoring. *2021 29th Mediterranean Conference on Control and Automation, MED 2021*, 770–775. <https://doi.org/10.1109/MED51440.2021.9480313>

SUMMARY

INTERPRETATION OF SENSOR IMPORTANCE AND COMPARATIVE ANALYSIS OF COMPUTATIONAL EFFICIENCY OF MACHINE LEARNING MODELS IN IOT-BASED SMART LIBRARY ENVIRONMENTS

Sarkan Mammadov

Nakhchivan State University, Nakhchivan, Azerbaijan

This study aims to interpret the impact of environmental sensors on the decision-making processes of machine learning models in IoT-based smart library environments and to evaluate computational efficiency. Based on multidimensional sensor data collected from Düzce University Library, seven machine learning algorithms — KNN, Random Forest, Decision Tree, SVM, XGBoost, Logistic Regression, and Naïve Bayes — were comparatively applied. The primary scientific contribution of this study is the parallel implementation of two sensor importance assessment approaches, namely Random Forest Feature Importance and Permutation Importance. This dual-method framework enabled a more precise analysis of the relationship between model-

internal decision mechanisms and the actual impact of sensor variables on predictive performance. The results indicate that while the Feature Importance approach identifies temperature and human occupancy as dominant factors, Permutation Importance analysis reveals that light intensity and CO₂-based variables exert a greater influence on model performance. This discrepancy demonstrates that relying solely on model-internal importance measures is insufficient for accurate sensor interpretation, highlighting the need for more robust evaluation methodologies. Furthermore, in terms of computational efficiency, KNN and Decision Tree models exhibited more suitable characteristics for real-time IoT applications, owing to their low resource consumption and high execution speed. Overall, the findings suggest that optimal model selection in IoT-based systems should be guided not only by predictive accuracy, but also by sensor interpretability and computational efficiency considered in an integrated manner.

Keywords: *IoT, Machine Learning, Sensor Analysis, Feature Importance, Permutation Importance*

РЕЗЮМЕ

ИНТЕРПРЕТАЦИЯ ЗНАЧИМОСТИ ДАТЧИКОВ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВЫЧИСЛИТЕЛЬНОЙ ЭФФЕКТИВНОСТИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ В ИНТЕЛЛЕКТУАЛЬНЫХ БИБЛИОТЕЧНЫХ СРЕДАХ НА ОСНОВЕ ИОТ

Серкан Мамедов

Нахчыванский государственный университет, Нахчыван, Азербайджан

Данное исследование направлено на интерпретацию влияния датчиков окружающей среды на процессы принятия решений моделей машинного обучения в интеллектуальных библиотечных средах на основе IoT, а также на оценку их вычислительной эффективности. На основе многомерных данных датчиков, собранных в библиотеке Университета Дюздже, семь алгоритмов машинного обучения — KNN, Random Forest, Decision Tree, SVM, XGBoost, Logistic Regression и Naïve Bayes — были применены в сравнительном контексте. Основным научным вкладом данного исследования является параллельное применение двух подходов к оценке значимости датчиков: метода важности признаков на основе Random Forest (Feature Importance) и метода перестановочной важности (Permutation Importance). Данная двухметодная структура позволила более точно проанализировать взаимосвязь между внутренними механизмами принятия решений модели и фактическим влиянием переменных датчиков на прогностическую эффективность. Результаты показывают, что метод Feature Importance определяет температуру и плотность присутствия людей как доминирующие факторы, тогда как анализ Permutation Importance выявляет, что интенсивность освещения и переменные на основе CO₂ оказывают более значительное влияние на производительность модели. Данное расхождение свидетельствует о том, что опора исключительно на внутренние показатели важности модели недостаточна для точной интерпретации данных датчиков, что подчёркивает необходимость применения более надёжных методологий оценки. Кроме того, с точки зрения вычислительной эффективности модели KNN и Decision Tree продемонстрировали более подходящие характеристики для IoT-приложений реального времени благодаря низкому потреблению ресурсов и высокой скорости выполнения. В целом полученные результаты свидетельствуют о том, что оптимальный выбор модели в IoT-системах должен определяться не только прогностической точностью, но и интерпретируемостью датчиков и вычислительной эффективностью, рассматриваемыми в интегрированном контексте.

Ключевые слова: *Интернет вещей, машинное обучение, анализ датчиков, важность признаков, перестановочная важность*