



Yapay Zeka 1: Veri Bilimi ve Makine Öğrenmesine Giriş

Doç. Dr. Taner Arsan–Ph.Dc. H. Fuat Alsan–Ph.Dc. Sena Kılınç

Bilgisayar Mühendisliği Bölümü

Mühendislik ve Doğa Bilimleri Fakültesi - Kadir Has Üniversitesi

Yapay Zeka 1 – Eğitim İçeriği

1. Python ile Programlama (Temel kodlamadan Nesneye Yönelik Programlama'ya kadar)
2. Veri Bilimine Giriş ve Temel Python Kütüphaneleri
3. Keşifsel Veri Analizi – Exploratory Data Analysis (EDA)
4. Makine Öğrenmesine Giriş
5. Denetimli Öğrenme (Supervised Learning)
6. Denetimsiz Öğrenme (Unsupervised Learning)
7. Model Değerlendirme ve Hiperparametre Optimizasyonu
8. Uygulamalı Final Projesi

(Kursta öğrenilen bilgileri gerçek hayatı bir yer uygulayıp pratik beceri kazanma)

Zeka'nın Tanımı

Zeka, temelde işleri ıçgündüsel veya otomatik olarak yapmak yerine düşünme ve anlama yeteneğidir. Düşünme, beyninizi bir problemi değerlendirmek veya bir fikir yaratmak için kullanma faaliyetidir.

Zekanın diğer bazı tanımları

- Deneyimden öğrenme yeteneği
- Düşünce gücü
- Akıl yürütme yeteneği
- İlişkileri algılama yeteneği
- Alet kullanma yeteneği
- Bilinç ve **duygusallık** içeren İçgörü gücü

Yapay Zeka

Organik zekanın makineler yoluyla taklidi

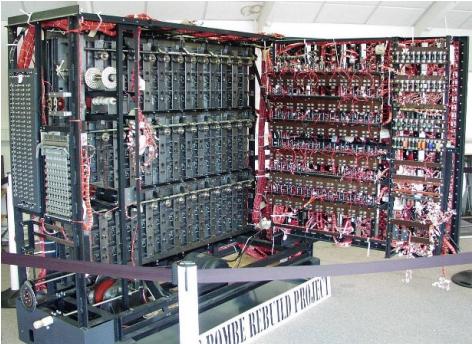
Yapay Zeka

İstenen işi başarmak için insanın **düşünme ve karar verme** yeteneğini taklit etmeye çalışan **akıllı sistemlerin** tasarıımı ve geliştirilmesi ile uğraşan bilim dalı.

- İnsanın zihinsel süreçlerini benzetimi için genelde bilgisayarlar kullanılır.
- İnsan dışı zihinsel süreçler tasarlanabilir.
 - ❖ İnsan zekası sadece olası bir zeka türüdür.
 - ❖ **Bir makinenin bir sorunu çözme yöntemi, insan yönteminden farklı olabilir,** ancak bu, daha az akıllı olduğu anlamına gelmez.

İnsan Zekasından Yapay Zekaya !

Basit bir soru ile başlayalım: **Makineler düşünebilir mi?**



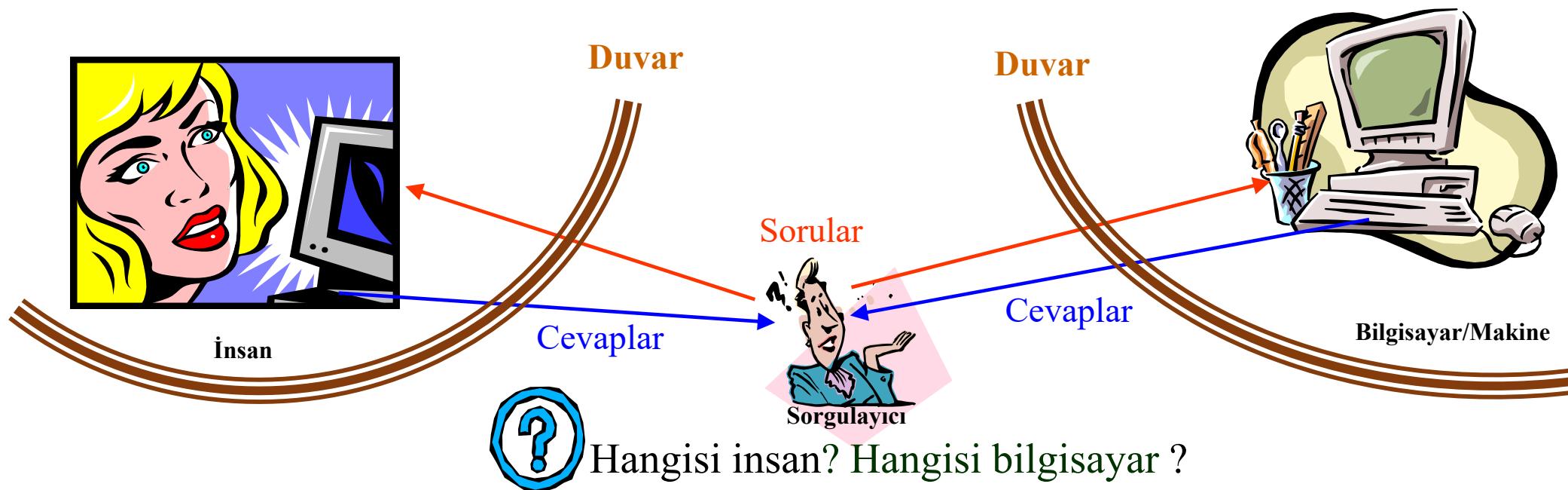
Alan Mathison Turing
(23 Haziran 1912 – 7 Haziran 1954),
İngiliz matematikçi, bilgisayar bilimcisi ve kriptolog
II. Dünya Savaşı'nı iki yıl kısaltan adam! 14M Avrupalı!
Bombe; tam otomatikleştirilmiş Enigma kodlarını kırmak makinesi



Turing Testi

Turing testinin amacı, bir makinenin düşünebildiğini söyleyebilmenin mantıksal olarak mümkün olup olmadığıdır.

