T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

İÇ MEKÂN NOKTA BULUTU VERİLERİNİN MAKİNE ÖĞRENİMİ YÖNTEMLERİ KULLANILARAK SINIFLANDIRILMASI

Sena VARBİL

YÜKSEK LİSANS TEZİ

Harita Mühendisliği Anabilim Dalı

Uzaktan Algılama ve CBS Programı

Danışman

Doç. Dr. Alper ŞEN

Haziran, 2025

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

İÇ MEKÂN NOKTA BULUTU VERİLERİNİN MAKİNE ÖĞRENİMİ YÖNTEMLERİ KULLANILARAK SINIFLANDIRILMASI

Sena VARBİL tarafından hazırlanan tez çalışması 16.06.2025 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı, Uzaktan Algılama ve CBS Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Alper ŞEN
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Doç. Dr. Alper ŞEN, Danışman
Yıldız Teknik Üniversitesi _____

Prof. Dr. Ali Melih BAŞARANER, Üye
Yıldız Teknik Üniversitesi _____

Prof. Dr. Nesibe Necla ULUĞTEKİN, Üye
İstanbul Teknik Üniversitesi _____

Danışmanım Doç. Dr. Alper ŞEN sorumluluğunda tarafımca hazırlanan “İç Mekân Nokta Bulutu Verilerinin Makine Öğrenimi Yöntemleri Kullanılarak Sınıflandırılması” başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Sena VARBİL

İmza

Sevgili aileme…

TEŞEKKÜR

Yüksek lisans eğitimim süresince bana destek olan, tezimin her aşamasında bana yol göstererek mesleki gelişimime katkı sağlayan ve yeni bilgiler edinmemde rol oynayan değerli danışman hocam Doç Dr. Alper ŞEN'e,

Yüksek lisans ders ve tez yazım sürecimde bana güvenen maddi manevi her konuda arkamda durup beni destekleyen ve hep cesaretlendiren sevgili aileme,

Yüksek lisansımın her aşamasında hep yanımda olan ve beni destekleyen sevgili arkadaşım İrem TEK'e teşekkür ederim.

Sena VARBİL

İÇİNDEKİLER

SİMGE LİSTESİ vii

KISALTMA LİSTESİ ix

ŞEKİL LİSTESİ x

TABLO LİSTESİ xi

HARİTA LİSTESİ xii

ÖZET xiii

ABSTRACT xv

1 GİRİŞ 1

2 LİTERATÜR ARAŞTIRMASI 6

3 MATERYAL VE YÖNTEM 12

3.1 Çalışmada Kullanılan Veri Seti 12

3.2 Çalışmada Kullanılan Programlar 14

3.2.1 Anaconda Navigator Programı 14

3.2.2 CloudCompare Programı 14

3.2.3 ArcGIS Pro Programı 15

3.3 Min-Max Normalizasyonu 16

3.4 Kullanılan Makine Öğrenimi Yöntemleri 17

3.4.1 Rastgele Orman (RO) Yöntemi 18

3.4.2 Çok Katmanlı Algılayıcı (ÇKA) Yöntemi 19

3.4.3 CatBoost Yöntemi 20

3.4.4 XGBoost Yöntemi 21

3.4.5 LightGBM Yöntemi 22

3.4.6 CART Yöntemi 23

3.5 Sınıflandırma için Performans Değerlendirme Ölçütleri 24

4 UYGULAMA VE BULGULAR 25

4.1 Çalışmada Elde Edilen Sınıflandırma Sonuçları 26

4.2 Tahmin Sonuçlarının Cloud Compare Yazılımında Görünümü 34

4.2.1 Ofis 3 - LightGBM Yöntemi 35

4.2.2 Ofis 4 - Catboost Yöntemi 35

4.2.3 Ofis 5 - Catboost Yöntemi 35

4.2.4 Ofis 7 - Çok Katmanlı Algılayıcı Yöntemi 36

4.2.5 Ofis 8 - Rastgele Orman Yöntemi 36

4.2.6 Ofis 10 - XGBoost Yöntemi 36

4.3 Planların Oluşturulması 37

4.3.1 Ofis 3 - LightGBM Yöntemi 37

4.3.2 Ofis 4 – CatBoost Yöntemi 38

4.3.3 Ofis 5 - CatBoost Yöntemi 39

4.3.4 Ofis 7- Çok Katmanlı Algılayıcı Yöntemi 40

4.3.5 Ofis 8 – Rastgele Orman Yöntemi 41

4.3.6 Ofis 10 – XGBOOST Yöntemi 42

4.4 Planı Oluşturulamayan Ofisler 43

SONUÇ 44

KAYNAKÇA 46

TEZDEN ÜRETİLMİŞ YAYINLAR 50

SİMGE LİSTESİ

$\boldsymbol{K}$	Ağaç Sayısı
y	Bağımlı Değişken
$\boldsymbol{x}$	Bağımsız Değişken
$\boldsymbol{y_i}$	Bağımsız Değişken
$\boldsymbol{g_i}$	Birinci Dereceden Gradyan İstatiği
$\boldsymbol{O_j}$	Çıkış Değeri
$\boldsymbol{DN}$	Doğru Negatif
$\boldsymbol{DP}$	Doğru Pozitif
$\boldsymbol{W_{ij}}$	Düğümler Arasındaki Ağırlık
$\boldsymbol{a_n}$	Eğitim Adım Boyutu
$\boldsymbol{x_{max}}$	En Büyük Öznitelik Değeri
$\boldsymbol{x_{min}}$	En Küçük Öznitelik Değeri
$\boldsymbol{Y_i}$	Etiketli Değerler
$\boldsymbol{F}$	Fonksiyon Uzayı
$\boldsymbol{PE^*}$	Genelleştirilmiş Hata
$\boldsymbol{F_{k-1}}$	Güçlü Öğrenici
$\boldsymbol{O_i}$	i Düğümünün Çıktısı
$\boldsymbol{h_i}$	İkinci Dereceden Gradyan İstatiği
$\boldsymbol{M_0}$	İterasyon

$\boldsymbol{N}$	İterasyon
$\boldsymbol{Q_j}$	j Düğümünün Eşik Değeri
$\boldsymbol{E}(.)$	Kayıp Fonksiyonu Beklentisi
$\boldsymbol{L}(.)$	Kayıp Fonksiyonu
$\boldsymbol{mg}()$	Margin Fonksiyonu
$\boldsymbol{\Omega}$	Modelin Karmaşıklığı
$\boldsymbol{I_j}$	Net Girdi Değeri
$\boldsymbol{M}$	Nihai Model
$\boldsymbol{X_i'}$	Normalizasyon Değeri
$\boldsymbol{x_i}$	Öznitelik Değeri
$\boldsymbol{t_k}$	Parametre
$\boldsymbol{I_j}$	Veri Noktalarının İndeks Kümesi
$\boldsymbol{YN}$	Yanlış Negatif
$\boldsymbol{YP}$	Yanlış Pozitif
$\boldsymbol{T}$	Yaprak Sayısı
$\boldsymbol{\omega}$	Yapraklardaki Skorların Vektörü

KISALTMA LİSTESİ

CART	Classification and Regression Trees
CatBoost	Categorical Boosting
ÇKA	Çok Katmanlı Algılayıcı
DGCNN	Dynamic Graph Convolutional Neural Network
FPN	Feature Pyramid Network
ISPRS	International Society for Photogrammetry and Remote Sensing
LightGBM	Light Gradient Boosting Machine
LRGNet	Learnable Region Growing Network
MLP	Multilayer Perceptron
NAS	Neural Architecture Search
RF	Random Forest
RGAM	Neural Network with Attention Module
RO	Rastgele Orman
S3DIS	Stanford 3D Indoor Scene
SACNN	Self-Augment Convolutional Neural Network
SAConv	Self-Augment Convolution
XGBoost	Extreme Gradient Boosting

ŞEKİL LİSTESİ

Şekil 1.1 Çalışmada kullanılan 10 ofis odasının Cloud Compare programındaki görünümü 5

Şekil 3.1 İç mekân nokta bulutu verisinin makine öğrenmesi yöntemleriyle sınıflandırılması ve görselleştirilmesi iş akış şeması 12

Şekil 3.2 Kullanılan ofis verilerinin iç görünümü 13

Şekil 3.3 Yapay zeka yöntemlerinin birbiriyle ilişkisi 18

Şekil 3.4 CART algoritması sınıflandırma diyagramı 23

Şekil 4.1 Ofis 3 verisi üzerindeki LightGBM yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 35

Şekil 4.2 Ofis 4 verisi üzerindeki CatBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 35

Şekil 4.3 Ofis 5 verisi üzerindeki CatBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 35

Şekil 4.4 Ofis 7 verisi üzerindeki Çok Katmanlı Algılayıcı yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 36

Şekil 4.5 Ofis 8 verisi üzerindeki Rastgele Orman yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 36

Şekil 4.6 Ofis 10 verisi üzerindeki XGBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN) 36

TABLO LİSTESİ

Tablo 1.1 Sınıfların birleştirilmeden önceki etiket değerleri 4
Tablo 1.2 Sınıfların birleştirildikten sonraki etiket değerleri 4
Tablo 3.1 Karışıklık Matrisi 24
Tablo 4.1 Ofis 2 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 26
Tablo 4.2 Ofis 3 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 27
Tablo 4.3 Ofis 4 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 28
Tablo 4.4 Ofis 5 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 29
Tablo 4.5 Ofis 6 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 30
Tablo 4.6 Ofis 7 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 31
Tablo 4.7 Ofis 8 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 32
Tablo 4.8 Ofis 9 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları ... 33
Tablo 4.9 Ofis 10 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları . 34

HARİTA LİSTESİ

Harita 4.1 LightGBM yöntemiyle Ofis 3 planı 37

Harita 4.2 CatBoost yöntemiyle ofis 4 planı 38

Harita 4.3 CatBoost yöntemiyle Ofis 5 planı 39

Harita 4.4 ÇKA yöntemiyle ofis 7 planı 40

Harita 4.5 RO yöntemiyle ofis 8 planı 41

Harita 4.6 XGBoost yöntemiyle ofis 10 planı 42

Harita 4.7 (a) Ofis 2 verisinin CatBoost yöntemine göre sınıflandırılması (b) Ofis 6 verisinin RO yöntemine göre sınıflandırılması; (c) Ofis 9 verisinin CatBoost yöntemine göre sınıflandırılması 43

ÖZET

İç Mekân Nokta Bulutu Verilerinin Makine Öğrenimi Yöntemleri Kullanılarak Sınıflandırılması

Sena VARBİL

Harita Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç. Dr. Alper ŞEN

İç mekânlara ait 3B nokta bulutu sınıflandırması, iç mekân harita yapımı, iç mekân navigasyonu, bina yenileme, tesis yönetimi vb. uygulamalarda iç mekân modellerinin oluşturulmasında büyük önem taşımaktadır. Bu çalışmada, Stanford Üniversitesi tarafından üretilen S3DIS veri setinde bulunan ofis odalarına ait nokta bulutları makine öğrenmesi yöntemlerinden Rastgele Orman, Çok Katmanlı Algılayıcı, Kategorik Artırma (CatBoost), Aşırı Gradyan Artırma (XGBoost), Hafif Gradyan Artırma Makinesi (LightGBM) ve Sınıflandırma ve Regresyon Ağaçları (CART) yöntemleri ile sınıflandırılarak iç mekân planları oluşturulmuştur. Giriş verileri için X, Y, Z koordinatları ve R, G, B renk değerleri öznitelik bilgileri olarak kullanılmıştır. Sınıflar tavan, zemin, duvar, kapı, pencere, kolon, masa, sandalye, pano, dağınıklık ve kitaplık nesnelerini kapsamaktadır. Eğitim ve test verilerinde iç mekân planlarının oluşturulması amacıyla duvar, kapı, pencere, kolon, pano ve kitaplık bir sınıf (birleştirilmiş sınıf-1); masa, sandalye ve dağınıklık bir sınıf (birleştirilmiş sınıf-2) halinde birleştirilmiştir. İç mekân nokta bulutlarının sınıflandırılması için kullanılan makine öğrenimi yöntemlerinden Rastgele Orman yöntemiyle ortalama %90, Çok Katmanlı Algılayıcı yöntemiyle ortalama %87, CatBoost yöntemiyle ortalama %91, XGBoost yöntemiyle ortalama %88,

LightGBM yöntemiyle ortalama %89 ve CART yöntemiyle ortalama %86 genel sınıflandırma doğruluğu elde edilmiştir.

Anahtar Kelimeler: İç mekân, Nokta bulutu, Sınıflandırma, Makine öğrenmesi

ABSTRACT

Indoor Point Cloud Data Classification Using Machine Learning Methods

Sena VARBİL

Department of Geomatics Engineering
Master of Science

Supervisor: Assoc. Prof. Dr. Alper ŞEN

The classification of 3D point clouds for indoor environments plays a crucial role in the creation of indoor models for applications such as indoor mapping, indoor navigation, building renovation, facility management, etc. In this study, indoor plans were generated by classifying point clouds of office rooms from the S3DIS dataset, produced by Stanford University, using machine learning methods such as Random Forest, Multi-Layer Perceptron, Categorical Boosting (CatBoost), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Classification and Regression Trees (CART). For the input data, X, Y, Z coordinates and R, G, B color values are used as attribute information. The classes include ceiling, floor, wall, door, window, column, table, chair, board, clutter, and bookcase objects. In the training and test datasets, to create indoor plans, wall, door, window, column, board, and bookcase were combined into a single class (merged class-1), while table, chair, and clutter were grouped into another class (merged class-2). Among the machine learning methods used for the classification of indoor point clouds, Random Forest method has an average classification accuracy of 90%, Multilayer Perceptron method has an average classification accuracy of 87%, CatBoost method has an average classification accuracy of 91%,

XGBoost method has an average classification accuracy of 88%, LightGBM method has an average classification accuracy of 89% and CART method has an average classification accuracy of 86%.

Keywords: Indoor, Point cloud, Classification, Machine learning

1
GİRİŞ

Bu çalışma, iç mekânlarda elde edilen nokta bulutu verilerinin makine öğrenmesi yöntemleriyle sınıflandırılması ve bu sınıflandırma sonuçlarından harita üretimine kadar olan süreç plan ve görselleştirme anlamında ele alınmıştır. Çalışmada, makine öğrenmesi yöntemleri kullanılarak iç mekân nokta bulutlarının yüksek doğrulukla sınıflandırılması hedeflenmiş; bu sınıflandırmanın, mekânsal bilginin görselleştirilmesi ve kullanımı açısından nasıl bir önem taşıdığı araştırılmıştır. İç mekânlarda bulunan eşyaların ve yapısal bileşenlerinin doğru bir şekilde ayrıştırılması, verinin çok boyutlu yapısı ve düzensiz dağılımı nedeniyle oldukça zorlu bir süreçtir. Bu tez çalışması, bu tür sınırlamaları aşmayı amaçlayarak, makine öğrenmesi teknikleriyle daha verimli ve yüksek doğruluklu sınıflandırma yapmayı hedeflemektedir.

Nokta bulutları, günümüzde hem açık alanlarda hem de kapalı mekânlarda geniş bir mühendislik ve bilimsel uygulama yelpazesinde kullanılabilen yaygın bir mekân sal bilgi kaynağıdır [1]. Son yıllarda lazer tarayıcıların ve dijital görüntülerin hızla gelişmesiyle birlikte, mekânsal 3-Boyutlu (3B) nokta bulutu verileri birçok alanda yaygın olarak kullanılmaya başlanmıştır [2]. İç mekân ortamlarıyla ilişkili 3B mekânsal bilgi risk ve afet yönetimi, insan hareket yollarının belirlenmesi ve tesis yönetimi gibi çeşitli uygulamalarda giderek daha fazla talep görmektedir [3]. Dış mekâna ait nokta bulutlarının sınıflandırılmasında önemli ilerlemeler kaydedilmiştir. Ancak iç mekân sahnelerinin dar bir çevreye, çeşitli yapısal özelliklere ve kolon, duvar vb. birçok engele sahip olması nedeniyle sınıflandırılması hâlâ bir sorundur [4]. Bu nedenle, iç mekânların doğru ve güncel durumunu yansıtan üç boyutlu modellere ve güncel mekânsal verilere olan ihtiyaç giderek artmaktadır. İç mekân modelleme, çeşitli araç ve teknolojilerle desteklenen bir süreçtir; iç mekânların fiziksel ve işlevsel özellikleri dijital olarak temsil edilir

ve yönetilir. İç mekân haritaları iç mekân ölçmelerinin sonucu olarak modelleme ile temsil, birçok iç odaklı uygulamanın temelini oluşturur [5]. İç mekân ortamlarının 3B modelleri, yeniden tasarımdan izleme ve simülasyona kadar bir dizi farklı uygulama alanında kullanılabildikleri için giderek daha fazla önem kazanmaktadır. Görüntülerden veya nokta bulutlarından otomatik olarak 3B iç mekânların yeniden oluşturulması süreci daha kolay, hızlı ve ekonomik hale getirmektedir [6]. Nokta bulutu elde edildikten sonra, genellikle içerdiği semantik bilgilerin anlaşılması için sınıflandırılması gereklidir. Bu nedenle nokta bulutu bölütleme teknolojisi birçok uygulama için büyük önem taşır [2]. Nokta bulutu semantik bölütlemesi, nokta kümesindeki her noktayı sınıflandırmayı hedefler. Semantik bölütleme otonom navigasyon, sanal gerçeklik, robot nesne etiketleme ve yüksek kaliteli iç mekân haritalama gibi birçok gerçek dünya uygulamasına büyük ölçüde fayda sağlar [7]. Nokta bulutu için geleneksel bölütleme yöntemleri birçok araştırmacının çalışmalarıyla sürekli olarak geliştirilmiştir. Bununla birlikte, geleneksel nokta bulutu bölütleme yöntemleri, manuel olarak etiketleme yoluyla tasarımcıların deneysel bilgiye sahip olmasını gerektirmektedir. Nokta bulutu bölütlemesinde kullanılması gereken bir dizi eşik değeri bulunur ve bu yöntemler genellikle karmaşık olduğundan, yalnızca belirli görevler için uygunken genel olarak performansları zayıf kalmaktadır [8].