Turing testine göre makine, gönüllü bir insanla birlikte, sorgulayıcının görüş alanının dışında bir yere saklanır. Sorgulayıcı yalnız soru sormak suretiyle hangisinin insan hangisinin bilgisayar olduğunu saptamaya çalışır. Sorgulayıcının soruları ve daha önemli aldığı yanıtlar, tamamen ses gizlenerek, yani ya bir klavye sisteminde yazılarak veya bir ekranда gösterilerek verilir. Sorgulayıcıya bu soru-cevap oturumunda elde edilen bilgiler dışında her iki taraf hakkında hiçbir bilgi verilmez. Dizi halinde tekrarlanan testler sonucunda sorgulayıcı, tutarlı bir şekilde insanı saptayamadığı takdirde makine **Turing testini** geçmiş sayılır.



Yapay Zeka ve Terminolojisi

Yapay Zeka

Organik zekanın makineler yoluyla taklidi

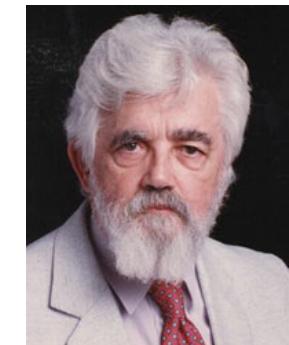
Professor John McCarthy - Stanford University ~1967



This picture is used for research, teaching and education and it is a transformative use. We consider that the use of this picture in this context represents fair use. Please, do not try to take down this picture before considering whether our conduct constitutes fair use!

Yapay Zeka Bilimi:

İstenen işi başarmak için insanın düşünme ve karar verme yeteneğini taklit etmeye çalışan akıllı sistemlerin tasarıımı ve geliştirilmesi ile uğraşan bilim dalı.



John McCarthy (4 Eylül 1927-24 Ekim 2011),
[Amerikalı bilgisayar bilimci](#)

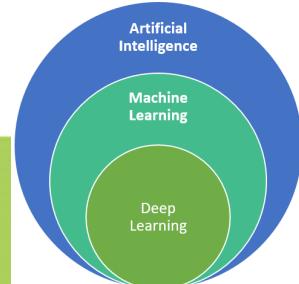
Yapay Zeka – AI
Artificial Intelligence

*

*

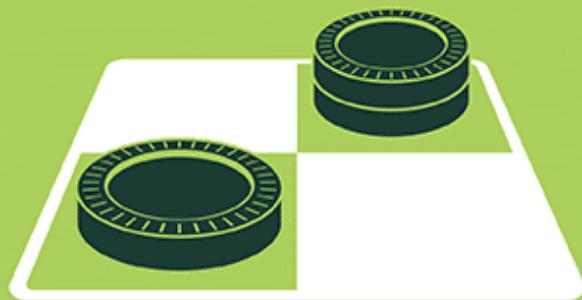
Hesaplamalı Zeka – CI
Computational Intelligence

Yapay Zeka - Makine Öğrenmesi – Derin Öğrenme



ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Yapay Zeka – Makine Öğrenmesi – Derin Öğrenme

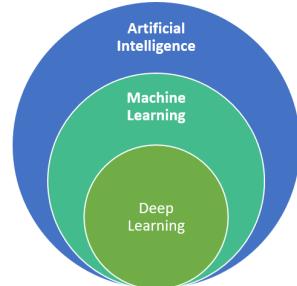
- **Yapay zeka:** Makinelerin zeka gerektiren görevleri gerçekleştirebilmesi

- **Makine öğrenimi:** bilgisayarların açıkça programlanmadan görevleri yerine getirebilmesidir.
- Örnek: bir arabayı tanıtmak için her özelliği manuel olarak kodlamaya gerek yoktur. Fakat **arabanın örnek olarak yeterli miktarda gösterilmesi** gereklidir.
- Yani çalışılacak alan hakkında yeterli veriye ihtiyaç vardır ve bu verilerin etiketlenmesi gereklidir. Ekleme süreci uzun ve maaliyetlidir (denetimli öğrenme).

- Derin öğrenme, makine öğrenmesi başlığı altında bir yöntemdir.
- Veriden özellik çıkartma işlemi (feature extraction) **makine öğrenmesinde insanlar tarafından yapılır**, derin öğrenmede **bu işlem otomatiktir**.

- (Klasik) **Makine öğrenmesine örnekler:** Linear/Polinom/Lojistik regresyon, Karar ağacı (decision tree), Destek vektör makinaları (support vector machines).
- Derin öğrenme yapay sinir ağları (**artifical neural network**) yaygındır ve çok fazla örnek veri gerektirir.

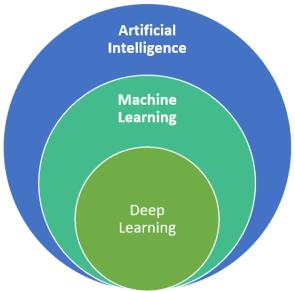
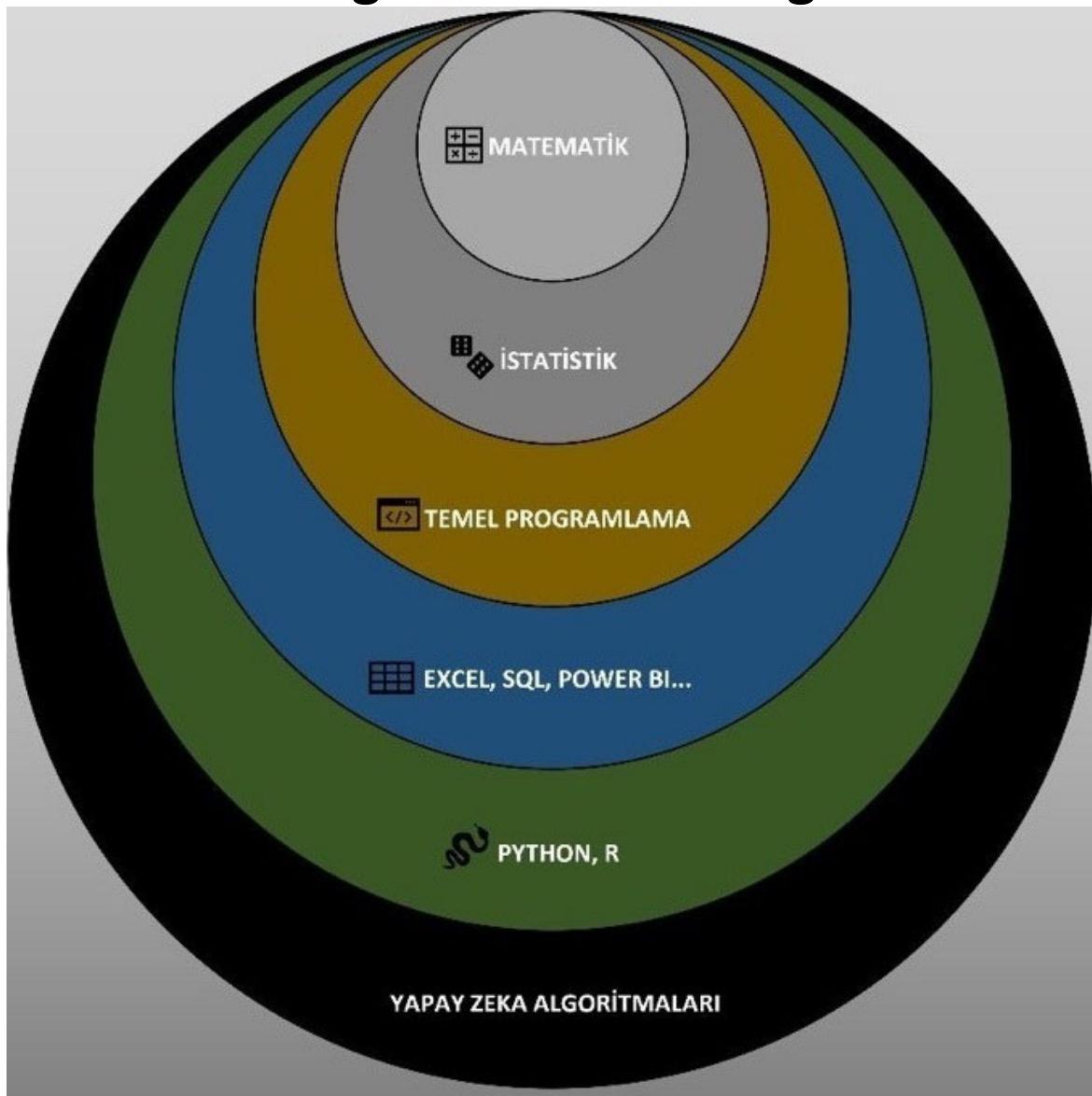
Yapay Zeka - Makine Öğrenmesi – Derin Öğrenme



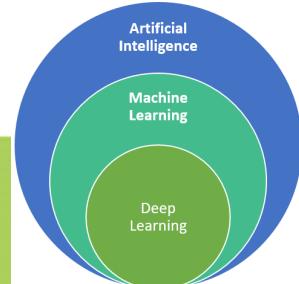
DEEP LEARNING AND MACHINE LEARNING

MACHINE LEARNING	DEEP LEARNING
Yapay Zekanın Alt Kümesi	Makine Öğrenmesinin Alt Kümesi
Küçük veri kümeleri ile eğitilir	Büyük veri kümeleri gereklidir
Düzelte ve öğrenme için insan desteği gereklidir	Önceki hatalardan kendi kendine öğrenir
Eğitim süresi kısa, doğruluk düşüktür	Eğitim süresi uzun, ancak doğruluk yüksektir

Yapay Zeka'ya Giden Yol . . . Long and the Winding Road

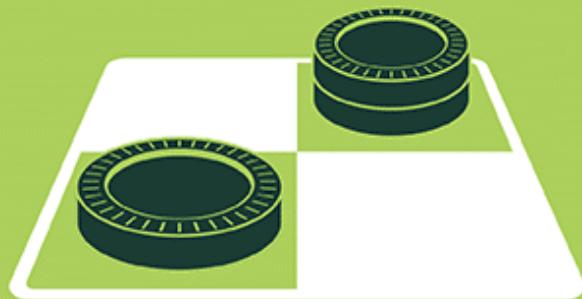


Yapay Zeka - Makine Öğrenmesi – Derin Öğrenme



ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.