İç mekân konumlandırma için birçok teknik mevcut olsa da iç mekân konum bilgisini ölçmelerden çıkarabilen sınırlı sayıda algoritma ve yöntem bulunmaktadır. Üç boyutlu mekânsal veri edinimi, şehir ve bina modelleme alanında bilimsel ve teknolojik olarak hızla ilerlemektedir. Bu ilerlemeler, iç mekân modelleme ve harita yapımı için daha gelişmiş donanım, yazılım, standartlar, teknikler ve uygulamaların ortaya çıkmasını sağlamaktadır. Sensörler, iç mekân bilgisi ediniminde anahtar bir rol oynamaktadır [3]. İç mekân nokta bulutu veri toplama sistemleri, fiziksel yapılarına göre el tipi, sırt çantası ve el arabası gibi sınıflara ayrılabilir. Araba konfigürasyonu, geniş bir sensör seçimi ve ağırlık esnekliği sunarak daha güçlü sensörlere erişim sağlar. Bu durum, daha güçlü sensörlerle daha yoğun nokta bulutu verisi sağlayabilir. Ancak, daha güçlü sensörlerin kullanımı araç maliyetini artırabilir ve manevra kabiliyetini azaltabilir. Özellikle dar alanlarda veya sınırlı alanlarda (merdivenler veya dar koridorlar gibi) fiziksel kısıtlamalarla karşılaşılabilir. Öte yandan, elde taşınabilir cihazlar daha fazla manevra kabiliyetine

sahiptir ve satın alırken daha fazla esneklik sağlayabilirler. Ancak, bu cihazların sınırlı ağırlık taşıma kapasitesi ve pil ömrü, harita yapımı sensörünün özelliklerini ve tamamlayıcı sensörlerin kullanımını kısıtlayabilir. Bu durum, belirli özelliklerin entegrasyonunda ve ek sensörlerin kullanımında bazı sınırlamalara yol açabilir. Veri işleme ve sensör senkronizasyonunda temel unsur, genellikle LiDAR veya RGB-D kamerası gibi donanımları içinde barındıran harita yapımı sensörüdür [9]. Böylece 3B tarayıcılar büyük miktarda veriyi yakalayarak yapay zekâ destekli makinelerin dünyayı daha doğru algılamasını ve tanımasını sağlamaktadır [10].

Kartografik harita yapımının başlıca odak noktasının mekânsal bilginin etkin bir şekilde aktarımını geliştirmek olduğu düşünüldüğünde, nokta bulutları ve örgülü (meshed) modeller genellikle harita olarak kabul edilmezken, mimari CAD çizimleri ve Yapı Bilgi Modellemesi (Building Information Modeling, BIM) ise ilkel haritalar olarak düşünülebilir. İç mekân ortamlara ait kartografik özellikler taşıyan haritalarının yapımı, iç mekânları modelleyen diğer disiplinlerle hem örtüşmekte hem de onlardan ayrılmakta ve iç mekân temsilinde nispeten keşfedilmemiş bir alanı temsil etmektedir. Bu bağlamda, nokta bulutlarının bölütlenmesi sonucunda elde edilen nesne sınıfları kullanılarak kartografik haritaların yapımı önemli bir araştırma konusudur [11]. İç mekânlarda yayalar için haritaların oluşturulması, özellikle büyük ve karmaşık yapılar için oldukça önemlidir. Bu tür haritalar, ziyaretçilere veya çalışanlara mekân içindeki konumlarını belirleme ve istedikleri hedefe yönlendirme konusunda yardımcı olabilir. İç mekânın güncel durumunu yansıtan nokta bulutlarının yüksek doğruluklu nesne sınıflarını sağlayacak biçimde bölütlenmesi ile iç mekâna ait haritalarının oluşturulması ve bu haritalardan iç mekân navigasyon ağ modellerinin otomatik olarak üretimi iç mekân navigasyonu sistemlerinin önemli bileşenleri arasında yer alır [12].

Bu çalışmada, Stanford Üniversitesi tarafından geliştirilen Stanford 3D Indoor Scene (S3DIS) veri setinde yer alan bazı ofis odalarına ait nokta bulutları kullanılarak iç mekân sınıflandırma ve plan oluşturma işlemleri gerçekleştirilmiştir. S3DIS veri seti, iç mekânlarda yer alan farklı türdeki nesnelerin üç boyutlu temsillerini içermesi nedeniyle, bu çalışmada iç mekân sınıflandırması için uygun bir zemin oluşturmuştur. Çalışmada, Rastgele Orman, Çok Katmanlı Algılayıcı, Kategorik Artırma (CatBoost), Aşırı Gradyan Artırma (XGBoost), Hafif Gradyan

Artırma Makinesi (LightGBM) ve Sınıflandırma ve Regresyon Ağaçları (CART) gibi makine öğrenmesi yöntemleri uygulanarak, sınıflandırma işlemleri gerçekleştirilmiştir. Sınıflandırma işlemleri sırasında, veri setinde yer alan sınıfların homojen olmaması nedeniyle, bazı sınıfların birleştirilmesi yoluna gidilmiştir. Bu süreçte, benzer özelliklere sahip sınıflar, daha anlamlı bir sınıflandırma yapabilmek amacıyla gruplandırılmıştır. Böylece sınıflandırma işlemi daha verimli hale getirilmiş ve karmaşık iç mekân yapılarının anlaşılmasını kolaylaştırmıştır. Bu gruplandırma sürecine bağlı olarak, veri setinde başlangıçta kullanılan sınıfların orijinal etiket değerleri Tablo 1.1'de; sınıfların birleştirilmesi sonrası oluşan etiket değerleri ise Tablo 1.2'de verilmiştir. Bu işlem sonucunda, daha sade ve anlaşılır bir sınıflandırma elde edilmiş ve elde edilen sınıflandırma sonuçları, iç mekân planlarının oluşturulmasında kullanılmıştır. Oluşturulan planlar, sınıflandırma sonuçlarına dayalı olarak iç mekânların nesnelerini görsel olarak temsil etmekte ve sınıflandırmanın doğruluğunu göstermek amacıyla da kullanılmaktadır.

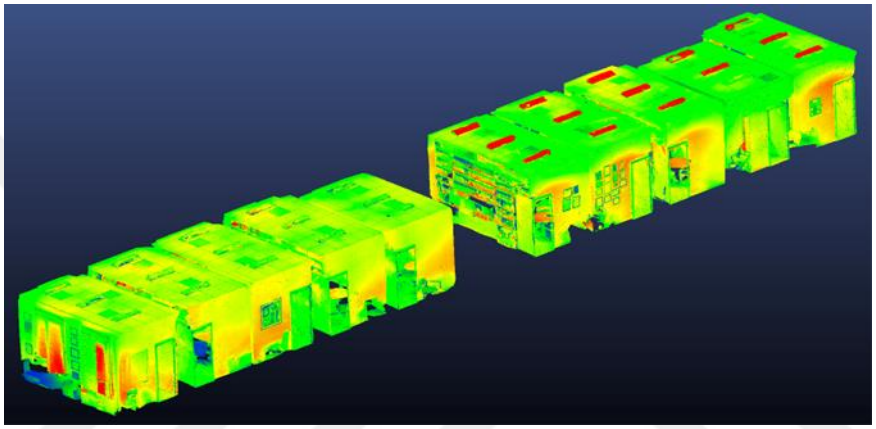
Tablo 1.1 Sınıfların birleştirilmeden önceki etiket değerleri

Sınıf	**Etiket**
Tavan	0
Zemin	1
Duvar	2
Kapı	3
Pencere	4
Kolon	5
Masa	6
Sandalye	7
Pano	8
Dağınıklık	9
Kitaplık	10

Tablo 1.2 Sınıfların birleştirildikten sonraki etiket değerleri

Sınıf	**Etiket**
Tavan	0
Zemin	1
Birleştirilmiş Sınıf-1 (duvar, kapı, pencere, kolon, pano, kitaplık)	2
Birleştirilmiş Sınıf-2 (masa, sandalye, dağınıklık)	3

Veri setindeki her sınıf için kesinlik, duyarlılık, F1 skoru ve doğruluk değerleri hesaplanmıştır. Tüm işlemler ve sınıflandırma aşamaları Python programlama diline ait Scikit-learn kütüphanesi kullanılarak gerçekleştirilmiş, görselleştirmeler ise Cloud Compare ve ArcGIS Pro yazılımları ile yapılmıştır. Çalışma sonucunda, Rastgele Orman yöntemiyle ortalama %90, Çok Katmanlı Algılayıcı yöntemiyle ortalama %87, CatBoost yöntemiyle ortalama %91, XGBoost yöntemiyle ortalama %88, LightGBM yöntemiyle ortalama %89 ve CART yöntemiyle ortalama %86 genel sınıflandırma doğruluğu elde edilmiştir. Çalışmada kullanılan 10 ofis odasının Cloud Compare yazılımındaki görünümü Şekil 1.1' de gösterilmiştir.



Şekil 1.1 Çalışmada kullanılan 10 ofis odasının Cloud Compare programındaki görünümü

Bu tez çalışmasının birinci bölümünde araştırma konusuna giriş yapılmış, ikinci bölümde ilgili literatür incelenmiş, üçüncü bölümde ise kullanılan materyal ve yöntemler detaylı olarak açıklanmıştır. Dördüncü bölümde uygulama süreci ve elde edilen bulgulara yer verilmiş, son olarak sonuç ve değerlendirme kısmında çalışma genel olarak ele alınmıştır.

2
LİTERATÜR ARAŞTIRMASI

İç mekân nokta bulutu verilerinin makine öğrenmesi yöntemleriyle sınıflandırılması ve bu verilerden haritaların üretilmesi, literatürde giderek daha fazla dikkat çeken bir araştırma alanı olarak karşımıza çıkmaktadır. Bu bölümde, konuyla ilgili mevcut literatür incelenerek, farklı sınıflandırma algoritmalarının performansı, kullanılan veri setleri, elde edilen doğruluk değerleri üzerine yapılan çalışmalar detaylandırılacaktır.

[13], yaptıkları çalışmada büyük ölçekli ve renkli S3DIS iç mekân veri seti üretilmiş ve hiyerarşik bir yaklaşım kullanılarak tüm bir binanın 3B nokta bulutunun anlamsal ayrıştırılması için bir yöntem önermişlerdir. Çalışmada önce noktalar arasındaki boşluk örüntülerini temel alan bir algoritma ile yatay ve düşey eksenler kullanılarak nokta bulutu yatay ve düşey alanlara bölütlenmiştir. Aşırı bölütlenmiş alanlar, topolojik düğüm noktaları arasındaki ilişkiler kullanılarak birleştirilmiştir. Böylece ana ayırıcı niteliğinde olan binaya ait duvar noktaları sınıflandırılmıştır. Diğer nesneler, belli bir koordinat sistemine getirilen bölütlenmiş alanlarda 3B kutular gezdirilerek sınıflandırılmıştır. Bu yöntem ile tahmin edilen sınıf etiketleri, uzmanlarca manuel olarak sınıflandırılarak üretilen referans (ground truth) etiketleri ile karşılaştırıldığında tavan ve zemin sınıfı için sırasıyla %71.61 ve %88.70 doğruluk; duvar, kolon ve pencere gibi temel bina elemanlarına ait sınıflar için ise %67.38 genel doğruluk elde edilmiştir.

[14], yaptıkları çalışmada iç mekân 3B nokta bulutlarını bölütlemek için denetimsiz bir bölge büyütme bölütleme yöntemi sunmuşlardır. Bu yaklaşım, verilerin etiketlenmesine ihtiyaç duymadan, geometrik ve yapısal benzerliklere dayalı olarak noktaları gruplar halinde ayırmayı hedeflemektedir. Bölge büyütme algoritması, rastgele başlangıç noktalarından başlayarak, benzer özellik gösteren komşu noktaları genişletir ve böylece her bir bölgeyi oluşturur. Çalışmada,

bölütleme sürecinin performansını artırmak için kendini uyarlayan, ver temelli sezgisel çıkarım süreci yaklaşımı geliştirilmiş ve bu yöntem, optimal parametrelerin seçilmesi için kullanılmıştır. Bu optimizasyon, bölütlemenin farklı iç mekân yapılarında tutarlı ve etkili olmasını sağlamayı amaçlamıştır. Bölütlemenin ardından, her bir bölütün belirli özelliklere (geometri, yoğunluk gibi) göre sınıflandırılması gerçekleştirilmiştir. Yöntemin doğruluğunu ve genel uygulanabilirliğini değerlendirmek için geniş bir iç mekân verisi olan S3DIS veri seti üzerinde testler yapılmıştır. Deneysel sonuçlar, yönteminin karmaşık ve büyük ölçekli iç mekân verileri üzerinde yüksek doğruluk sağladığını göstermiştir ve %88.10 F1 skor elde edilmiştir.

[15], yaptıkları çalışmada iç mekân sahnelerine odaklanılarak 3B nokta bulutu semantik bölütlemesi için yenilikçi bir ağ mimarisi geliştirmişlerdir. Özellikle, Dikkat Modülü ile Halka Gruplama Sinir Ağı (Neural Network with Attention Module - RGAM) adı verilen yeni bir ağ yapısı önermişlerdir. İlk olarak, her noktanın yüksek boyutlu uzaydaki özelliklerini çıkarmak için Çok Katmanlı Algılayıcı kullanılmıştır. Daha sonra, temsilci noktalar en uzak nokta örnekleme yöntemiyle seçilmiş ve bu noktalardan çok ölçekli bilgileri çıkarmak amacıyla çok ölçekli halka gruplama öğrenimi uygulanmıştır. Bu adım, tekrarlanan noktaların hesaplanmasını azaltarak algoritmanın hızını artırırken, farklı ölçeklerdeki nesnelere uyarlanabilirliği sağlamaktadır. Ardından, uzun menzilli bağlamsal bilgileri yakalamak amacıyla çok ölçekli komşuluklar arasında mekânsal dikkat modülleri oluşturulmuştur. Böylece, uzaydaki nesneler arasındaki anlamsal ilişkiler kullanılarak sınıflandırma doğruluğu artırılmıştır. Ayrıca, farklı ölçeklerdeki bağımsız ve komşu özellikleri birleştirmek için ters mesafe ağırlıklı bir yöntem kullanılmıştır. Bu sayede, yerel anlamsal bilgiler daha iyi öğrenilmiş ve modelin doğruluğu geliştirilmiştir. Son olarak, kanal dikkat modülü eklenerek kanallar arasındaki korelasyonların tanınması sağlanmış ve bölütleme doğruluğu artırılmıştır. RGAM, özellikle karmaşık iç mekân ortamlarındaki 3B nokta bulutlarının segmentasyonu için tasarlanmıştır ve bu tür ortamların içerdiği yapısal unsurları daha iyi ayırt etmeyi hedeflemektedir. Yapılan araştırmada, önerilen ağ mimarisinin performansı, yaygın olarak kullanılan S3DIS, ScanNet ve NYU-V2 veri kümeleri üzerinde kapsamlı bir şekilde değerlendirilmiştir. Bu veri kümeleri, özellikle bina içi sahnelerin modellenmesi ve analizi açısından zengin veri

sağlayarak, modelin gerçek dünya uygulamalarında ne kadar etkili olabileceğini test etme olanağı sunmuştur. Deneysel sonuçlar, S3DIS veri setinde önerilen modelin en iyi genel doğruluk ve ortalama kesişim-birleşim oranı (mean intersection-over-union, mIoU) değerlerini elde ettiğini göstermiştir. Özellikle S3DIS veri setinde %87.05 oranında genel doğruluk elde edilmiştir.

[16], yaptıkları çalışmada 3B nokta bulutu semantik bölütlemesi için yenilikçi bir yöntem önermişlerdir. Bu yöntem, iç mekân sahnelerindeki nesnelerin daha doğru bir şekilde tanımlanmasını sağlamak amacıyla iki ana modül üzerine kurulmuştur. İlk modül, çok ölçekli seyrek evrişimsel sinir ağı yapısını kullanır. Bu modül, nokta bulutu verisinin seyrek doğasından yararlanarak farklı ölçeklerdeki öznitelikleri bir araya getirmiş bu sayede hem küçük nesneler hem de büyük yapılar gibi çeşitli boyutlardaki yapıların daha hassas bir şekilde bölütlenmesi mümkün olmuştur. Çok ölçekli yapısı sayesinde, hem ayrıntılı yerel özellikler hem de geniş kapsamlı global özellikler etkili bir şekilde işlenmiştir. İkinci modül ise kanal dikkat mekanizması adı verilen bir strateji ile özelliklerin daha verimli kullanılmasını sağlamaktadır. Bu mekanizma, öğrenme süreci sırasında modelin önemli özelliklere odaklanmasını sağlarken, daha az önemli olan özelliklerin etkisini azaltmıştır. Kanal dikkat mekanizması, modelin daha doğru ve verimli bir şekilde çalışmasına katkı sağlamıştır çünkü veri içerisindeki kritik özellikler öne çıkarmış ve bölütleme performansını arttırmıştır. Yöntemin etkinliği, yaygın olarak kullanılan S3DIS veri seti üzerinde test edilmiş ve %67 mIoU (mean Intersection over Union) skoru elde edilmiştir.