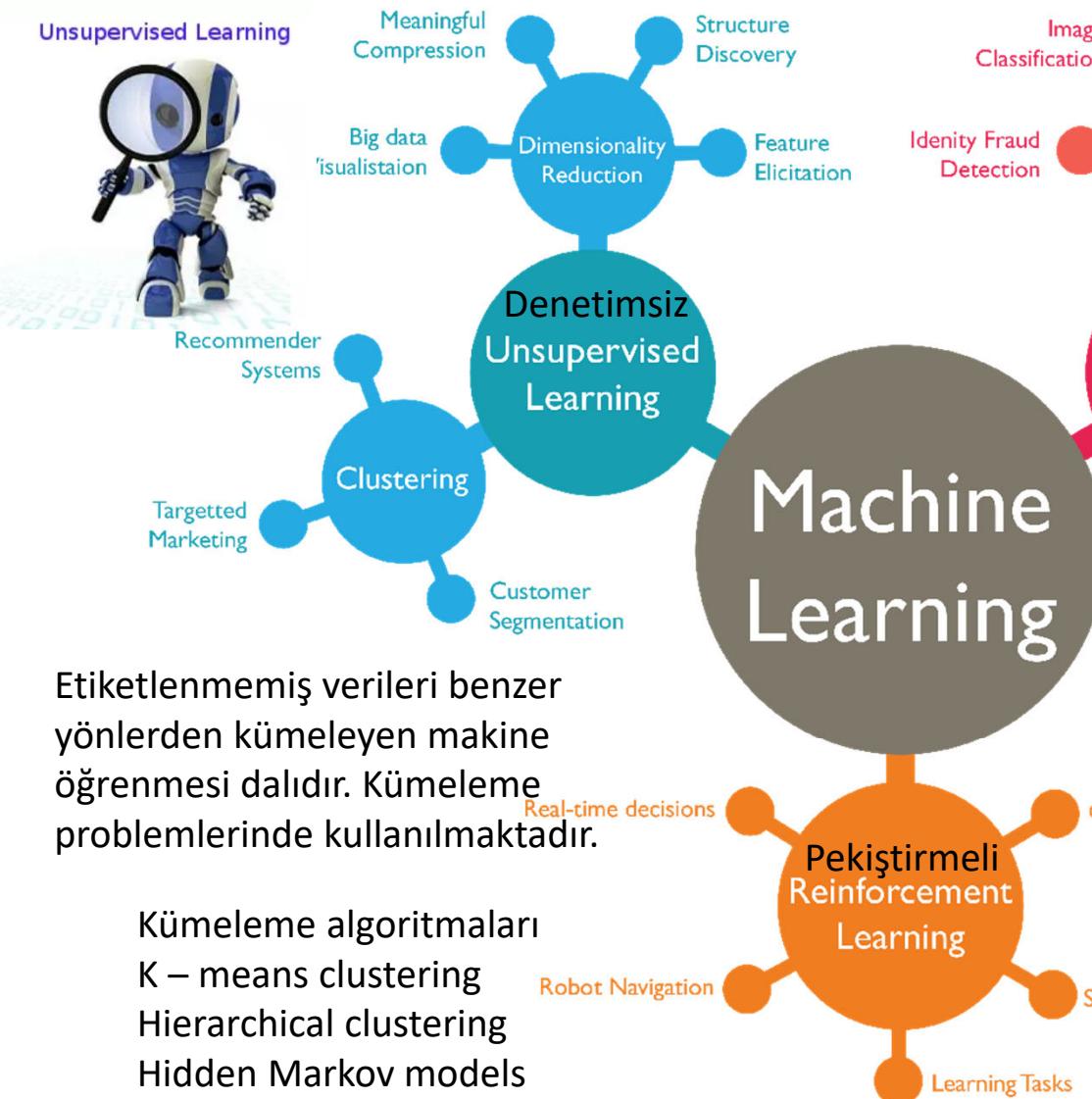


DEEP LEARNING

Deep learning breakthroughs drive AI boom.

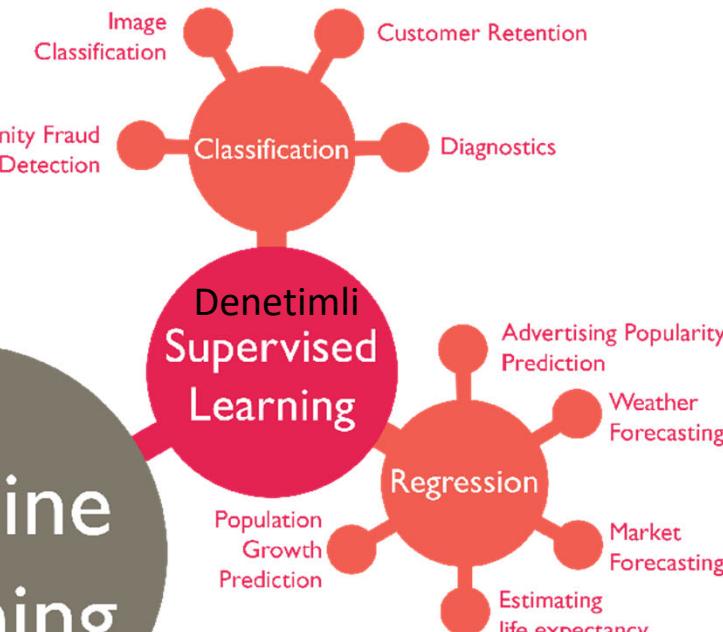


Makine Öğrenmesi



Etiketlenmemiş verileri benzer yönlerden kümeleyen makine öğrenmesi dalıdır. Kümeleme problemlerinde kullanılmaktadır.

Kümeleme algoritmaları
K – means clustering
Hierarchical clustering
Hidden Markov models



Algoritma içinde bulunduğu dinamik ortamla ilişki içindedir ve belirli bir amacı gerçekleştirmek için problem uzayında dolaşırken ödül ve ceza sistemi ile geribesleme alır.

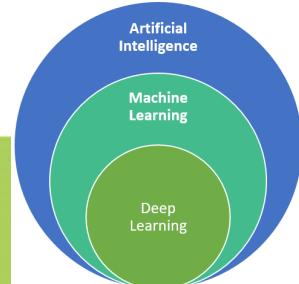


Supervised Learning

Etiketli verilerden eğitim ve test kümesi oluşturarak eğitim kümesinden bir model oluşturulmasını sağlayan ve modelin performansını test kümesi üzerinden ölçen makine öğrenmesi çeşididir. Sınıflandırma ve regresyon problemlerinde kullanılır.

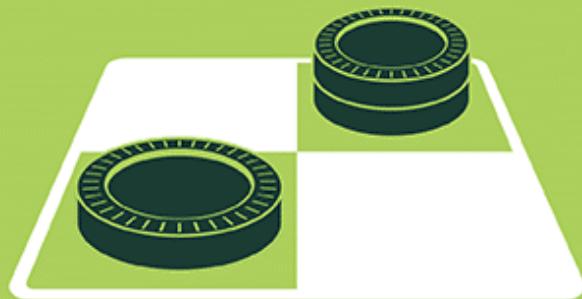
Sınıflandırma ve Regresyon
Logistic Regression
Decision trees
Support vector machine (SVM)
k-Nearest Neighbors
Naive Bayes
Random forest
Linear regression
Polynomial regression
SVM for regression

Yapay Zeka - Makine Öğrenmesi – Derin Öğrenme



ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Yapay Zekaya Neden İhtiyaç var? .:CCTV Örneği:.



15 Yıl Önce

Yapay Zekaya Neden İhtiyaç var? .:CCTV Örneği:.



Güvenlik Personeli
Çoklu ekranda **12 dakika** boyunca
kayıt izlediğinde bu kayıtların
%45'ini kaçırdığı tespit edildi.

22 dakika sonra bu oran **%95**

Yapay Zekaya Neden İhtiyaç var? ~~CCTV~~ Video Analitik



Yapay Zekaya Neden İhtiyaç var? ~~CCTV~~ Video Analitik



Video Analistik'e Geçiş Süreci

- * Kamera teknolojileri
- * Sıkıştırılmış kayıt teknolojileri
- * Verinin etiketlenmesi ve sınıflandırılması
- * Veri işleme hızının artması
- * Bilgisayar Ağ hızının artması - video/ses/data



IP Camera



Analog Camera



Multi Megapixel



Main



Amaç: Beyaz x marka aracın havaalanına ne zaman geldiği bulmak

Video Analitik Çözümleri

Nesne Sınıflandırma

Renk

Tip

Çoklu arama



Amaç: Siyah araçtan çıkan siyah kıyafetli sürücüyü bulmak

Video Analitik Çözümleri

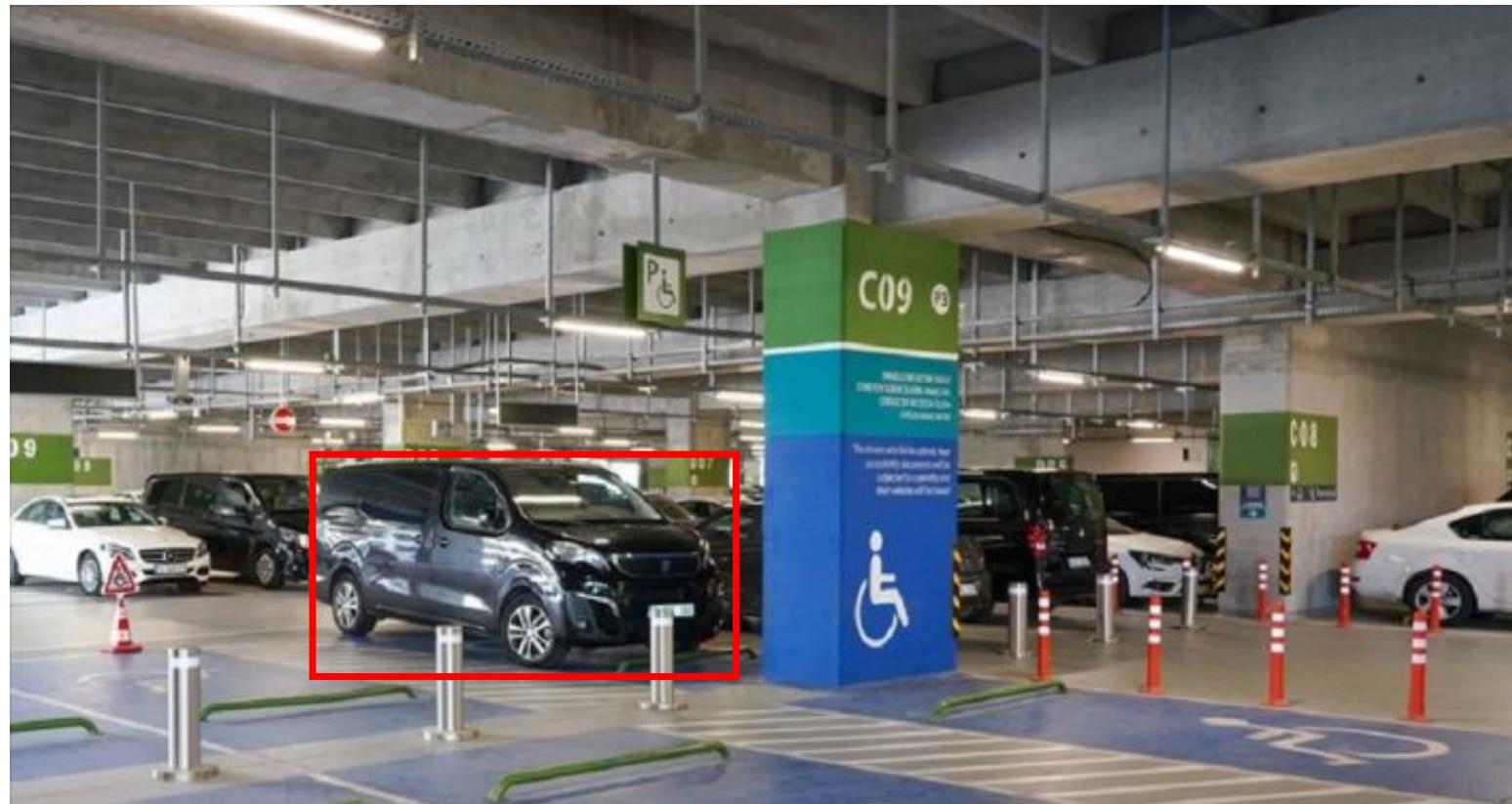
Nesne Sınıflandırma

Renk

Tip

Cinsiyet (Gender)