[2], yaptıkları çalışmada iç mekân noktalarının otomatik olarak anlamsal bölütlenmesine (semantic segmentation) odaklanmışlardır. Araştırmacılar, özellikle sütun, kiriş, duvar, zemin ve tavan gibi iç mekân yapısal unsurlarını yüksek doğrulukla tespit etmek için Dinamik Grafik Evrişimli Sinir Ağı (Dynamic Graph Convolutional Neural Network-DGCNN) kullanmışlardır. DGCNN, geleneksel derin öğrenme yaklaşımlarından farklı olarak, nokta bulutu verilerinin düzensiz ve karmaşık yapısını daha iyi öğrenebilen, üç boyutlu bir derin öğrenme mimarisidir. Bu yöntem, nokta bulutları arasında dinamik olarak bağlantılar kurarak, noktaların çevresel ilişkilerini dikkate almakta ve bu sayede bölütleme doğruluğunu önemli ölçüde arttırmaktadır. Araştırmacılar, DGCNN modelini eğitmek ve test etmek için S3DIS veri setini kullanmışlardır. DGCNN'nin

kullanımıyla model, çeşitli yapısal unsurların bölütlemesi sırasında yüksek doğruluk oranlarına ulaşmış ve S3DIS veri setinde %86.90 doğruluk elde edilmiştir.

[17], yaptıkları çalışmada bölge büyütme sürecinin kendisinin veriden öğrenildiği bir nokta bulutu bölütleme yöntemi olan LRGNet (Learnable Region Growing Network) yöntemini tanıtmışlardır. Bu yöntem, geleneksel bölge büyütme algoritmalarından esinlenerek geliştirilmiş olup, bu süreci tamamen öğrenilebilir bir hale getirmektedir. LRGNet, başlangıç noktalarını öğrenilebilir bir şekilde seçerek, bu noktalar etrafında bölütler oluşturur ve her adımda hangi noktaların mevcut bölüte dahil edileceğini öğrenir. Bu yaklaşım, modelin nokta bulutlarının yapısını otomatik olarak öğrenmesini ve bölütleme işlemini dinamik bir şekilde gerçekleştirmesini sağlamıştır. Uygulama için ScanNet ve S3DIS veri setleri kullanılmıştır. Uygulama önce ScanNet veri setinde eğitilen verinin S3DIS veri setinde test edilmesi ve sonrasında tersi olacak şekilde S3DIS üzerinde eğitilip ScanNet üzerinde test edilmesine dayanmaktadır.

[18], yaptıkları çalışmada kapalı alan modelleme alanında bir referans noktası oluşturmayı amaçlayan ISPRS (International Society for Photogrammetry and Remote Sensing) tarafından başlatılan bir benchmark çalışmasının ön sonuçlarını sunmaktadır. Farklı yöntemlerin performanslarının, yeniden oluşturulan modeller ile kesinlik, doğruluk ve duyarlılık bakımından karşılaştırılması amaçlanmıştır. Kesinlik, doğruluk ve duyarlılık ölçütleri, her bir gönderilen model için hesaplanmış, ancak bu hesaplamalar sadece duvar unsurları için yapılmıştır. Bu seçim, duvar unsurlarının tüm gönderilen modellerde yeniden oluşturulan tek öğeler olması nedeniyle yapılmış olup, diğer öğelerin (bölmeler, kapılar, pencereler gibi) yüklenen her modelde bulunmadığı belirtilmiştir. Kesinlik ve doğruluk değerleri, 1 cm ila 15 cm arasında değişen tampon boyutları için hesaplanırken, duyarlılıklarda 1 cm ila 15 cm arasında kesme mesafeleri için hesaplanmıştır. Sonuçların karşılaştırılması için, 10 cm'lik bir tampon boyutunda kesinlik ve doğruluk değerleri ile 10 cm'lik bir kesme mesafesinde duyarlılıklar dikkate alınarak, tüm yüklenen veriler için değerlendirme sonuçları sunulmuştur. Genel olarak, yeniden oluşturma yöntemlerinin daha yüksek bir kesinlik ve daha düşük bir doğruluk elde ettiği, duyarlılıkların ise 0.5 cm ila 6.7 cm arasında değiştiği gözlemlenmiştir. Ayrıca, sonuçların iç mekânın karmaşıklığı, dağınıklık düzeyi,

nokta bulutunun doğruluğu ve yoğunluğu gibi faktörlere bağlı olarak değiştiği vurgulanmıştır.

[4], yaptıkları çalışmada az sayıda etiketli örneklerle iç mekân bileşenlerinin sınıflandırılması üzerine odaklanmaktadır. Çalışmanın amacı, büyük ölçekte etiketlenmiş veri ihtiyacını azaltarak, nokta bulutları üzerinden iç mekân bileşenlerini etkili bir şekilde sınıflandırmaktır. Bu amaç doğrultusunda yazarlar, az sayıda etiketli örnekle öğrenme yeteneğine sahip bir yöntem geliştirmişlerdir. İlk olarak birkaç etiketli nokta bulutu etiketlenmemiş nokta bulutları ile karıştırılmıştır. Ardından karışık yüksek boyutlu özellikler derin öğrenme çerçevesi kullanılarak çıkarılmıştır. Daha sonra karışık özellikleri düşük boyutlu bir uzaya gömmek için boyut indirgeme işlemi yapılmıştır. Son olarak her kümedeki birkaç etiketli nokta bulutu tanımlanmış ve aynı kümedeki etiketlenmemiş nokta bulutları için anlamsal etiketler sağlanmıştır. Çalışmada kullanılan yöntem PointNet, PointNet++ ve PointNetCNN yöntemleriyle karşılaştırılmış ve önerilen yöntem ile %87.62 doğruluk elde edilmiştir.

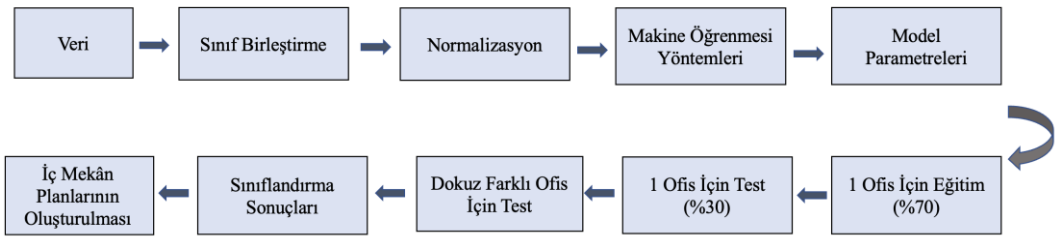
[7], yaptıkları çalışmada iç mekân 3B Lidar nokta bulutlarının anlamsal bölütlemesi için yenilikçi bir yöntem geliştirmişlerdir. Bu çalışmanın amacı, derin öğrenme modellerinin nokta bulutları üzerinde daha yüksek doğrulukla bölütleme yapabilmesi için etkili bir Özellik Piramidi Ağı (Feature Pyramid Network-FPN) modülü tasarlamaktır. Bunun için Sinir Ağı Mimarisi Arama (Neural Architecture Search-NAS) adı verilen bir yöntem kullanılarak, mevcut mimarilerin performansını artıracak yeni bir FPN yapısı keşfedilmiştir. Bu yapı, çok seviyeli özelliklerin daha verimli bir şekilde birleştirilmesini sağlayarak sınıflandırma doğruluğunu arttırmaktadır. Araştırmada önerilen FPN modülü, SFPN olarak adlandırılmış ve farklı seviyelerdeki özelliklerin daha etkin bir şekilde birleştirilmesi için tüm potansiyel iletişim yolları optimize edilmiştir. SFPN, mevcut bölütleme ağlarına eklenebilen, esnek ve güçlü bir modül olarak tasarlanmış olup, farklı mimarilere entegre edildiğinde bölütleme performansını belirgin ölçüde artırmaktadır. Bu yöntemin önemli özelliği, sadece tek bir derin öğrenme modeline özgü olmayıp, çok çeşitli klasik bölütleme ağlarına entegre edilebilir olmasıdır. Bu da SFPN'nin genel bir modül olarak geniş bir uygulama yelpazesinde kullanılabileceğini göstermektedir. Çalışmada kullanılan veri setleri arasında, iç mekân bölütleme araştırmalarında sıklıkla kullanılan iki büyük veri seti olan S3DIS

ve ScanNet v2 bulunmaktadır. Yapılan deneylerde, önerilen SFPN modülü ile entegre edilen klasik bölütleme ağlarının performansında iyileşmeler gözlemlenmiştir. Örneğin, PointNet++-SFPN modeli, ScanNet v2 üzerinde %7.8'lik bir mean Intersection over Union (mIoU) artışı sağlarken, S3DIS veri setinde %4.7'lik bir iyileşme elde etmiştir. Benzer şekilde, PointConv-SFPN modeli, ScanNet v2 üzerinde %4.5 ve S3DIS üzerinde %3.7'lik mIoU iyileşmeleri sağlamıştır. Bu sonuçlar, SFPN modülünün mevcut bölütleme yöntemlerine önemli katkılar sağladığını ve farklı veri setlerinde tutarlı bir performans artışı sunduğunu göstermektedir.

[19], yaptıkları çalışmada 3B nokta bulutu verisinin sınıflandırılması ve bölütlemesi için geliştirilen yeni bir yöntemi tanıtmaktadır. Bu çalışmada, Kendini Artıran Evrişimli Sinir Ağı (Self-augment Convolutional Neural Network- SACNN) yöntemi kullanılarak, nokta bulutlarındaki veri sınırlamalarının üstesinden gelinmeye çalışılmıştır. Modelin yenilikçi SACNN yapısı, özellikle sınırlı miktardaki etiketli veriyi daha verimli kullanarak veri çeşitliliğini artırmayı hedeflemiştir. SACNN'in kendini artıran mekanizması, modelin öğrenme kapasitesini genişletirken, etiketlenmemiş veriler üzerinde veri genişlemesi yaparak modelin genelleme yeteneğini iyileştirmektedir. Çalışmanın temel katkılarından biri, Self-Augment Convolution (SAConv) modülünün tanıtılmasıdır. SAConv modülü, modelin varyansı azaltmak için noktaların konumsal bilgilerini hizalamakta ve böylece daha tutarlı ve güçlü özellikler elde etmektedir. Bu yaklaşım, özellikle iç mekânlarda nokta bulutu verisinin karmaşıklığı ile başa çıkmak adına büyük avantaj sağlamaktadır. Çalışmada ayrıca, Yerel Karışık Birleştirme Modülü (Local Mixed Aggregation Module) önerilmiştir. Bu modül, noktaların komşu ilişkilerini daha iyi analiz ederek hem geniş ölçekli genel bakış hem de ince ayrıntılı betimleyici özellikleri birleştirmeyi amaçlamaktadır. Böylece, modelin daha zengin bir özellik kümesi öğrenmesi sağlanarak sınıflandırma ve bölütleme performansı artırılmaktadır. Çalışmada sınıflandırma işlemi, ModelNet40 ve ScanObjectNN veri kümeleri üzerinde gerçekleştirilmiştir. Önerilen yöntem ile, ModelNet40 veri setinde %93.2 kesinlik, ScanObjectNN veri setinde ise %80.4 doğruluk elde edilmiştir.

3
MATERYAL VE YÖNTEM

Bu çalışma, veri işleme aşamalarıyla başlayıp modelleme ve değerlendirme süreçleriyle devam eden bir yaklaşım sunmaktadır. İlk aşamada veri düzenleme, sınıfların birleştirilmesi ve normalizasyon işlemleri yapılmış; ardından Rastgele Orman, Çok Katmanlı Algılayıcı, CatBoost, XGBoost, LightGBM ve CART gibi ağ modelleri geliştirilmiştir. Bu süreçte model parametreleri belirlenmiş, bir ofis için %70 eğitim ve %30 test setleri oluşturularak modeller eğitilmiştir. Son aşamada, dokuz farklı ofise ait nokta bulutu sınıflandırma sonuçları doğruluk metrikleri ile değerlendirilmiş ve bu sonuçlar planlara dönüştürülmüştür. Çalışmanın iş akış şeması Şekil 3.1' de gösterilmiştir.



Şekil 3.1 İç mekân nokta bulutu verisinin makine öğrenmesi yöntemleriyle sınıflandırılması ve görselleştirilmesi iş akış şeması

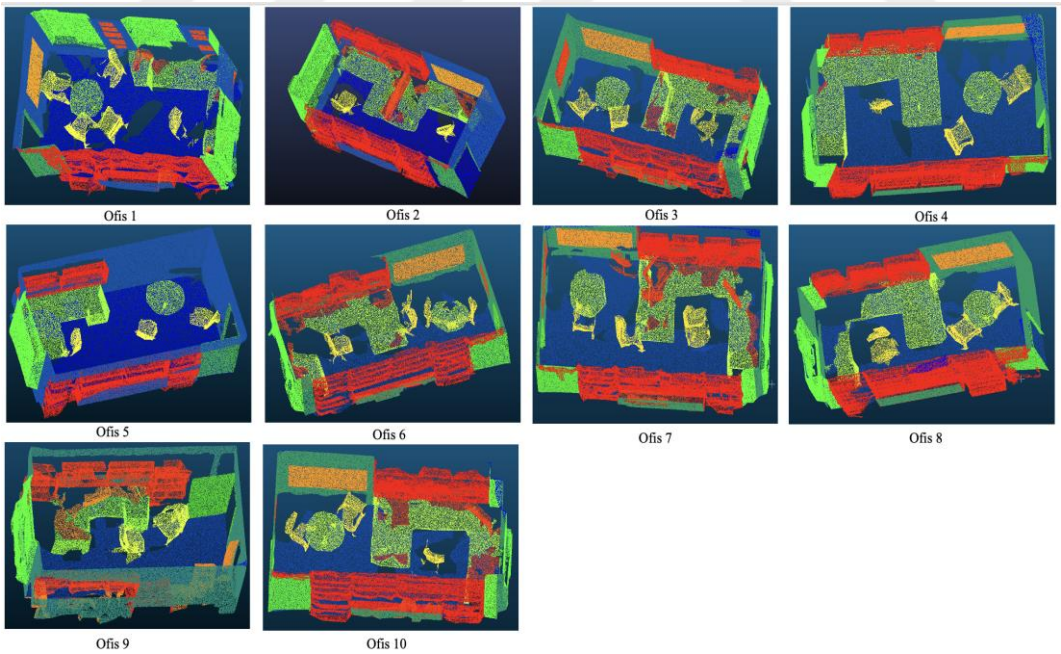
3.1 Çalışmada Kullanılan Veri Seti

Bu çalışmada, S3DIS veri seti kullanılmıştır. Veri seti, her bir tarama konumunda, 360°'lik dönüş boyunca farklı açılardan üç yapılandırılmış ışık sensörünü birleştirerek 18 RGB ve derinlik görüntüsü yakalayabilen Matterport Kamerası ile oluşturulmuştur. Tarama yapılan alanlar, renkli nokta bulutları şeklinde kaydedilmiş olup, veri seti üç farklı binadan toplanan altı büyük ölçekli iç mekân taramasını içermektedir. Bu binalar, ağırlıklı olarak eğitim ve ofis ortamlarında

kullanılmaktadır. Veri setinde toplamda 70.496 düzenli RGB ve 1.413 eş dikdörtgen RGB görüntüsü yer almaktadır. Tüm görüntüler 1080x1080 çözünürlüğe sahiptir.

S3DIS veri setinde, 13 sınıf bulunmaktadır. Bu sınıflar tavan, zemin, duvar, kiriş, sütun, pencere, kapı, masa, sandalye, kanepe, kitaplık, pano, düzensiz eşyalar(dağınıklık) nesnelerini kapsamaktadır. Nokta bulutu verisi, X, Y, Z koordinatlarının yanı sıra R, G, B renk bilgisi ve uzmanlar tarafından manuel olarak etiketlenen referans (ground truth) etiketlerini içermektedir [13].

Bu çalışmada, S3DIS veri setindeki ofis kullanımına yönelik bölümler, makine öğrenme yöntemleri ile sınıflandırma amacıyla eğitim ve test verisi olarak ayrılmıştır. Örneğin, Ofis 1 verisi, %70 eğitim ve %30 test olacak şekilde bölünmüştür. Bu oranların öğrenme performansı için uygun olduğu deneysel olarak gözlemlenmiştir. Ayrıca, dokuz ayrı ofis verisi üzerinde elde edilen test sonuçları, modelin aşırı uyuma (overfitting) neden olmadığını ortaya koymaktadır. Kullanılan ofis odalarının iç kısımlarının CloudCompare yazılımı üzerindeki görünümü Şekil 3.2'de gösterilmiştir.



Şekil 3.2 Kullanılan ofis verilerinin iç görünümü

3.2 Çalışmada Kullanılan Programlar

3.2.1 Anaconda Navigator Programı

Bu çalışmada, veri işleme, model eğitimi ve performans analizi için Python programlama dili, Anaconda Navigator platformu üzerinde kullanılmıştır. Çalışmada kullanılan S3DIS veri seti, ofis odalarına ait .txt uzantılı verilerden oluşmaktadır. Verilere ilk olarak min-max normalizasyon işlemi uygulanmış ve bu işlem için MinMaxScaler kullanılmıştır. Normalizasyon sonrasında, sınıflandırma için veri setinde yer alan sınıfların homojen olmaması nedeniyle bazı sınıflar birleştirilmiştir. Başlangıçta veri seti 11 sınıftan oluşmaktadır. Eğitim ve test verilerinde, iç mekân planlarının oluşturulmasını kolaylaştırmak amacıyla, duvar, kapı, pencere, kolon, pano ve kitaplık bir sınıf (birleştirilmiş sınıf-1); masa, sandalye ve dağınıklık ise başka bir sınıf (birleştirilmiş sınıf-2) olarak birleştirilmiştir. Bu sınıflandırma yöntemi, modellerin performansını artırmak ve veri setindeki sınıflar arasındaki dengesizlik minimize edilmiştir. Veri sınıflandırması için çeşitli modeller oluşturulmuş ve veriler bu modellerle eğitilerek test edilmiştir. Kullanılan algoritmalar arasında Scikit-learn kütüphanesindeki RandomForestClassifier, Multi-Layer Perceptron (MLP) ve Classification and Regression Trees (CART); CatBoost kütüphanesindeki CatBoostClassifier; XGBoost kütüphanesindeki XGBoostClassifier ve LightGBM kütüphanesindeki LightGBMClassifier bulunmaktadır. Veri manipülasyonu ve analizi için NumPy ve Pandas kütüphaneleri kullanılmıştır. Model performansını değerlendirmek amacıyla kesinlik, duyarlılık ve F1 skor metrikleri hesaplanmıştır. Eğitim sürecinde elde edilen modellerin kaydedilmesi ve gelecekte yeniden kullanılabilmesi için Joblib kütüphanesinden yararlanılmıştır. Tüm bu adımlar sonucunda, veri seti işlenmiş, farklı algoritmalarla modeller eğitilmiş ve test edilmiştir.