Çoklu arama



Amaç: Beyaz t-shirt'lu, mor renkli çantalı yolcuyu bulmak

Video Analitik Çözümleri

Nesne Sınıflandırma

Renk

Tip

Cinsiyet (Gender)

Çoklu arama



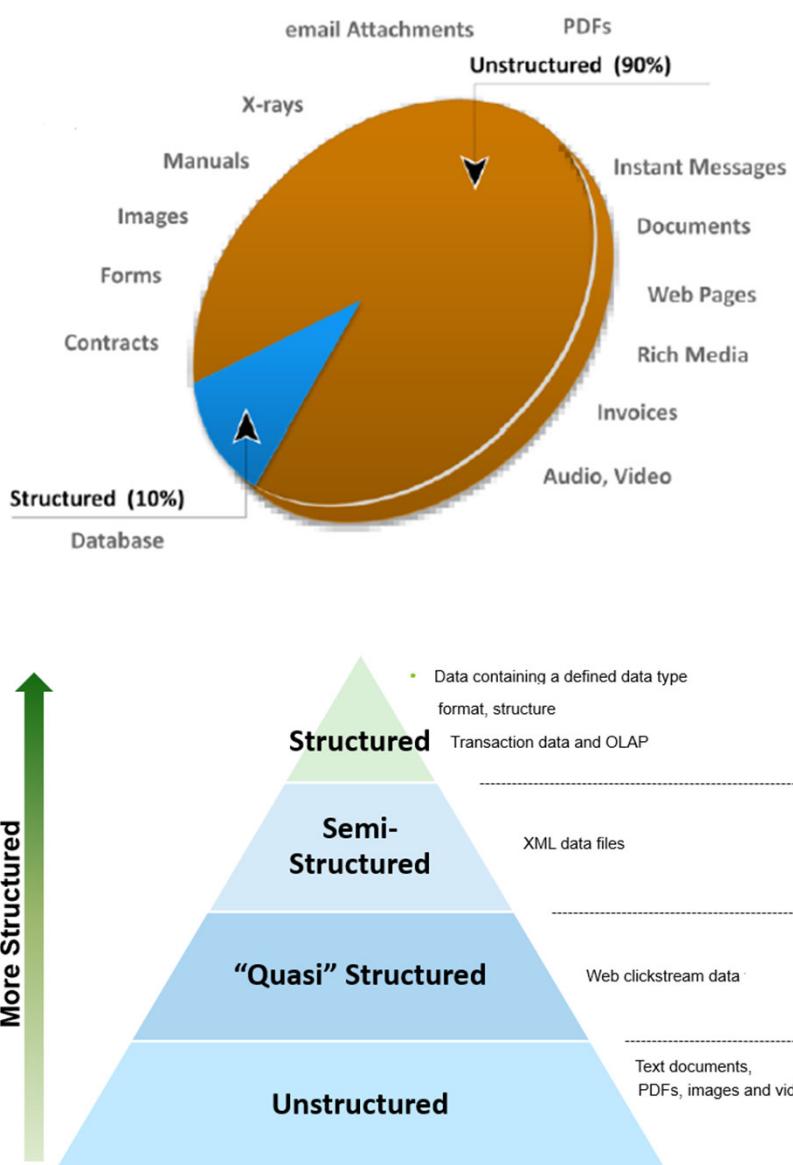
Veri Bilimine Giriş – Büyük Veri

Büyük Veri (Big Data), hacmi hızlıca artan yeni veri kaynaklarından elde edilen daha büyük, daha karmaşık veri kümeleridir.

Bu veri kümeleri o kadar hacimlidir ki geleneksel veri işleme yazılımı onları bu verilerle başa çıkamaz.

2005 yılında, Facebook, YouTube ve diğer çevrimiçi programları kullananların çok fazla veri ürettiği fark edildi. Aynı yıl Hadoop geliştirildi. Hadoop'un amacı büyük veri kümelerini depolamak ve analiz etmek için özel olarak oluşturulmuş açık kaynaklı bir yazılım.

Ancak bu büyük hacimli veriler, daha önce üstesinden gelemeyeceğiniz iş sorunlarını çözmek için kullanılabilir.