3.2.2 CloudCompare Programı

Bu çalışmada, sınıflandırma algoritmaları ile elde edilen tahmin sonuçları Python programlama dili kullanılarak işlenmiş ve tahmin dosyaları oluşturulmuştur. Hazırlanan tahmin dosyaları, her bir noktanın referans sınıflarını (gerçek sınıflar), tahmin edilen sınıflarını ve doğru/yanlış sınıflandırma durumlarını içermektedir. Dosyada, nokta bulutu sınıflarını analiz etmek amacıyla referans sınıfları, tahmin edilen sınıflar ve bu sınıflandırmaya ilişkin doğru pozitif (DP), doğru negatif (DN),

yanlış pozitif (YP) ve yanlış negatif (YN) bilgileri yer almıştır. Bu tahmin dosyası, sınıflandırma sonuçlarının doğruluğunu analiz etmek ve bu doğruluğu görsel olarak ifade edebilmek için CloudCompare yazılımına aktarılmıştır. CloudCompare ortamında, sınıflandırma sonuçları nokta bulutunda renklerle görselleştirilmiştir. Bu yaklaşım, sınıflandırma performansını görsel olarak değerlendirme imkanı sunmuş ve doğru ile yanlış sınıflandırmaların yoğunlaştığı bölgelerin kolaylıkla tespit edilmesini sağlamıştır. Elde edilen görselleştirmeler, sınıflandırma algoritmasının ofis içi veriler üzerindeki performansını detaylı bir şekilde analiz etmeye olanak tanımış ve bu bulgular tez kapsamında detaylı bir şekilde tartışılmıştır.

3.2.3 ArcGIS Pro Programı

Nokta bulutunun sınıflandırılmasının ardından elde edilen tahmin sonuçları ArcGIS Pro yazılımı kullanılarak yarı otomatik bir şekilde planları hazırlanmıştır. İzlenen işlem adımları aşağıdaki gibi sıralanmıştır:

3.2.3.1 Tahmin Dosyasının Eklenmesi

Tahmini içeren her ofis verisi, .txt formatındaki dosya kullanılarak, Map modülünden Add Data > XY Point Data seçeneğiyle noktalar eklenmiştir.

3.2.3.2 Etiket Bazlı Verilerin Ayrılması ve Dışa Aktarılması

Eklenen ofis dosyası, Map modülünde bulunan Select by Attributes seçeneği kullanılarak işlenmiştir. Bu işlem, her bir sınıf için (tavan / 0 etiketi, zemin / 1 etiketi, birleştirilmiş sınıf-1 / 2 etiketi, birleştirilmiş sınıf-2 / 3 etiketi) tahmin edilen ve gerçek etiketleri ayrı ayrı seçmek amacıyla gerçekleştirilmiştir. Seçilen veriler, ardından Export Features aracı ile dışa aktarılmıştır. Bu işlem sırasında, her bir sınıf için uygun dosya isimlendirmeleri yapılmış ve böylece her sınıfın tahmin ve gerçek etiket verilerini içeren ayrı katmanlar oluşturulmuştur. Bu sayede, her sınıfın verileri bağımsız olarak analiz edilebilmiş ve görselleştirilebilmiştir.

3.2.3.3 Plan Oluşturma

Çizim sürecinde, tahmin verileri kullanılarak her sınıfın geometrik sınırları oluşturulmuştur. Sınırların oluşturulmasında ise aşağıdaki adımlar takip edilmiştir:

Tavan için;

Analysis>Tools>Geoprocessing sekmesinden Minimum Bounding Geometry aracı seçilerek ve Geometry Type olarak Convex Hull belirlenerek tavan sınırları otomatik olarak oluşturulmuştur. Oluşan sınırların köşeleri Add vertex kullanılarak simetrik hale getirilmiştir.

Masa için;

Create Features>Circle seçeneği kullanılarak masa sınırları manuel olarak çizilmiştir.

Sandalye için;

Create Features>Polygon sekmesi seçilerek sandalye sınırları manuel olarak çizilmiştir.

Kitaplık için;

Create Features > Polygon sekmesi seçilerek manuel olarak çizilmiştir.

Kapı ve pencereler için;

Kapı ve pencere sınırları, Create Features > Line sekmesi seçilerek manuel olarak çizilmiştir.

Bu adımlar sonucunda, sınıflandırılan verilerin tahminleri ve gerçek değerleri temel alınarak planlar oluşturulmuş ve her bir sınıf için lejantlar eklenerek planda gösterilmiştir.

3.3 Min-Max Normalizasyonu

Min-max normalizasyon tekniği, değişkenlerin ortalama ve varyans değerleri arasında belirgin farklar olduğunda tercih edilir ve verileri belirli bir aralığa ölçeklendiren doğrusal bir dönüşüm uygular [20]. Bu yöntem, verileri normalleştirmek için en yaygın kullanılan tekniklerden biridir. Bu teknikte, bir özelliğin değerleri belirli bir aralığa, genellikle [0–1] arasına ölçeklendirilir. Min-max normalizasyonu, veriler arasındaki mevcut ilişkileri koruyarak dönüşümü gerçekleştirir [21]. Bu çalışmada, X, Y, Z konum koordinatları ile R, G, B renk değerleri, min-max normalizasyon yöntemi kullanılarak normalize edilmiştir.

$$X_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{3.1}$$

3.4 Kullanılan Makine Öğrenimi Yöntemleri

Makine öğrenmesi ve derin öğrenme kavramları uzun bir geçmişe sahiptir. Ancak son yıllarda, büyük veriye erişim ve hesaplama gücündeki ilerlemelerle birlikte bu teknolojilerde önemli bir dönüşüm yaşanmıştır [22]. Makine öğrenimi, verilerden bilgi analiz etme ve çıkarma konusunda temel bir rol oynamaktadır [23]. Makine öğrenmesi, verilerin daha verimli bir şekilde işlenmesini sağlamak amacıyla kullanılan bir yöntemdir. Veriler incelendiğinde, bazen semantik bilgilerin doğrudan çıkarılması mümkün olmayabilir. Bu tür durumlarda, makine öğrenimi teknikleri devreye girerek verilerden anlamlı bilgilerin elde edilmesini sağlar. Günümüzde veri setlerinin hızla artmasıyla birlikte makine öğrenimine olan talep de yükselmektedir. Birçok sektör, büyük veri setlerinden semantik bilgiler çıkarmak ve karar destek sistemleri oluşturmak amacıyla makine öğrenimi yöntemlerinden yararlanmaktadır. Makine öğreniminin temel amacı, verilerden bilgiler öğrenerek tahminleme ve sınıflandırma gibi görevleri otomatikleştirmektir [24].

Klasik makine öğrenimi, bir algoritmanın daha doğru tahminler yapmayı nasıl öğrendiğine bağlı olarak sınıflandırılabilir. Bu bağlamda, dört temel öğrenme yöntemi bulunmaktadır [25]:

- Denetimli Öğrenme
- Denetimsiz Öğrenme
- Yarı Denetimli Öğrenme
- Pekiştirmeli Öğrenme

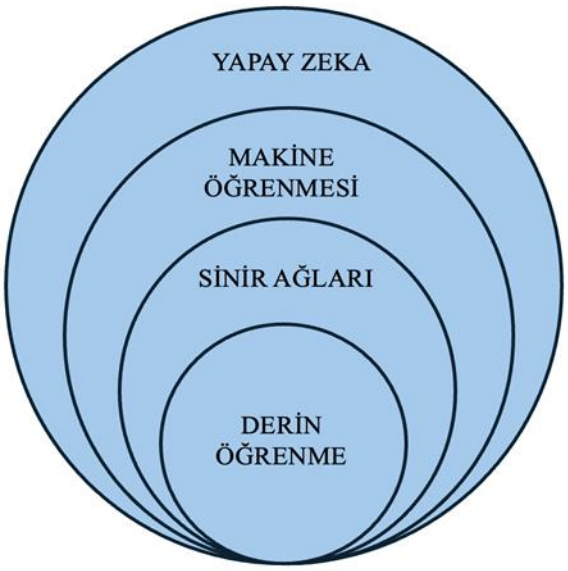
Denetimli öğrenme, algoritmanın etiketli veriler kullanılarak eğitildiği bir makine öğrenimi yöntemidir. Etiketli veriler, her bir girdinin karşılık geldiği doğru çıktıyla birlikte bulunduğu veri çiftlerinden oluşur. Öğrenme süreci boyunca algoritma, tahminlerini sürekli olarak geliştirerek, ürettiği çıktı ile gerçek değer arasındaki farkı en aza indirmeye çalışır.

Denetimsiz öğrenme, denetimli öğrenmeden farklı olarak, öğrenme algoritmasını etiketli çıktılar olmadan yalnızca giriş verileriyle eğitir ve etiketlenmemiş verilerden özellikler çıkarmak, bunları kategorize etmek veya etiketlemek için kullanılır.

Yarı denetimli öğrenme, büyük miktarda giriş verisi kullanır; bu verilerin bir kısmı etiketli iken geri kalanı etiketsizdir ve denetimli ile denetimsiz öğrenme arasında bir yerde konumlanır. Yarı denetimli öğrenmede, öğrenme algoritması hem denetimli hem de denetimsiz eğitim yaklaşımlarıyla eğitilebilir [26].

Pekiştirmeli öğrenme, bir makinenin belirli bir hedefe ulaşmak için nasıl davranması gerektiğini öğrenen bir makine öğrenme yöntemidir. Bu sistem, karşılaştığı durumlara tepki verir ve buna karşılık sayısal bir ödül alır. Pekiştirmeli öğrenmenin en belirgin özelliği olan deneme-yanılma yöntemiyle sistem, aldığı ödülü maksimize etmeyi hedefler [27].

Bu çalışmada, Rastgele Orman, Çok Katmanlı Algılayıcı, CatBoost, XGBoost, LightGBM ve CART gibi makine öğrenmesi yöntemleri kullanılmıştır. Bu yöntemler, sınıflandırma ve regresyon gibi temel makine öğrenimi görevlerini yerine getirirken, aynı zamanda büyük veri setlerinde yüksek performans ve verimlilik sağlamaktadır. Yapay zekâ yöntemlerinin birbiriyle olan ilişkisi Şekil 3.3' te verilmiştir.



Şekil 3.3 Yapay zekâ yöntemlerinin birbiriyle ilişkisi [28]

3.4.1 Rastgele Orman (RO) Yöntemi

Rastgele Orman (RO) algoritması, Breiman (2001) tarafından geliştirilmiş olup, karar ağaçlarını belirli bir kombinasyonla bir araya getiren bir topluluk öğrenme yöntemi olarak literatüre kazandırılmıştır [29]. Rastgele Orman algoritması, basit ve etkili bir topluluk öğrenme yöntemidir. Bu algoritmanın kullanımı oldukça

kolaydır, çünkü her düğümde rastgele seçilen değişkenlerin sayısı ve ormandaki ağaç sayısı olmak üzere sadece iki parametre gereklidir. RO, bir dizi karar ağacının oluşturulmasından oluşur; bu ağaçlar, rastgele seçilen eğitim kümesinin alt kümeleri üzerinde eğitilir. Test nesnesinin sınıfı, bu ağaçlardan alınan oyların birleştirilmesiyle belirlenir. Böylece, her bir karar ağacının katkısı, doğru sınıflandırma için nihai bir karar verir [30]. Bu süreçte elde edilen hata oranı genelleştirilmiş hata olarak tanımlanır ve hesaplama yöntemi (3.2) eşitliğinde verilmiştir.

$$PE^* = P_{X,Y}(mg(X,Y) < 0) \tag{3.2}$$

Burada mg(), margin fonksiyonu olarak tanımlanır. Margin, (X, Y) çiftinde doğru sınıfın aldığı ortalama oy sayısının, diğer sınıfların ortalama oylarından ne kadar fazla olduğunu gösterir. Margin değeri büyüdükçe, sınıflandırmanın doğruluğu ve güvenilirliği artar [31].

3.4.2 Çok Katmanlı Algılayıcı (ÇKA) Yöntemi

Çok Katmanlı Algılayıcı, bir veri kümesi üzerinde eğitim yaparak karmaşık ilişkileri öğrenebilen denetimli bir makine öğrenmesi yöntemidir [32]. ÇKA sinir ağı, yapay nöronlardan oluşan bir yapıya sahiptir ve ağın temel yapı taşlarını bu nöronlar oluşturur. Her nöron, girişlerini ve bias değerini ağırlıklandıran bir aktivasyon fonksiyonuna sahip hesaplama birimi olarak çalışır [33]. ÇKA, bilginin giriş katmanından çıkış katmanına doğru, gizli katmanlar aracılığıyla tek yönlü olarak aktığı, katmanlı ileri beslemeli geri yayılımlı bir sinir ağı modelidir [34]. ÇKA'nın giriş katmanında veriler alınır ve toplama fonksiyonu kullanılarak aktivasyon fonksiyonuna iletilir. Her bir katmanda, bu aktivasyon fonksiyonları kullanılarak hesaplamalar yapılır ve bu hesaplamalar, bir sonraki katmana iletilir. Bu süreç, çıkış katmanına ulaşana kadar tekrarlanır. Çıkış katmanında hesaplanan hatalar, geri yayılım algoritmaları ile ağırlıkların güncellenmesinde kullanılır ya da nihai karar verilir. ÇKA modelinin performansı, eğitim verilerinin sayısına, değişken seçimine, gizli katman ve düğüm sayısına, öğrenme oranı ve momentum katsayısı gibi eğitim parametrelerine bağlıdır. Bu parametreler, öğrenme sürecinin hızını ve etkinliğini kontrol eder, bu nedenle dikkatli bir şekilde ayarlanması,

modelin başarı performansını doğrudan etkiler [35]. Net girdi değeri I_j (3.3) eşitliğinde gösterilmiştir.

$$I_j = \sum_i W_{ij}\, O_i + Q_j \tag{3.3}$$

Burada I_j j düğümünün çıktısını; W_{ij} I ve j düğümü arasındaki ağırlığı, O_i bir önceki katmandan gelen i düğümünün çıktısını ve Q_j j düğümünün eşik değerini ifade eder. Çıkış değeri O_j (3.4) eşitliğinde gösterilmiştir.

$$O_j = \frac{1 - e^{-2l_j}}{1 + e^{-2l_j}} \tag{3.4}$$

Her gizli ve çıktı katmanındaki birim, net girdisine bir aktivasyon fonksiyonu uygular. Bu çalışmada, tangent-sigmoid (tanh) aktivasyon fonksiyonu kullanılmıştır. Bu fonksiyon, net girdiyi [–1, 1] aralığına dönüştürür [36].

3.4.3 CatBoost Yöntemi

CatBoost sınıflandırıcısı, veri hazırlığı açısından oldukça basittir ve sayısal değişkenlerdeki eksik değerlerle birlikte kodlanmamış kategorik değişkenleri otomatik olarak işler [37].CatBoost sınıflandırıcısı, temelinde gradyan artırma ile güçlendirilmiş karar ağaçları bulunan bir algoritmadır. Eğitim sürecinde, ardışık karar ağaçları oluşturulur ve her yeni ağaç, bir önceki ağacın yaptığı hataları düzeltmeye odaklanır. Bu şekilde, her bir ağaç sırasıyla kayıpları azaltarak daha doğru sonuçlar elde etmek üzere modellenir [38].

Verilen n örnek ve $D = \{(X_i, Y_i)_{i=1,2,\ldots\ldots,n}\}$ boyutlu bir eğitim seti için $X_i = (x_i^1, x_i^2, \ldots., x_i^m)$ m boyutlu giriş özelliklerini ve Yi ∈ R etiketli değerleri gösterir. Eğitimden sonra oluşturulan güçlü öğrenici F_{k-1}'dir ve bir sonraki turdaki eğitim hedefi, kayıp fonksiyonu L (·)'nin beklentisi E (·)'yi minimize etmek için CART karar ağacı kümesi T' den bir ağaç t_k elde etmektir. Parametre t_k (3.5) eşitliğinde olduğu gibi hesaplanır:

$$t_k = argminEL\big(y, F_{k-1}(x) + t(x)\big) \tag{3.5}$$

Burada (x, y) eğitim setinden bağımsız test örnekleridir. GBDT (Gradient Boosting Decision Trees), eğitilmiş CART karar ağacını t_k'ye uydurmak için kayıp fonksiyonunun negatif gradyanını kullanır ve eşitlik 3.6'da gösterilen nihai model M, başlangıçtaki zayıf öğrenici M_0ve N iterasyondan sonra eğitim adımı boyutu a_n'nin n'inci turundan elde edilir:

$$M = M_0 + \sum_{k=1}^{N} a_n t_k \tag{3.6}$$

CatBoost, geleneksel Gradient Boosting Decision Trees (GBDT) yöntemine dayanan çeşitli iyileştirmeler sunarak diğer artırma algoritmalarından ayrılır. İlk olarak, eğitim setindeki gürültüye karşı daha dayanıklı olan sıralı artırma (order boosting) tekniğini kullanır. Ayrıca, kategorik özellikleri otomatik olarak sayısal verilere dönüştüren Sıralı Hedef İstatistikleri (Ordered Target Statistics) yöntemini tanıtarak bu tür verilere doğrudan destek sağlar. CatBoost, kategorik özelliklerin boyutunu önemli ölçüde artırarak daha verimli modeller oluşturur. Ayrıca, her bir katmanda aynı bölme kriterini uygulayan simetrik ağaç yapısıyla istikrarı ve tahmin hızını artırmayı başarır [39].

3.4.4 XGBoost Yöntemi

XGBoost (Aşırı Gradyan Artırma), 2016 yılında Tianqi Chen ve Carlos Guestrin tarafından geliştirilen, karar ağaçları temelli yenilikçi bir makine öğrenmesi algoritmasıdır. Gradyan artırma algoritmasının optimize edilmiş bir versiyonu olan XGBoost, yüksek tahmin gücü ve hızlı işlem kabiliyeti sayesinde veri bilimciler arasında popülerlik kazanmış ve özellikle Kaggle gibi yarışmalarda sıklıkla tercih edilmiştir [40]. XGBoost'un en önemli avantajlarından biri, gürültü ve aykırı değerleri etkili bir şekilde yöneterek modelin kararlılığını ve güvenilirliğini artırmasıdır [41]. Bu modelin matematiksel ifadesi (3.7) eşitliğindeki gibidir:

$$y_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F \tag{3.7}$$

Bu denklemde, K ağaçların sayısını, f fonksiyonu fonksiyonel uzaydaki bir fonksiyonu ve F tüm olası ağaçların bulunduğu fonksiyon uzayını temsil eder.

Algoritmanın her aşamasında, ormanın öğrenme sürecinin ne kadar ilerlediği değerlendirilir ve her defasında bir ağaç eklenir. Eklenen yeni ağaçlar, hedef fonksiyonu en iyi şekilde optimize edecek şekilde seçilir. Yani, her adımda, algoritma mevcut ağaçların katkılarını göz önünde bulundurarak, öğrenim sürecini iyileştirecek yeni bir ağaç ekleyerek devam eder. Adım t için belirlenen spesifik hedef ise (3.8) eşitliğindeki gibi hesaplanır:

$$\sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{3.8}$$

Bu denklemde, g_i ve h_i, kayıp fonksiyonu üzerinde sırasıyla birinci ve ikinci dereceden gradyan istatistiklerini temsil eder. Ω, modelin karmaşıklığını (yani, regresyon ağacı fonksiyonlarını) düzenleyen bir terimdir. Bu terim, yeni ağacın optimizasyon hedefi haline gelir. Denkleme dönüşümler uygulandıktan sonra, t. ağaç için amaç fonksiyonu aşağıdaki (3.9) ve (3.10) eşitliğindeki gibi hesaplanır:

$$obj^{(t)} = \sum_{j=1}^{T}[G_j \omega_J + \frac{1}{2}(H_j + \lambda)\omega_j^2] + \gamma T, \tag{3.9}$$

$$I_J = \big(i \setminus q(x_i)\big) = j, G_J = \sum_{\mathrm{i}\epsilon I_j} g_i\,, H_j = \sum_{i\epsilon I_j} h_i\,, \tag{3.10}$$

Bu formülde; I_j, j'inci yaprağa atanan veri noktalarının indekslerinin kümesi olarak tanımlanmıştır. Burada T, yaprak sayısını; ω yapraklardaki skorların vektörünü ifade eder. Ayrıca, g_i ve h_i sırasıyla birinci ve ikinci dereceden gradyan istatistiklerinin toplamlarını ifade etmektedir [42].

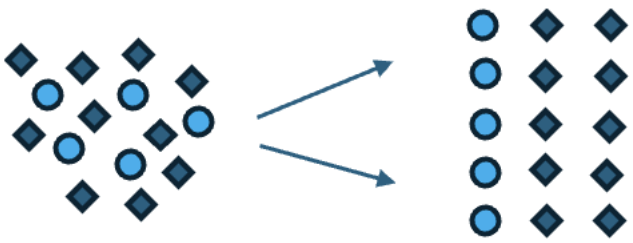
3.4.5 LightGBM Yöntemi

Microsoft, 2017 yılında gradyan artırma karar ağaçları (Gradient Boosting Decision Trees - GBDT) algoritmasına dayalı olarak geliştirilen LightGBM modelini tanıtmıştır [43]. LightGBM, gradyan artırmalı makinelerde hesaplama verimliliğini artırmak amacıyla özellikle büyük veri setleriyle çalışmak için tasarlanmıştır [44]. LightGBM, diğer ağaç tabanlı makine öğrenmesi algoritmalarına benzer

şekilde çalışır, ancak farkı, ağacın yatay yerine dikey olarak büyümesidir. Bu özellik, LightGBM'in ağaç yapısını seviye bazında büyütmesiyle, diğer algoritmaların yaprak bazında büyütmesinden farklılık gösterir. Bu sayede, LightGBM büyük veri setleriyle etkin bir şekilde başa çıkabilir ve daha az bellek kullanarak sonuçların doğruluğuna odaklanır [45]. Ayrıca, LightGBM'in gradyan artırma karar ağaçları üzerindeki etkinliği, geleneksel GBDT yöntemlerine kıyasla eğitim süresini hızlandırarak benzer doğruluk seviyeleri sağladığı yapılan araştırmalarda görülmüştür [46].

3.4.6 CART Yöntemi

CART modeli, veri sınıflandırması için en yaygın kullanılan algoritmalardan biridir. Eksik verilere karşı dayanıklıdır ve değişkenlerin normal dağılıma sahip olması gerekmez [47]. CART algoritması, veri setindeki değişkenleri inceleyerek veriyi daha homojen alt gruplara ayırmayı amaçlar. Bu süreç, her düğümde, veri setini en uygun şekilde bölecek değişken ve eşik değerinin seçilmesiyle gerçekleştirilir. Bölünme işlemi, veri setindeki heterojenliği minimuma indirgemeyi hedefler ve böylece oluşan yaprak düğümlerde daha homojen gruplar elde edilir. Özellikle büyük veri setleri ve karmaşık ilişkiler içeren problemlerde etkili bir çözüm sunar [48]. CART algoritması, parametrik olmayan bir yöntem olup, herhangi bir fonksiyonel formun önceden belirlenmesini gerektirmez. Bu özelliği sayesinde, veri dağılımına ilişkin herhangi bir varsayımda bulunmadan çalışabilir. Ayrıca, değişkenlerin seçim sürecini otomatik olarak gerçekleştirerek, en anlamlı değişkenleri belirler ve amaç bakımından önemsiz olanları elimine eder. Bu yönüyle, modelin öğrenme sürecinde gereksiz değişkenlerin etkisini ortadan kaldırarak daha sade ve etkili bir karar ağacı oluşturulmasını sağlar [49].



Şekil 3.4 CART algoritması sınıflandırma diyagramı [47]

3.5 Sınıflandırma için Performans Değerlendirme Ölçütleri

Oluşturulan modellerin eğitimi tamamlandıktan sonra performans değerlendirmesi yapmak için sınıflandırma problemlerinde sık kullanılan kesinlik, duyarlılık, doğruluk ve F1 skor metrikleri kullanılmıştır. Bu metrikler, modelin tahminlerinin kalitesini değerlendirmek için genellikle karışıklık matrisinden türetilmektedir

Tablo 3.1 Karışıklık Matrisi [50]

		Referans	
		Pozitif	Negatif
Tahmin	Pozitif	DP	YP
	Negatif	YN	DN

Karışıklık matrisi, gerçek ve tahmin edilen sınıf etiketlerini karşılaştıran bir tablo olarak, modelin sınıflandırma performansını detaylı bir şekilde analiz etmeye olanak tanır. İkili sınıflandırma problemleri için dört temel bileşeni bulunur: Doğru Pozitif (True Positive, TP), modelin pozitif olarak tahmin ettiği ve gerçekten pozitif olan örnekleri ifade eder. Doğru Negatif (True Negative, TN), modelin negatif olarak tahmin ettiği ve gerçekten negatif olan örneklerdir. Yanlış Pozitif (False Positive, FP), modelin pozitif olarak tahmin ettiği ancak gerçekte negatif olan örnekleri temsil ederken, Yanlış Negatif (False Negative, FN) ise modelin negatif olarak tahmin ettiği ancak gerçekte pozitif olan örnekleri ifade eder. Bu bileşenler, modelin hatalarını ve doğruluk oranlarını detaylı bir şekilde değerlendirmek için kullanılır [51].

$$\text{Kesinlik} = \frac{DP+DN}{DP+YP} \tag{3.11}$$

$$\text{Duyarlılık} = \frac{DP}{DP+YN} \tag{3.12}$$

$$\text{Doğruluk} = \frac{DP+DN}{DP+YP+DN+YN} \tag{3.13}$$

$$\text{F1 Skoru} = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik+Duyarlılık} \tag{3.14}$$

4
UYGULAMA VE BULGULAR

İç mekânlarda elde edilen nokta bulutu verileri 11 farklı sınıf altında sınıflandırılmıştır. Bu sınıflar; tavan, zemin, duvar, kapı, pencere, kolon, masa, sandalye, pano, dağınıklık ve kitaplıktan oluşmaktadır. Hem eğitim hem de test verilerinde bu sınıflandırma, mekânsal özelliklerin belirlenmesi ve iç mekân planlarının oluşturulması amacıyla kullanılmıştır. Sınıflandırma sürecinde, benzer özelliklere sahip sınıflar birleştirilmiş ve daha etkin bir analiz yapılabilmesi için gruplandırılmıştır. Bu bağlamda, duvar, kapı, pencere, kolon, pano ve kitaplık birleştirilmiş sınıf-1 olarak adlandırılmıştır. Benzer şekilde, masa, sandalye ve dağınıklık gibi nesneler birleştirilmiş sınıf-2 olarak gruplanmıştır. Bu sınıflar, Python programlama dili kullanılarak kodlanmış ve her bir sınıfa belirli etiketler atanmıştır: tavan için 0, zemin için 1, birleştirilmiş sınıf-1 için 2 ve birleştirilmiş sınıf-2 için 3 etiketi verilmiştir.

Aynı zamanda bu çalışmada Rastgele Orman ve Çok Katmanlı Algılayıcı modellerinde, literatürde sıklıkla tercih edilen ve yaygın olarak kullanılan varsayılan (default) parametreler kullanılmıştır.

CatBoost, LightGBM, XGBoost ve CART algoritmaları için hiperparametre ayarlaması yapılmıştır. Bu modellerin sınıflandırma başarımını artırmak ve veri setine özgü en uygun yapılandırmayı elde etmek amacıyla çeşitli parametre kombinasyonları denenmiş, en iyi sonuçları veren yapılandırmalar deneysel olarak belirlenmiştir.

4.1 Çalışmada Elde Edilen Sınıflandırma Sonuçları

Tablo 4.1 Ofis 2 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,98	0,96	0,98	0,98	0,98	0,98
Duyarlılık-0	0,99	0,99	1,00	0,99	0,99	1,00
F1 Skor-0	0,99	0,97	0,99	0,99	0,98	0,99
Kesinlik-1	1,00	1,00	1,00	0,99	1,00	1,00
Duyarlılık-1	0,58	0,49	0,77	0,46	0,40	0,43
F1 Skor-1	0,73	0,66	0,87	0,63	0,57	0,61
Kesinlik-2	0,89	0,90	0,92	0,92	0,87	0,80
Duyarlılık-2	0,89	0,83	0,89	0,85	0,89	0,85
F1 Skor-2	0,89	0,86	0,90	0,88	0,88	0,82
Kesinlik-3	0,42	0,39	0,48	0,39	0,35	0,37
Duyarlılık-3	0,60	0,72	0,65	0,72	0,52	0,47
F1 Skor-3	0,49	0,50	0,55	0,51	0,42	0,41
Genel Doğruluk	0,83	0,80	0,86	0,81	0,80	0,77

Ofis 2 için RO yöntemiyle %83, ÇKA yöntemiyle %80, CatBoost yöntemiyle %86, XGBoost yöntemiyle %81, LightGBM yöntemiyle %80, CART yöntemiyle %77 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %97, CatBoost yöntemiyle %99, XGBoost yöntemiyle %99, LightGBM yöntemiyle %98, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %73, ÇKA yöntemiyle %66,CatBoost yöntemiyle %87, XGBoost yöntemiyle %63, LightGBM yöntemiyle %57, CART yöntemiyle %61; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %89, ÇKA yöntemiyle %86, CatBoost yöntemiyle %90, XGBoost yöntemiyle %88, LightGBM yöntemiyle %88, CART yöntemiyle %82; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %49, ÇKA yöntemiyle %50, CatBoost yöntemiyle %55, XGBoost yöntemiyle %51, LightGBM yöntemiyle %42, CART yöntemiyle %41 F1 doğruluğu elde edilmiştir.

Tablo 4.2 Ofis 3 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,99	0,98	0,99	0,99	0,99	0,99
Duyarlılık-0	0,99	0,93	1,00	0,99	0,99	1,00
F1 Skor-0	0,99	0,96	0,99	0,99	0,99	0,99
Kesinlik-1	0,97	0,98	0,97	0,97	0,98	0,97
Duyarlılık-1	1,00	1,00	1,00	1,00	1,00	1,00
F1 Skor-1	0,99	0,99	0,99	0,99	0,99	0,99
Kesinlik-2	0,88	0,86	0,88	0,88	0,89	0,88
Duyarlılık-2	0,96	0,89	0,98	0,92	0,97	0,93
F1 Skor-2	0,92	0,88	0,93	0,90	0,93	0,91
Kesinlik-3	0,87	0,69	0,94	0,75	0,92	0,79
Duyarlılık-3	0,64	0,63	0,61	0,63	0,64	0,63
F1 Skor-3	0,74	0,66	0,74	0,69	0,75	0,70
Genel Doğruluk	0,91	0,87	0,92	0,89	0,92	0,90

Ofis 3 için RO yöntemiyle %91, ÇKA yöntemiyle %87, CatBoost yöntemiyle %92, XGBoost yöntemiyle %89, LightGBM yöntemiyle %92, CART yöntemiyle %90 genel doğruluk elde edilmiştir.Tavan (0 etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %96, CatBoost yöntemiyle %99, XGBoost yöntemiyle %99, LightGBM yöntemiyle %99, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %99,CatBoost yöntemiyle %99, XGBoost yöntemiyle %99, LightGBM yöntemiyle %99, CART yöntemiyle %99; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %92, ÇKA yöntemiyle %88, CatBoost yöntemiyle %93, XGBoost yöntemiyle %90, LightGBM yöntemiyle %93, CART yöntemiyle %91; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %74, ÇKA yöntemiyle %66, CatBoost yöntemiyle %74, XGBoost yöntemiyle %69, LightGBM yöntemiyle %75, CART yöntemiyle %70 F1 doğruluğu elde edilmiştir.

Tablo 4.3 Ofis 4 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,98	0,96	0,99	0,98	0,98	0,99
Duyarlılık-0	0,98	0,99	0,99	0,98	0,97	1,00
F1 Skor-0	0,98	0,97	0,99	0,98	0,98	0,99
Kesinlik-1	0,97	0,98	0,97	0,98	0,97	0,98
Duyarlılık-1	0,99	0,99	1,00	0,99	0,99	0,98
F1 Skor-1	0,98	0,98	0,99	0,98	0,98	0,98
Kesinlik-2	0,95	0,95	0,96	0,96	0,95	0,93
Duyarlılık-2	0,87	0,82	0,89	0,83	0,89	0,84
F1 Skor-2	0,91	0,88	0,93	0,89	0,92	0,88
Kesinlik-3	0,54	0,46	0,60	0,50	0,56	0,46
Duyarlılık-3	0,77	0,77	0,81	0,84	0,75	0,69
F1 Skor-3	0,63	0,58	0,69	0,63	0,64	0,55
Genel Doğruluk	0,90	0,87	0,91	0,88	0,90	0,87

Ofis 4 için RO yöntemiyle %90, ÇKA yöntemiyle %87, CatBoost yöntemiyle %91, XGBoost yöntemiyle %88, LightGBM yöntemiyle %90, CART yöntemiyle %87 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %98, ÇKA yöntemiyle %97, CatBoost yöntemiyle %99, XGBoost yöntemiyle %98, LightGBM yöntemiyle %98, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %98, ÇKA yöntemiyle %98,CatBoost yöntemiyle %99, XGBoost yöntemiyle %98, LightGBM yöntemiyle %98, CART yöntemiyle %98; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %91, ÇKA yöntemiyle %88, CatBoost yöntemiyle %93, XGBoost yöntemiyle %89, LightGBM yöntemiyle %92, CART yöntemiyle %88; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %63, ÇKA yöntemiyle %58, CatBoost yöntemiyle %69, XGBoost yöntemiyle %63, LightGBM yöntemiyle %64, CART yöntemiyle %55 F1 doğruluğu elde edilmiştir.

Tablo 4.4 Ofis 5 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,97	0,95	0,97	0,97	0,97	0,97
Duyarlılık-0	0,98	0,94	0,98	0,99	0,98	0,99
F1 Skor-0	0,97	0,95	0,98	0,98	0,97	0,98
Kesinlik-1	0,96	0,96	0,96	0,96	0,96	0,96
Duyarlılık-1	0,99	1,00	1,00	0,96	1,00	0,96
F1 Skor-1	0,97	0,98	0,98	0,96	0,98	0,96
Kesinlik-2	0,95	0,94	0,96	0,95	0,94	0,92
Duyarlılık-2	0,88	0,85	0,89	0,86	0,88	0,85
F1 Skor-2	0,91	0,90	0,92	0,90	0,91	0,89
Kesinlik-3	0,52	0,49	0,56	0,49	0,49	0,42
Duyarlılık-3	0,70	0,74	0,78	0,76	0,63	0,59
F1 Skor-3	0,59	0,59	0,65	0,59	0,55	0,49
Genel Doğruluk	0,90	0,88	0,91	0,89	0,89	0,87

Ofis 5 için RO yöntemiyle %90, ÇKA yöntemiyle %88, CatBoost yöntemiyle %91, XGBoost yöntemiyle %89, LightGBM yöntemiyle %89, CART yöntemiyle %87 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %97, ÇKA yöntemiyle %95, CatBoost yöntemiyle %98, XGBoost yöntemiyle %98, LightGBM yöntemiyle %97, CART yöntemiyle %98; zemin (1etiketi) RO yöntemiyle %97, ÇKA yöntemiyle %98, CatBoost yöntemiyle %98, XGBoost yöntemiyle %96, LightGBM yöntemiyle %98, CART yöntemiyle %96; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %91, ÇKA yöntemiyle %90, CatBoost yöntemiyle %92, XGBoost yöntemiyle %90, LightGBM yöntemiyle %91, CART yöntemiyle %89; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %59, ÇKA yöntemiyle %59, CatBoost yöntemiyle %65, XGBoost yöntemiyle %59, LightGBM yöntemiyle %55, CART yöntemiyle %49 F1 doğruluğu elde edilmiştir.