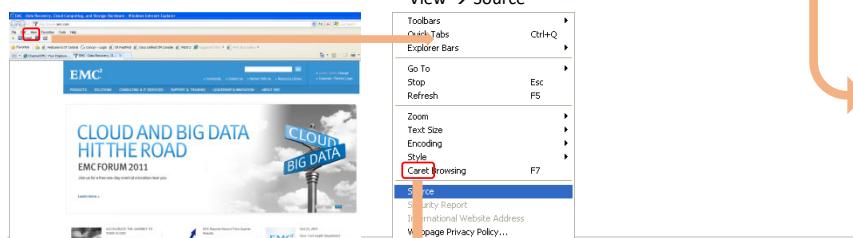


Four Main Types of Data Structures

Structured Data

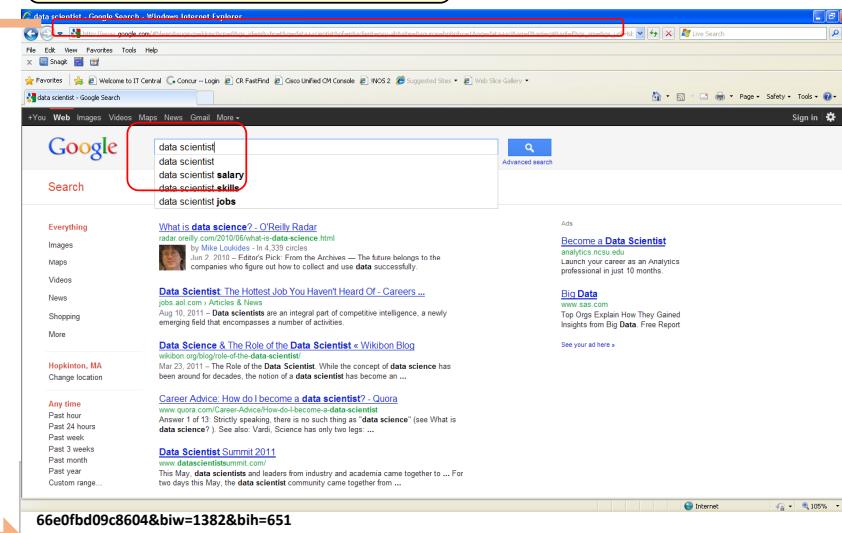
SUMMER FOOD SERVICE PROGRAM 1]				
(Data as of August 01, 2011)				
Fiscal Year	Number of Sites	Peak (July) Participation	Meals Served	Total Federal Expenditures 2)
1969	1.2	99	2.2	0.3
1970	1.9	227	8.2	1.8
1971	3.2	569	29.0	8.2
1972	6.5	1,080	73.5	21.9
1973	11.2	1,437	65.4	26.6
1974	10.6	1,403	63.6	33.6
1975	12.0	1,785	84.3	50.3
1976	16.0	2,453	104.8	73.4
TO 3)	22.4	3,455	198.0	88.9
1977	23.7	2,791	170.4	114.4
1978	22.4	2,333	120.3	100.3
4079	22.4	2,429	124.8	108.6
			2	110.1

Semi-Structured Data



```
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
    <META name="y_key" content="59b4020e1c9acec">
        <link rel="canonical" href="http://www.emc.com/index.htm" />
            <META NAME="verify-v1" CONTENT="y1Zc9V0P4eV0jFd1PeVVifRP32g4qtwFE0I2UvTMfSU
<title>EMC - Data Recovery, Cloud Computing, and Storage Hardware</title>
<META NAME="description" CONTENT="EMC is a leading provider of storage hardware solutions t</title>
data recovery and improve cloud computing." />
        <!-- Start :styleSheet includes -->
<link rel="stylesheet" href="/_admin/css/styles.css" />
<link rel="stylesheet" href="/_admin/css/styles_nav.css" />
<!---[ IE]>
```

Quasi-Structured Data



Unstructured Data

The Red Wheelbarrow, by William Carlos Williams

so much depends
upon
a red wheel
barrow
glazed with rain
water
beside the white
chickens.



Veri Bilimi ve Veri Bilimci

Veriler üzerinde iş için anlamlı öngörüler elde edebilmek amacıyla gerçekleştirilen çalışmalara **Veri Bilimi** denir.

Büyük miktardaki verileri analiz etmek için matematik, istatistik, yapay zeka ve bilgisayar mühendisliği alanlarının ilke ve uygulamalarını bir araya getiren, disiplinler arası bir yaklaşımdır.

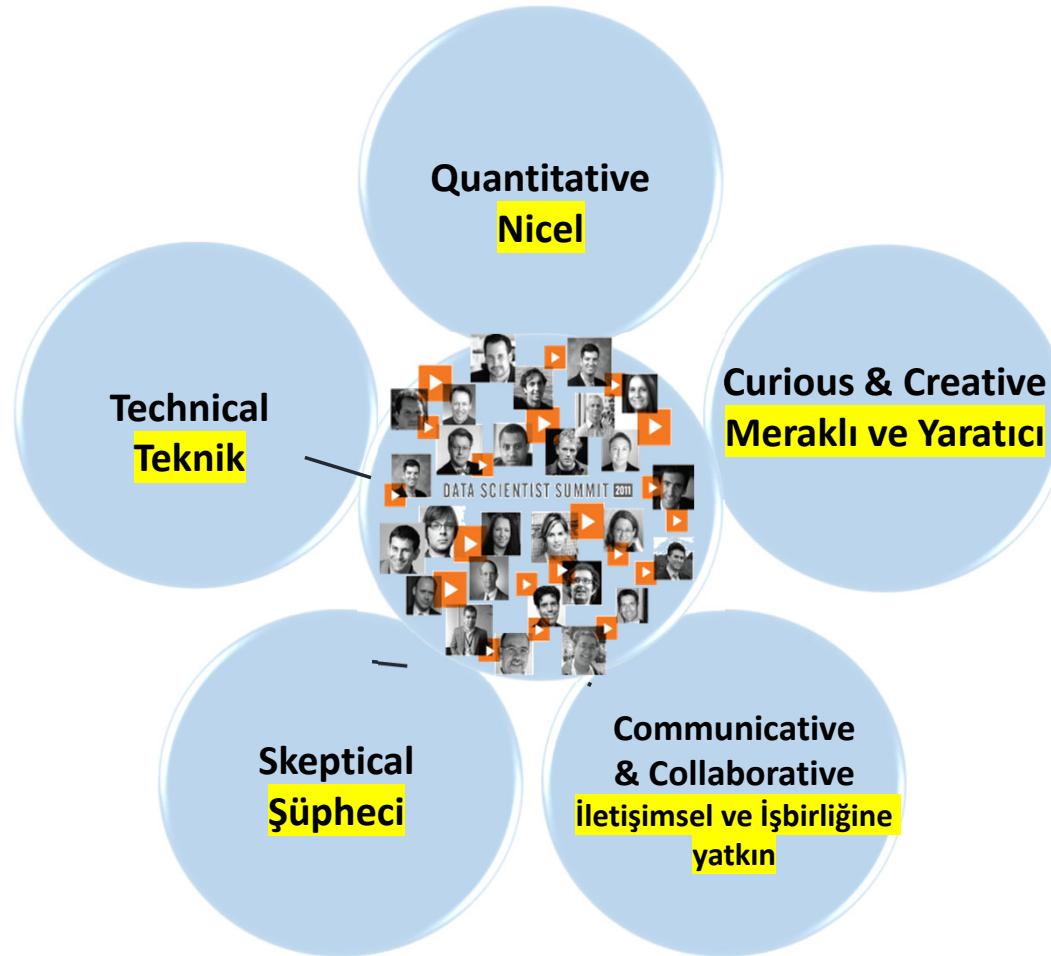
Verileri analiz etmek için strateji geliştiren, verileri analiz için hazırlayan, verileri modelleme, analiz etme ve görselleştirme, Python ve R gibi programlama dillerini kullanarak verilerle birlikte modeller oluşturma ve modelleri uygulama yazılımları için hazırlayan uzmana **Veri Bilimci** denir.