Tablo 4.5 Ofis 6 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,97	0,96	0,98	0,98	0,98	0,98
Duyarlılık-0	0,98	1,00	0,96	0,97	0,94	0,99
F1 Skor-0	0,98	0,98	0,97	0,97	0,96	0,99
Kesinlik-1	0,99	0,99	0,99	0,99	0,99	0,99
Duyarlılık-1	0,94	0,92	0,99	0,91	0,88	0,95
F1 Skor-1	0,97	0,96	0,99	0,95	0,93	0,97
Kesinlik-2	0,91	0,92	0,92	0,92	0,90	0,88
Duyarlılık-2	0,86	0,82	0,87	0,81	0,87	0,83
F1 Skor-2	0,88	0,86	0,89	0,86	0,88	0,85
Kesinlik-3	0,55	0,51	0,58	0,48	0,52	0,50
Duyarlılık-3	0,67	0,73	0,70	0,74	0,67	0,63
F1 Skor-3	0,61	0,60	0,63	0,58	0,58	0,55
Genel Doğruluk	0,87	0,85	0,88	0,84	0,85	0,84

Ofis 6 için RO yöntemiyle %87, ÇKA yöntemiyle %85, CatBoost yöntemiyle %88, XGBoost yöntemiyle %84, LightGBM yöntemiyle %85, CART yöntemiyle %84 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %98, ÇKA yöntemiyle %98, CatBoost yöntemiyle %97, XGBoost yöntemiyle %97, LightGBM yöntemiyle %96, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %97, ÇKA yöntemiyle %96,CatBoost yöntemiyle %99, XGBoost yöntemiyle %95, LightGBM yöntemiyle %93, CART yöntemiyle %97; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %88, ÇKA yöntemiyle %86, CatBoost yöntemiyle %89, XGBoost yöntemiyle %86, LightGBM yöntemiyle %88, CART yöntemiyle %85; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %61, ÇKA yöntemiyle %60, CatBoost yöntemiyle %63, XGBoost yöntemiyle %58, LightGBM yöntemiyle %58, CART yöntemiyle %55 F1 doğruluğu elde edilmiştir.

Tablo 4.6 Ofis 7 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,98	0,98	0,99	0,98	0,98	0,99
Duyarlılık-0	0,88	0,96	0,83	0,82	0,78	0,95
F1 Skor-0	0,93	0,97	0,90	0,90	0,87	0,97
Kesinlik-1	0,98	0,98	0,98	0,98	0,98	0,98
Duyarlılık-1	1,00	0,99	1,00	0,99	0,99	1,00
F1 Skor-1	0,99	0,99	0,99	0,98	0,98	0,99
Kesinlik-2	0,92	0,91	0,91	0,92	0,91	0,91
Duyarlılık-2	0,96	0,89	0,98	0,93	0,97	0,89
F1 Skor-2	0,94	0,90	0,94	0,93	0,94	0,90
Kesinlik-3	0,77	0,68	0,74	0,65	0,70	0,64
Duyarlılık-3	0,71	0,73	0,67	0,73	0,70	0,69
F1 Skor-3	0,74	0,70	0,70	0,69	0,70	0,66
Genel Doğruluk	0,91	0,89	0,90	0,89	0,89	0,88

Ofis 7 için RO yöntemiyle %91, ÇKA yöntemiyle %89, CatBoost yöntemiyle %90, XGBoost yöntemiyle %89, LightGBM yöntemiyle %89, CART yöntemiyle %88 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %93, ÇKA yöntemiyle %97, CatBoost yöntemiyle %90, XGBoost yöntemiyle %90, LightGBM yöntemiyle %87, CART yöntemiyle %97; zemin (1etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %99,CatBoost yöntemiyle %99, XGBoost yöntemiyle %98, LightGBM yöntemiyle %98, CART yöntemiyle %99; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %94, ÇKA yöntemiyle %90, CatBoost yöntemiyle %94, XGBoost yöntemiyle %93, LightGBM yöntemiyle %94, CART yöntemiyle %90; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %74, ÇKA yöntemiyle %70, CatBoost yöntemiyle %70, XGBoost yöntemiyle %69, LightGBM yöntemiyle %70, CART yöntemiyle %66 F1 doğruluğu elde edilmiştir.

Tablo 4.7 Ofis 8 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,98	0,97	0,99	0,99	0,99	0,99
Duyarlılık-0	0,99	0,98	0,99	0,99	0,98	1,00
F1 Skor-0	0,99	0,98	0,99	0,99	0,99	0,99
Kesinlik-1	0,95	0,95	0,95	0,95	0,95	0,96
Duyarlılık-1	0,99	0,98	1,00	0,97	0,99	0,98
F1 Skor-1	0,97	0,96	0,97	0,96	0,97	0,97
Kesinlik-2	0,96	0,96	0,97	0,97	0,96	0,95
Duyarlılık-2	0,89	0,79	0,90	0,84	0,90	0,85
F1 Skor-2	0,92	0,86	0,93	0,90	0,93	0,90
Kesinlik-3	0,53	0,38	0,57	0,46	0,54	0,43
Duyarlılık-3	0,79	0,83	0,83	0,86	0,77	0,71
F1 Skor-3	0,63	0,52	0,67	0,60	0,64	0,54
Genel Doğruluk	0,91	0,84	0,92	0,88	0,91	0,88

Ofis 8 için RO yöntemiyle %91, ÇKA yöntemiyle %89, CatBoost yöntemiyle %90, XGBoost yöntemiyle %88, LightGBM yöntemiyle %89, CART yöntemiyle %88 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %98, CatBoost yöntemiyle %99, XGBoost yöntemiyle %99, LightGBM yöntemiyle %99, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %97, ÇKA yöntemiyle %96,CatBoost yöntemiyle %97, XGBoost yöntemiyle %96, LightGBM yöntemiyle %97, CART yöntemiyle %97; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %92, ÇKA yöntemiyle %86, CatBoost yöntemiyle %93, XGBoost yöntemiyle %90, LightGBM yöntemiyle %93, CART yöntemiyle %90; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %63, ÇKA yöntemiyle %52, CatBoost yöntemiyle %67, XGBoost yöntemiyle %60, LightGBM yöntemiyle %64, CART yöntemiyle %54 F1 doğruluğu elde edilmiştir.

Tablo 4.8 Ofis 9 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,95	0,94	0,96	0,96	0,96	0,96
Duyarlılık-0	1,00	0,97	1,00	1,00	1,00	1,00
F1 Skor-0	0,98	0,95	0,98	0,98	0,98	0,98
Kesinlik-1	0,93	0,93	0,92	0,93	0,93	0,93
Duyarlılık-1	1,00	1,00	1,00	0,99	1,00	0,99
F1 Skor-1	0,96	0,96	0,96	0,96	0,96	0,96
Kesinlik-2	0,90	0,91	0,92	0,91	0,89	0,87
Duyarlılık-2	0,95	0,93	0,97	0,94	0,96	0,89
F1 Skor-2	0,93	0,92	0,94	0,92	0,92	0,88
Kesinlik-3	0,86	0,80	0,93	0,82	0,87	0,66
Duyarlılık-3	0,65	0,71	0,70	0,68	0,59	0,56
F1 Skor-3	0,74	0,76	0,80	0,74	0,70	0,60
Genel Doğruluk	0,91	0,90	0,93	0,91	0,90	0,86

Ofis 9 için RO yöntemiyle %91, ÇKA yöntemiyle %90, CatBoost yöntemiyle %93, XGBoost yöntemiyle %91, LightGBM yöntemiyle %90, CART yöntemiyle %86 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %98, ÇKA yöntemiyle %95, CatBoost yöntemiyle %98, XGBoost yöntemiyle %98, LightGBM yöntemiyle %98, CART yöntemiyle %98; zemin (1etiketi) RO yöntemiyle %96, ÇKA yöntemiyle %96,CatBoost yöntemiyle %96, XGBoost yöntemiyle %96, LightGBM yöntemiyle %96, CART yöntemiyle %96; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %93, ÇKA yöntemiyle %92, CatBoost yöntemiyle %94, XGBoost yöntemiyle %92, LightGBM yöntemiyle %92, CART yöntemiyle %88; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %74, ÇKA yöntemiyle %76, CatBoost yöntemiyle %80, XGBoost yöntemiyle %74, LightGBM yöntemiyle %70, CART yöntemiyle %60 F1 doğruluğu elde edilmiştir.

Tablo 4.9 Ofis 10 verisinin kullanılan yöntemlere göre sınıflandırma sonuçları

Değerler-Sınıflar	Rastgele Orman	Çok Katmanlı Algılayıcı	CatBoost	XGBoost	LightGBM	CART
Kesinlik-0	0,98	0,98	0,98	0,98	0,98	0,99
Duyarlılık-0	0,98	0,99	0,99	0,98	0,98	0,99
F1 Skor-0	0,98	0,98	0,99	0,98	0,98	0,99
Kesinlik-1	0,98	0,98	0,98	0,98	0,98	0,98
Duyarlılık-1	1,00	1,00	1,00	1,00	1,00	1,00
F1 Skor-1	0,99	0,99	0,99	0,99	0,99	0,99
Kesinlik-2	0,93	0,93	0,92	0,93	0,93	0,92
Duyarlılık-2	0,95	0,88	0,98	0,91	0,98	0,89
F1 Skor-2	0,94	0,90	0,95	0,92	0,95	0,90
Kesinlik-3	0,80	0,65	0,92	0,71	0,91	0,65
Duyarlılık-3	0,73	0,75	0,70	0,75	0,73	0,72
F1 Skor-3	0,77	0,70	0,80	0,73	0,81	0,68
Genel Doğruluk	0,93	0,90	0,94	0,91	0,94	0,90

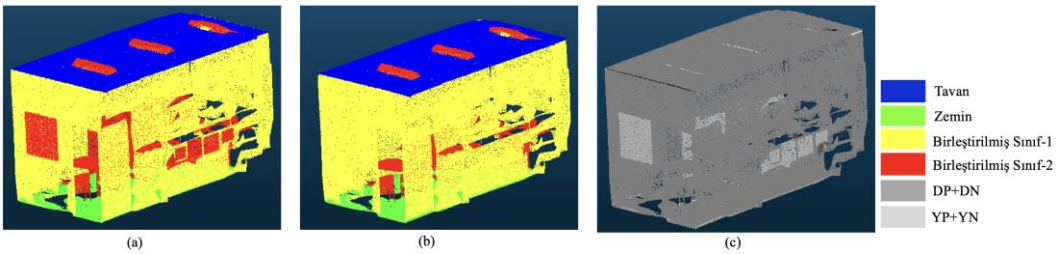
Ofis 10 için RO yöntemiyle %93, ÇKA yöntemiyle %90, CatBoost yöntemiyle %94, XGBoost yöntemiyle %91, LightGBM yöntemiyle %94, CART yöntemiyle %90 genel doğruluk elde edilmiştir. Tavan (0 etiketi) RO yöntemiyle %98, ÇKA yöntemiyle %98, CatBoost yöntemiyle %99, XGBoost yöntemiyle %98, LightGBM yöntemiyle %98, CART yöntemiyle %99; zemin (1etiketi) RO yöntemiyle %99, ÇKA yöntemiyle %99,CatBoost yöntemiyle %99, XGBoost yöntemiyle %99, LightGBM yöntemiyle %99, CART yöntemiyle %99; birleştirilmiş sınıf-1(2 etiketi) RO yöntemiyle %94, ÇKA yöntemiyle %90, CatBoost yöntemiyle %95, XGBoost yöntemiyle %92, LightGBM yöntemiyle %95, CART yöntemiyle %90; birleştirilmiş-2 (3 etiketi) RO yöntemiyle %77, ÇKA yöntemiyle %70, CatBoost yöntemiyle %80, XGBoost yöntemiyle %73, LightGBM yöntemiyle %81, CART yöntemiyle %68 F1 doğruluğu elde edilmiştir.

4.2 Tahmin Sonuçlarının Cloud Compare Yazılımında Görünümü

Bu çalışmada yapılan uygulama sonucunda hem genel doğruluk hem de nesne bazındaki doğruluklar incelendiğinde en yüksek doğruluk Ofis 3 verisi için %92 doğrulukla LightGBM yöntemi , Ofis 4 verisi için % 91 doğrulukla CatBoost yöntemi, Ofis 5 verisi için %91 doğrulukla CatBoost yöntemi, Ofis 7 verisi için %89 doğrulukla Çok Katmanlı Algılayıcı yöntemi , Ofis 8 verisi için %91 doğrulukla Rastgele Orman yöntemi, Ofis 10 verisi için %91 doğrulukla XGBoost

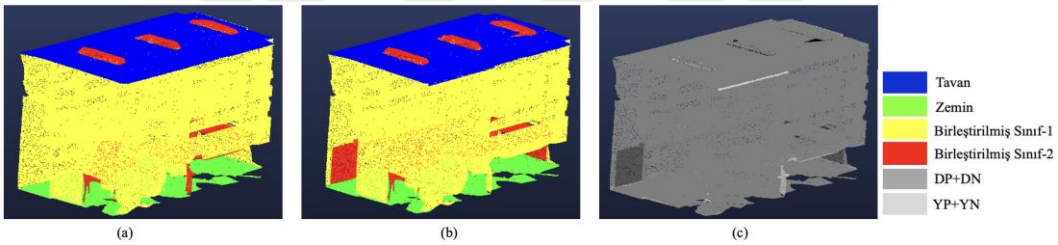
yöntemi olmuştur.Bu ofis verileri ile ilgili referans değerleri tahmin edilen sınıflar ve DP, DN, YP ve YN sınıfları lejantta belirtilen renklerle Şekil 4.1, Şekil 4.2, Şekil 4.3, Şekil 4.4, Şekil 4.5 ve Şekil 4.6'da gösterilmiştir.

4.2.1 Ofis 3 - LightGBM Yöntemi



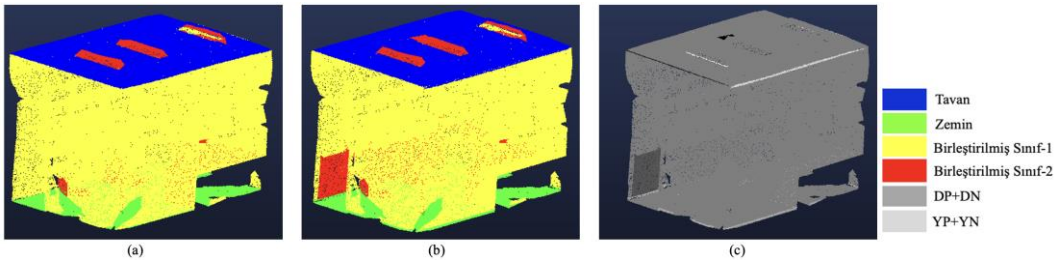
Şekil 4.1 Ofis 3 verisi üzerindeki LightGBM yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

4.2.2 Ofis 4 - Catboost Yöntemi



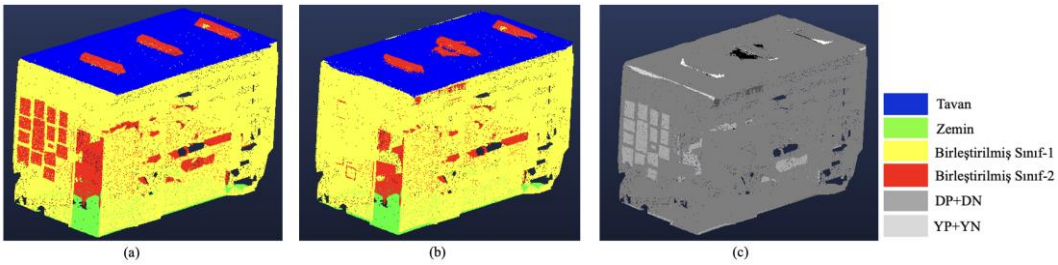
Şekil 4.2 Ofis 4 verisi üzerindeki CatBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

4.2.3 Ofis 5 - Catboost Yöntemi



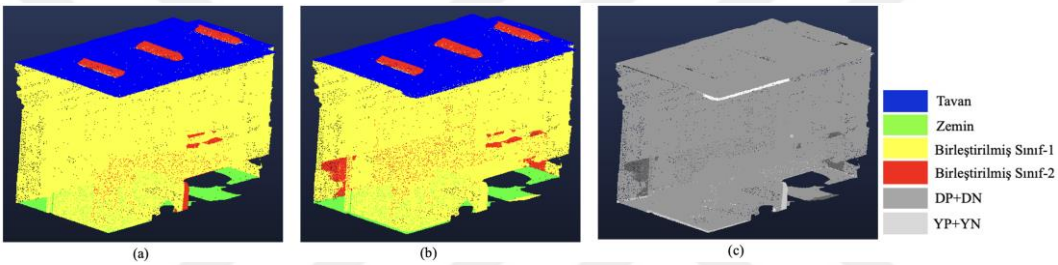
Şekil 4.3 Ofis 5 verisi üzerindeki CatBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

4.2.4 Ofis 7 - Çok Katmanlı Algılayıcı Yöntemi



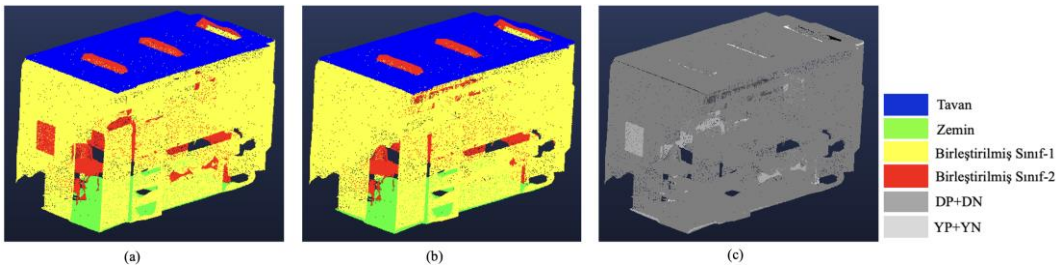
Şekil 4.4 Ofis 7 verisi üzerindeki Çok Katmanlı Algılayıcı yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

4.2.5 Ofis 8 - Rastgele Orman Yöntemi



Şekil 4.5 Ofis 8 verisi üzerindeki Rastgele Orman yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

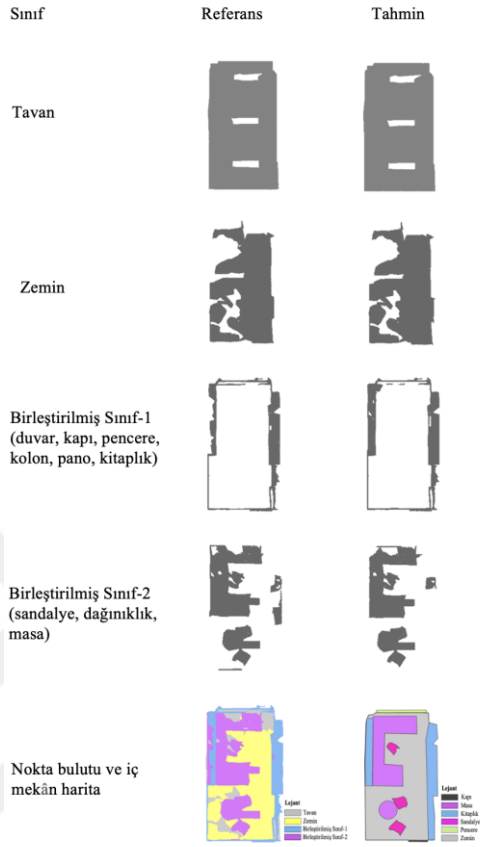
4.2.6 Ofis 10 - XGBoost Yöntemi



Şekil 4.6 Ofis 10 verisi üzerindeki XGBoost yöntemine göre doğruluk sonuçları: (a) referans sınıflar, (b) tahmin edilen sınıflar, (c) doğru sınıflandırma (DP ve DN) ve yanlış sınıflandırma (YP ve YN)

4.3 Planların Oluşturulması

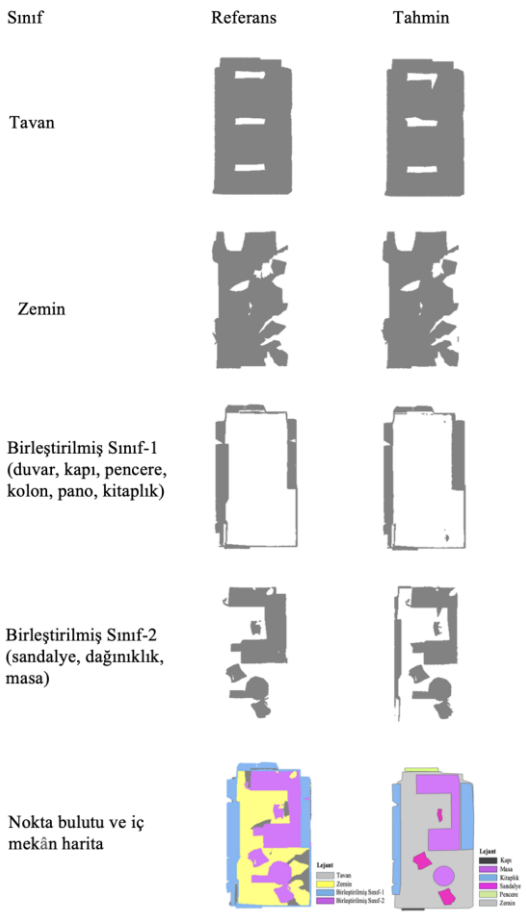
4.3.1 Ofis 3 - LightGBM Yöntemi



Harita 4.1 LightGBM yöntemiyle Ofis 3 planı

Ofis 3 için en yüksek doğruluk LightGBM yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan ve zemin yüksek doğrulukla sınıflandırılmış ve başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Alt tarafta bulunan masa ve sandalye nesneleri, net bir şekilde sınıflandırılmış ve şekilleri anlaşılır biçimde planda yer almıştır. Sağ tarafta yer alan kitaplık, çizilebilir düzeyde sınıflandırılmış ve planda başarılı bir şekilde gösterilmiştir. Bununla birlikte, sol üst köşede masa ve kitaplık nesneleri üst üste geldiğinden, bu bölgede sınıflandırma karışıklığı yaşanmıştır. Masa net bir şekilde çizilebilmişken, kitaplık nesnesi belirgin şekilde ayrıştırılamadığı için planda yeterince temsil edilememiştir. Ayrıca, sağdaki kitaplığın küçük bir bölümü başka bir sınıfa dahil edilerek yanlış sınıflandırılmıştır.

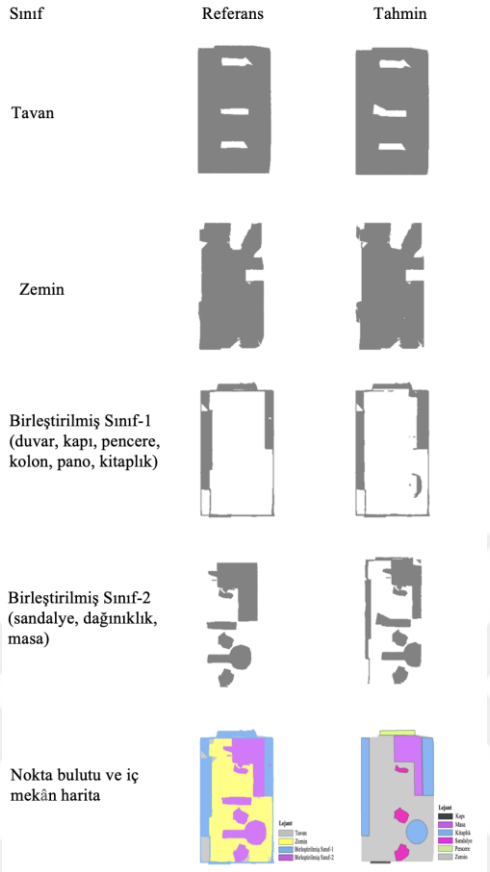
4.3.2 Ofis 4 – CatBoost Yöntemi



Harita 4.2 CatBoost yöntemiyle ofis 4 planı

Ofis 4 için en yüksek sınıflandırma doğruluğu CatBoost yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan ve zemin yüksek doğrulukla sınıflandırılmış olup tavanda bazı bölgelerde yanlış sınflandırılan yerler olsa da başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Alt tarafta bulunan masa ve sandalye nesneleri, net bir şekilde sınıflandırılmış ve şekilleri anlaşılır biçimde planda yer almıştır. Sol tarafta yer alan kitaplığın bir kısmı birleştirilmiş sınıf 2'de sınıflandırılmış olmasına rağmen çizilebilmiştir. Bununla birlikte, sol üst köşede masa ve kitaplık nesneleri üst üste geldiğinden, bu bölgede sınıflandırma karışıklığı yaşanmış ama yine de çizilip temsil edilmiştir.

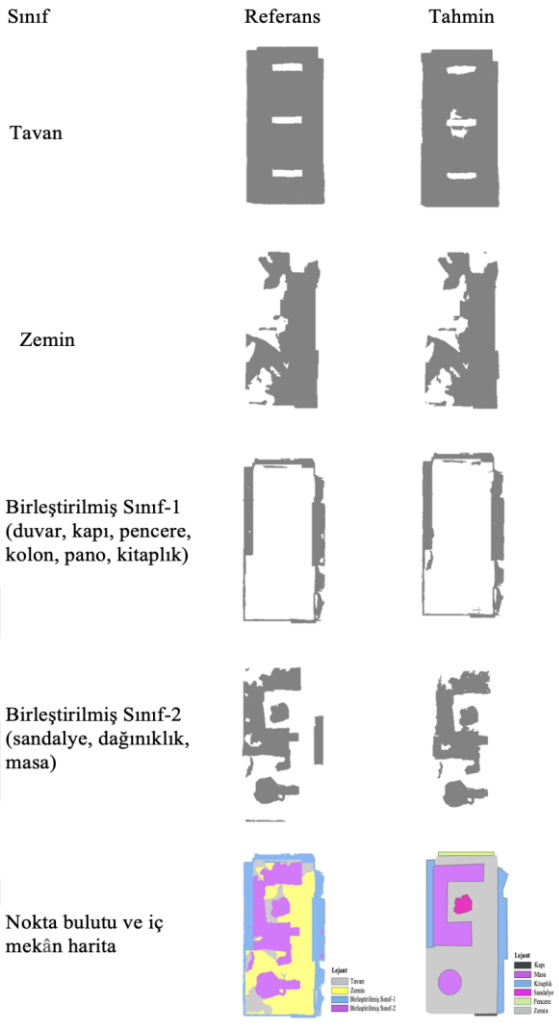
4.3.3 Ofis 5 - CatBoost Yöntemi



Harita 4.3 CatBoost yöntemiyle Ofis 5 planı

Ofis 5 için en yüksek sınıflandırma doğruluğu CatBoost yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan ve zemin yüksek doğrulukla sınıflandırılmış olup başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Alt tarafta bulunan masanın bir kısmı birleştirilmiş sınıf-2'de sınıflandırılmış ama daire biçiminde çizilebilmiş sandalyeler ise net bir şekilde sınıflandırılmış ve şekilleri anlaşılır biçimde planda yer almıştır. Sağ tarafta yer alan kitaplığın bir kısmı birleştirilmiş sınıf 2'de sınıflandırılmış olmasına rağmen çizilebilmiştir. Bununla birlikte, sol üst köşede masa ve kitaplık nesneleri üst üste geldiğinden, bu bölgede sınıflandırma karışıklığı yaşanmış ama yine de çizilebilip temsil edilmiştir.

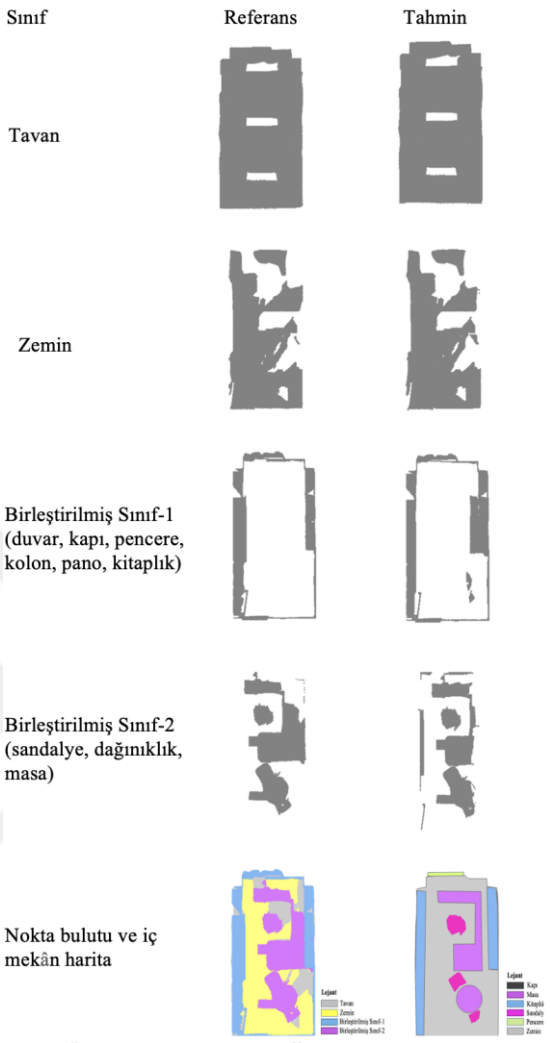
4.3.4 Ofis 7- Çok Katmanlı Algılayıcı Yöntemi



Harita 4.4 ÇKA yöntemiyle ofis 7 planı

Ofis 7 için en yüksek sınıflandırma doğruluğu Çok Katmanlı Algılayıcı yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan yüksek doğrulukla sınıflandırılmasına rağmen bazı kısımlarında sınıflandırma yanlışlığı yapılmış ama çizilebilmiş; zemin ise yüksek doğrulukla sınıflandırılmış olup başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Alt tarafta bulunan masa net bir şekilde sınıflandırılmış ve anlaşılır biçimde planda yer almıştır. İki kitaplık ve masa net bir şekilde sınıflandırılmış ve anlaşılır biçimde planda temsil edilmiştir.

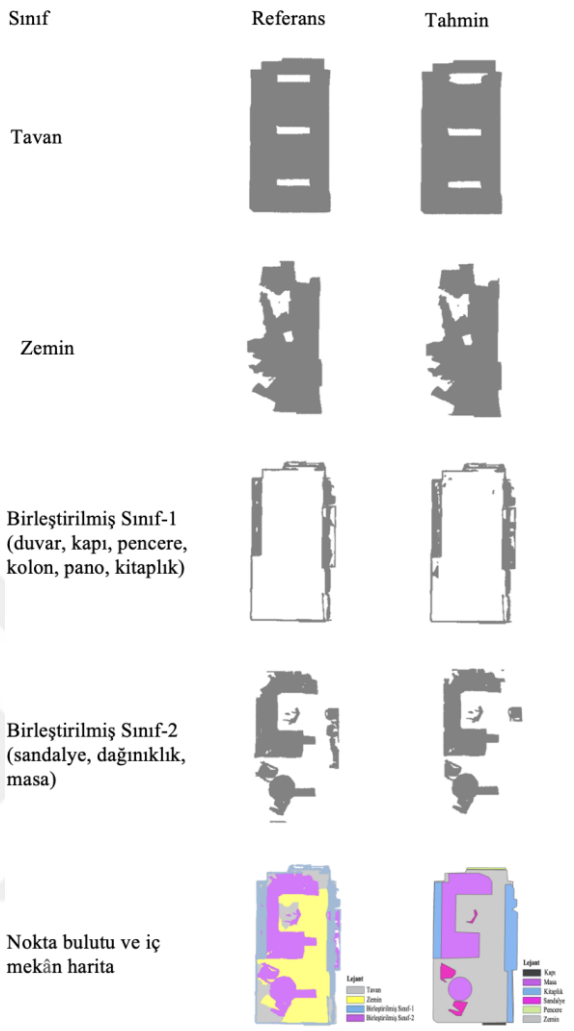
4.3.5 Ofis 8 – Rastgele Orman Yöntemi



Harita 4.5 RO yöntemiyle ofis 8 planı

Ofis 8 için en yüksek sınıflandırma doğruluğu Rastgele Orman yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan ve zemin yüksek doğrulukla sınıflandırılmış olup başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Sağ tarafta bulunan masanın bir kısmı birleştirilmiş sınıf-1'de yanlış sınıflandırılmıştır. Yine de doğru sınıflandırılan kısmı çizilmiş ve planda gösterilmiştir. Sol taraftaki kitaplığın bir kısmı birleştirilmiş sınıf-2' de yanlış sınıflandırılmış olsa da çizilebilmiştir. Aşağıda bulunan masa ve bütün sandalyeler doğru sınıflandırılmış ve net bir şekilde çizilmiştir.

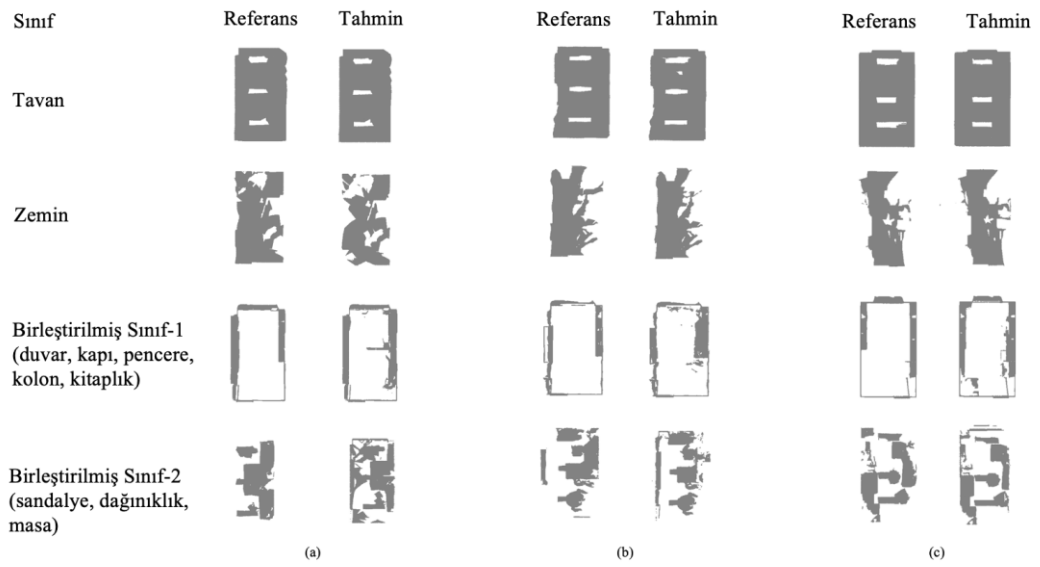
4.3.6 Ofis 10 – XGBOOST Yöntemi



Harita 4.6 XGBoost yöntemiyle ofis 10 planı

Ofis 10 için en yüksek sınıflandırma doğruluğu XGBoost yöntemiyle elde edilmiştir. Sınıflandırma sonuçlarına dayanarak oluşturulan planda, tavan ve zemin yüksek doğrulukla sınıflandırılmış olup başarılı bir şekilde çizilmiştir. Kapı ve pencere, çizilebilir seviyede doğru şekilde sınıflandırılmış ve planda temsil edilmiştir. Kitaplıkların büyük bir kısmı doğru sınıflandırılmış olup anlaşılır şekilde çizilmiştir. Masa ve sandalyeler doğru sınıflandırılmış olup anlaşılır şekilde çizilmiştir.

4.4 Planı Oluşturulamayan Ofisler



Harita 4.7 (a) Ofis 2 verisinin CatBoost yöntemine göre sınıflandırılması (b) Ofis 6 verisinin RO yöntemine göre sınıflandırılması; (c) Ofis 9 verisinin CatBoost yöntemine göre sınıflandırılması

Ofis 2, Ofis 6 ve Ofis 9 verileri, Şekil 4.7'de gösterildiği gibi sınıflandırma işlemine dahil edilmiştir. Ancak, veri toplama sürecinde gerçekleşen boşluklar nedeniyle, bu ofislere ait planlar beklenen doğrulukta oluşturulamamıştır. Boşluklar, sınıflandırma sonuçlarının doğruluğunu ve planların netliğini olumsuz etkilemiş, bu da ilgili bölgelerin detaylı bir şekilde görselleştirilmesini zorlaştırmıştır.

SONUÇ

Bu çalışmada, iç mekân nokta bulutlarının sınıflandırılması için kullanılan makine öğrenimi yöntemlerinden Rastgele Orman yöntemiyle ortalama %90, Çok Katmanlı Algılayıcı yöntemiyle ortalama %87, CatBoost yöntemiyle ortalama %91, XGBoost yöntemiyle ortalama %88, LightGBM yöntemiyle ortalama %89 ve CART yöntemiyle ortalama %86 genel sınıflandırma doğruluğu elde edilmiştir. Nesne bazında ise tavan için Rastgele Orman yöntemiyle %98, ÇKA yöntemiyle %97, CatBoost yöntemiyle %98, XGBoost yöntemiyle %97, LightGBM yöntemiyle %97, CART yöntemiyle %99; zemin için RO yöntemiyle %95, ÇKA yöntemiyle %94, CatBoost yöntemiyle %97, XGBoost yöntemiyle %93, LightGBM yöntemiyle %93, CART yöntemiyle %94; birleştirilmiş sınıf-1 için Rastgele Orman yöntemiyle %92, ÇKA yöntemiyle %88, CatBoost yöntemiyle %93, XGBoost yöntemiyle %90, LightGBM yöntemiyle %92, CART yöntemiyle %88; birleştirilmiş sınıf-2 için Rastgele Orman yöntemiyle %66, ÇKA yöntemiyle %62, CatBoost yöntemiyle %69, XGBoost yöntemiyle %64, LightGBM yöntemiyle %64, CART yöntemiyle %58 F1 Skor elde edilmiştir. Genel olarak, kullanılan yöntemlerde tavan, zemin ve birleştirilmiş sınıf-1 (duvar, kapı, pencere, kolon, pano, kitaplık) gibi temel sınıfları başarılı bir şekilde tahmin edilmiştir. Bu sınıflar, harita yapımı ve yaya navigasyonu gibi uygulamalar için gerekli doğruluk seviyelerine ulaşmıştır. Ancak, birleştirilmiş sınıf-2'de yer alan masa, sandalye ve dağınıklık gibi karmaşık nesnelerin sınıflandırılmasında tüm yöntemlerin performansı sınırlı kalmıştır. Buna rağmen, bu sınıflar için yapılan tahminlerin planlarla temsil edilmesine yeterli olduğu gözlemlenmiştir. Tüm süreçler, MacBook Air (Apple M2 çip, 8 çekirdekli CPU, 8 GB RAM) donanım özelliklerine sahip bir bilgisayar üzerinde gerçekleştirilmiştir. Kullanılan tüm makine öğrenmesi yöntemleri verimli bir şekilde çalışmış olmakla birlikte, özellikle LightGBM ve XGBoost algoritmaları en hızlı sonuç veren yöntemler olmuştur. Nokta bulutlarının manuel olarak sınıflandırılması yerine makine öğrenmesi tabanlı yöntemlerle

sınıflandırılması ve bu sınıflandırmalara dayalı planların üretilmesi; zaman, hız ve işlem performansı açısından önemli avantajlar sağlamaktadır. Ayrıca, elde edilen semantik sınıflar için ortalama yükseklik değerleri kullanılarak üç boyutlu gösterimler yapılabilir. Mekân içerisindeki eşyaların sınıflandırılması sayesinde nokta bulutu verisinden ayrılarak mekânın üç boyutlu modelleri elde edilebilir. Literatürde iç mekân nokta bulutu sınıflandırmalarında sıklıkla kullanılan derin öğrenme tabanlı yöntemler, model karmaşıklığı ve nokta bulutu verisinin yüksek hacmi göz önünde bulundurulduğunda, donanım kullanımını daha verimli hale getiren makine öğrenmesi algoritmalarının alternatif olarak etkili bir şekilde kullanılabileceğini ortaya koymuştur.

KAYNAKÇA

[1] M. Štroner, R. Urban, and L. Línková, "Color-based point cloud classification using a novel gaussian mixed modeling-based approach versus a deep neural network," *Remote Sens.,* vol. 16, no. 1, p. 115, 2024, doi: 10.3390/rs16010115.

[2] C. S. Hsieh and X. J. Ruan, "Automated semantic segmentation of indoor point clouds from close-range images with three-dimensional deep learning," *Buildings*, vol. 13, no. 2, p. 468, 2023, doi: 10.3390/buildings13020468.

[3] M. Gunduz, U. Isikdag, and M. Basaraner, "A review of recent research in indoor modelling & mapping," in Int. Arch. Photogramm. *Remote Sens. Spatial Inf. Sci.*, 2016, pp. 289–294, doi: 10.5194/isprsarchives-XLI-B4-289-2016.

[4] H. Wu et al., "Classification of point clouds for indoor components using few labeled samples," *Remote Sens.,* vol. 12, no. 14, 2020, doi: 10.3390/rs12142181.

[5] Y. Deng, H. Ai, Z. Deng, W. Gao, and J. Shang, "An overview of indoor positioning and mapping technology standards," *Standards,* vol. 2, no. 2, pp. 157–183, 2022, doi: 10.3390/standards2020012.

[6] L. Díaz-Vilariño, K. Khoshelham, J. Martínez-Sánchez, and P. Arias, "3D modeling of building indoor spaces and closed doors from imagery and point clouds," *Sensors,* vol. 15, no. 2, pp. 3491–3512, 2015, doi: 10.3390/s150203491.

[7] H. Lin et al., "Semantic segmentation of 3D indoor LiDAR point clouds through feature pyramid architecture search," ISPRS J. Photogramm. *Remote Sens.,* vol. 177, pp. 279–290, 2021, doi: 10.1016/j.isprsjprs.2021.05.009.

[8] J. Zhao, X. Zhang, and Y. Wang, "Indoor 3D point clouds semantics segmentation base on the modified PointNet network," in Int. Arch. Photogramm. *Remote Sens. Spatial Inf. Sci.,* 2020, pp. 369–373, doi: 10.5194/isprs-archives-XLIII-B2-2020-369-2020.

[9] R. Otero, S. Lagüela, I. Garrido, and P. Arias, "Mobile indoor mapping technologies: A review," Autom. Constr., Dec. 2020, doi: 10.1016/j.autcon.2020.103399.

[10] D. Shan et al., "An Efficient Point Cloud Semantic Segmentation Method Based on Bilateral Enhancement and Random Sampling," *Electronics,* vol. 12, no. 24, 2023, doi: 10.3390/electronics12244927.

[11] J. Chen and K. C. Clarke, "Indoor cartography," *Cartogr. Geogr. Inf. Sci.*, vol. 47, no. 2, pp. 95–109, 2020.

[12] A. Bilgili, A. Şen, and M. Başaraner, "İç mekân navigasyonu ağ modelleri: Karşılaştırmalı bir inceleme," *J. Geodesy Geoinf.*, c. 9, no. 2, ss. 108–126, 2022, doi: 10.9733/jgg.2022r0008.t.

[13] I. Armeni et al., "3D semantic parsing of large-scale indoor spaces (a) raw point cloud (b) space parsing and alignment in canonical 3D space (c) building element detection enclosed spaces." [Online]. Available: http://buildingparser.stanford.edu/

[14] F. Poux, C. Mattes, and L. Kobbelt, "Unsupervised segmentation of indoor 3D point cloud: application to object-based classification," in Int. Arch. Photogramm. *Remote Sens. Spatial Inf. Sci.*, 2020, pp. 111–118, doi: 10.5194/isprs-archives-XLIV-4-W1-2020-111-2020.

[15] X. T. Chen, Y. Li, J. H. Fan, and R. Wang, "RGAM: A novel network architecture for 3D point cloud semantic segmentation in indoor scenes," *Inf. Sci.*, vol. 571, pp. 87–103, 2021, doi: 10.1016/j.ins.2021.04.069.

[16] Y. Su, L. Jiang, and J. Cao, "Point cloud semantic segmentation using multi scale sparse convolution neural network," 2022. [Online]. Available: http://arxiv.org/abs/2205.01550

[17] J. Chen, Z. Kira, and Y. K. Cho, "LRGNet: learnable region growing for class-agnostic point cloud segmentation," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2799–2806, 2021.

[18] K. Khoshelham et al., "The ISPRS benchmark on indoor modelling: preliminary results," Int. Arch. Photogramm. *Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B4-2020, pp. 207–211, 2020.

[19] X. Meng et al., "A new self-augment CNN for 3D point cloud classification and segmentation," Int. J. Mach. Learn. Cybern., vol. 15, no. 3, pp. 807–818, 2024, doi: 10.1007/s13042-023-01940-4.

[20] A. Kiran and D. Vasumathi, "Data mining: Min–max normalization based data perturbation technique for privacy preservation," in Adv. Intell. Syst. Comput., *Springer*, 2020, pp. 723–734, doi: 10.1007/978-981-15-1480-7_66.

[21] A. Alsarhan, F. Hussein, S. Moh, and F. S. El-Salhi, "The effect of preprocessing techniques, applied to numeric features, on classification algorithms' performance," *Data*, vol. 6, no. 2, p. 11, 2021, doi: 10.3390/data.

[22] K. Sharifani and M. Amini, "Machine learning and deep learning: A review of methods and applications." [Online]. Available: https://ssrn.com/abstract=4458723

[23] T. Emmanuel et al., "A survey on missing data in machine learning," *J. Big Data*, vol. 8, no. 1, pp. 1–37, 2021, doi: 10.1186/s40537-021-00516-9.

[24] B. Mahesh, "Machine learning algorithms - A review," Int. J. Sci. Res., vol. 9, no. 1, pp. 381–386, Jan. 2020, doi: 10.21275/art20203995.

[25] R. Sengupta et al., "Journal of critical reviews artificial intelligence and quantum computing for a smarter wireless network," *Artif. Intell.*, vol. 7, no. 19, 2020.

[26] M. M. Taye, “Understanding of machine learning with deep learning: Architectures, workflow, applications and future directions,” *Computers,* vol. 12, no. 5, p. 91, 2023, doi: 10.3390/computers12050091.

[27] O. Engin and B. Sarıcan, “Makine çizelgeleme problemlerinin çözümünde pekiştirmeli öğrenme etkisinin analizi,” 2024.

[28] H. Sotoudeh et al., “Artificial intelligence in the management of glioma: Era of personalized medicine,” *Front. Oncol.,* vol. 9, p. 768, 2019, doi: 10.3389/fonc.2019.00768.

[29] Z. Sun et al., “An improved random forest based on the classification accuracy and correlation measurement of decision trees,” *Expert Syst. Appl.,* vol. 237, p. 121549, 2024, doi: 10.1016/j.eswa.2023.121549.

[30] M. Zeybek, “Classification of UAV point clouds by random forest machine learning algorithm,” Turk. J. Eng., vol. 5, no. 2, pp. 48–57, 2021, doi: 10.31127/tuje.669566.

[31] M. E. Atik, Z. Duran, and D. Z. Seker, “Machine learning-based supervised classification of point clouds using multiscale geometric features,” *ISPRS Int. J. Geo-Inf.,* vol. 10, no. 3, 2021, doi: 10.3390/ijgi10030187.

[32] Z. Duran, K. Ozcan, and M. E. Atik, “Classification of photogrammetric and airborne lidar point clouds using machine learning algorithms,” *Drones,* vol. 5, no. 4, p. 104, 2021, doi: 10.3390/drones5040104.

[33] Y. Megahed, A. Shaker, and W. Y. Yan, “Fusion of airborne lidar point clouds and aerial images for heterogeneous land-use urban mapping,” *Remote Sens.,* vol. 13, no. 4, pp. 1–36, Feb. 2021, doi: 10.3390/rs13040814.

[34] C. M. Bishop, Neural Networks for Pattern Recognition. Oxford: Oxford Univ. Press, 1995.

[35] H. Taud and J.-F. Mas, “Multilayer perceptron (MLP),” in Geomatic Approaches for Modeling Land Change Scenarios, Cham: *Springer,* 2017, pp. 451–455.

[36] A. Sen and K. Gumus, “Comparison of different parameters of feedforward backpropagation neural networks in DEM height estimation for different terrain types and point distributions,” *Systems,* vol. 11, no. 5, p. 261, 2023, doi: 10.3390/systems11050261.

[37] V. N. Ogar, S. Hussain, and K. A. A. Gamage, “Transmission line fault classification of multi-dataset using catboost classifier,” *Signals,* vol. 3, no. 3, pp. 468–482, 2022, doi: 10.3390/signals3030027.

[38] Z. Alenzi et al., “A semantic classification approach for indoor robot navigation,” *Electronics,* vol. 11, no. 13, 2022, doi: 10.3390/electronics11132063.

[39] D. Wang and H. Qian, “Catboost-based automatic classification study of river network,” *ISPRS Int. J. Geo-Inf.,* vol. 12, no. 10, 2023, doi: 10.3390/ijgi12100416.

[40] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 785–794.

[41] Z. Shao, M. N. Ahmad, and A. Javed, “Comparison of random forest and XGBoost classifiers using integrated optical and SAR features for mapping urban impervious surface,” *Remote Sens.*, vol. 16, no. 4, p. 665, 2024, doi: 10.3390/rs16040665.

[42] M. Kuprowski and P. Drozda, “Feature selection for airborne LiDAR point cloud classification,” *Remote Sens.*, vol. 15, no. 3, p. 561, 2023, doi: 10.3390/rs15030561.

[43] E. A. R. Hamed et al., “An efficient combination of convolutional neural network and LightGBM algorithm for lung cancer histopathology classification,” *Diagnostics*, vol. 13, no. 15, 2023, doi: 10.3390/diagnostics13152469.

[44] E. Sevgen and S. Abdikan, “Classification of large-scale mobile laser scanning data in urban area with LightGBM,” *Remote Sens.*, vol. 15, no. 15, p. 3787, 2023, doi: 10.3390/rs15153787.

[45] J. Ahmad, J. Sidén, and H. Andersson, “A proposal of implementation of sitting posture monitoring system for wheelchair utilizing machine learning methods,” *Sensors*, vol. 21, no. 19, p. 6349, 2021, doi: 10.3390/s21196349.

[46] G. Ke et al., “LightGBM: A highly efficient gradient boosting decision tree.” [Online]. Available: https://github.com/Microsoft/LightGBM

[47] S. Elmahdy, T. Ali, and M. Mohamed, “Regional mapping of groundwater potential in Ar Rub al Khali, Arabian Peninsula using the classification and regression trees model,” *Remote Sens.*, vol. 13, no. 12, p. 2300, 2021, doi: 10.3390/rs13122300.

[48] D. Korkmaz, H. E. Çelik, and M. Kapar, “Sınıflandırma ve regresyon ağaçları ile rastgele orman algoritması kullanarak botnet tespiti: Van Yüzüncü Yıl Üniversitesi örneği,” *Yüzüncü Yıl Üniv. Fen Bilim. Ens. Derg.*, c. 23, no. 3, ss. 297–307, 2018.

[49] Classification and Regression Trees (CART) Theory and Applications, 2004.

[50] A. Alan and M. Karabatak, “Veri seti-sınıflandırma ilişkisinde performansa etki eden faktörlerin değerlendirilmesi,” *Fırat Üniv. Müh. Bilim. Derg.*, c. 32, no. 2, ss. 531–540, 2020.

[51] G. Alparslan and M. Dursun, “Konvolüsyonel sinir ağları tabanlı Türkçe metin sınıflandırma,” *Bilisim Teknol. Derg.*, c. 16, no. 1, ss. 21–31, 2023, doi: 10.17671/gazibtd.1165291.

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirileri

S. Varbil ve A. Şen, "İç mekân nokta bulutlarının makine öğrenmesi yöntemleriyle sınıflandırılması," 19. Türkiye Harita Bilimsel ve Teknik Kurultayı, Ankara, Türkiye, ss.1, 2023

Makaleler

S. Varbil Ve A. Şen, "İç mekân harita yapımı için makine öğrenmesiyle nokta bulutlarının sınıflandırılması,"*Jeodezi ve jeoinformasyon dergisi* (Online), cilt.11, no.1, ss. 30-40, 2024