Veri bilimciler ekipler halinde çalışırlar ve bu ekiple veri bilimciye ek olarak, sorunu tanımlayan bir **İş Analisti**, verileri hazırlayan ve verilere nasıl erişileceğini belirleyen bir **Veri Mühendisi**, altta yatan süreçlere ve altyapıya ilişkin gözetim faaliyetlerini yürüten bir **BT Uzmanı** ve modelleri veya analiz sonuçlarını uygulama yazılımlarına yerlestiren bir **Uygulama Yazılımı Geliştiricisinden** oluşur.

Veri Bilimi Projesinin Anahtar Rollerİ

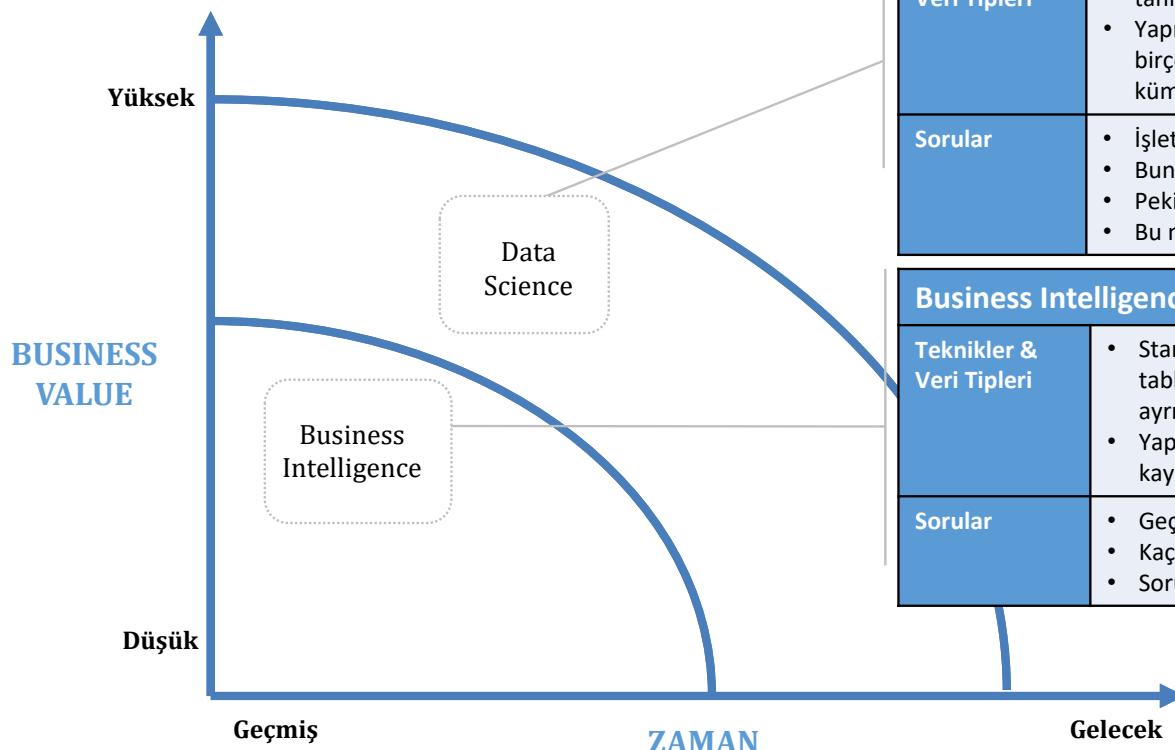
Rol	Tanım
Şirket Çalışanı	Nihai sonuçlardan yararlanan ve proje ekibine nihai sonuçların değeri ve bunların nasıl operasyonel hale getirileceği konusunda danışmanlık yapabilecek ve tavsiyelerde bulunabilecek kişi.
Proje Sponsoru	Projenin doğusundan sorumlu olan, projeye ve temel iş sorununa ivme kazandıran kişi, genellikle finansmanı sağlar ve çalışma ekibinin nihai çıktılarının değer derecesini ölçecektir.
Proje Yürüttürücüsü	Önemli kilometre taşılarının ve hedeflerin zamanında ve beklenen kalitede olmasını sağlar.
İş Zekası Analisti	Verilerin, key performance indicator KPI'ların, temel ölçümlerin ve iş zekasının raporlama perspektifinden derinlemesine anlaşılmasını sağlayan uzman
Veri Mühendisi	Veri yönetimi, çıkarma ve analitik sanal alana veri alımını desteklemek için SQL sorgularının ayarlanması yardımcı olacak derin teknik becerilere sahip olan mühendis
Veritabanı Yöneticisi (DBA)	Çalışma ekibinin analitik ihtiyaçlarını desteklemek için veritabanı ortamını hazırlayan ve yapılandıran yönetici kadrosunda uzman.
Veri Bilimcisi	Analitik teknikler, veri modelleme, geçerli analitik tekniklerin belirli iş sorunlarına uygulanması ve genel analitik hedeflerin karşılanması sağlanmak için konu uzmanlığı sağlamak.

Veri Bilimci'nin Profili



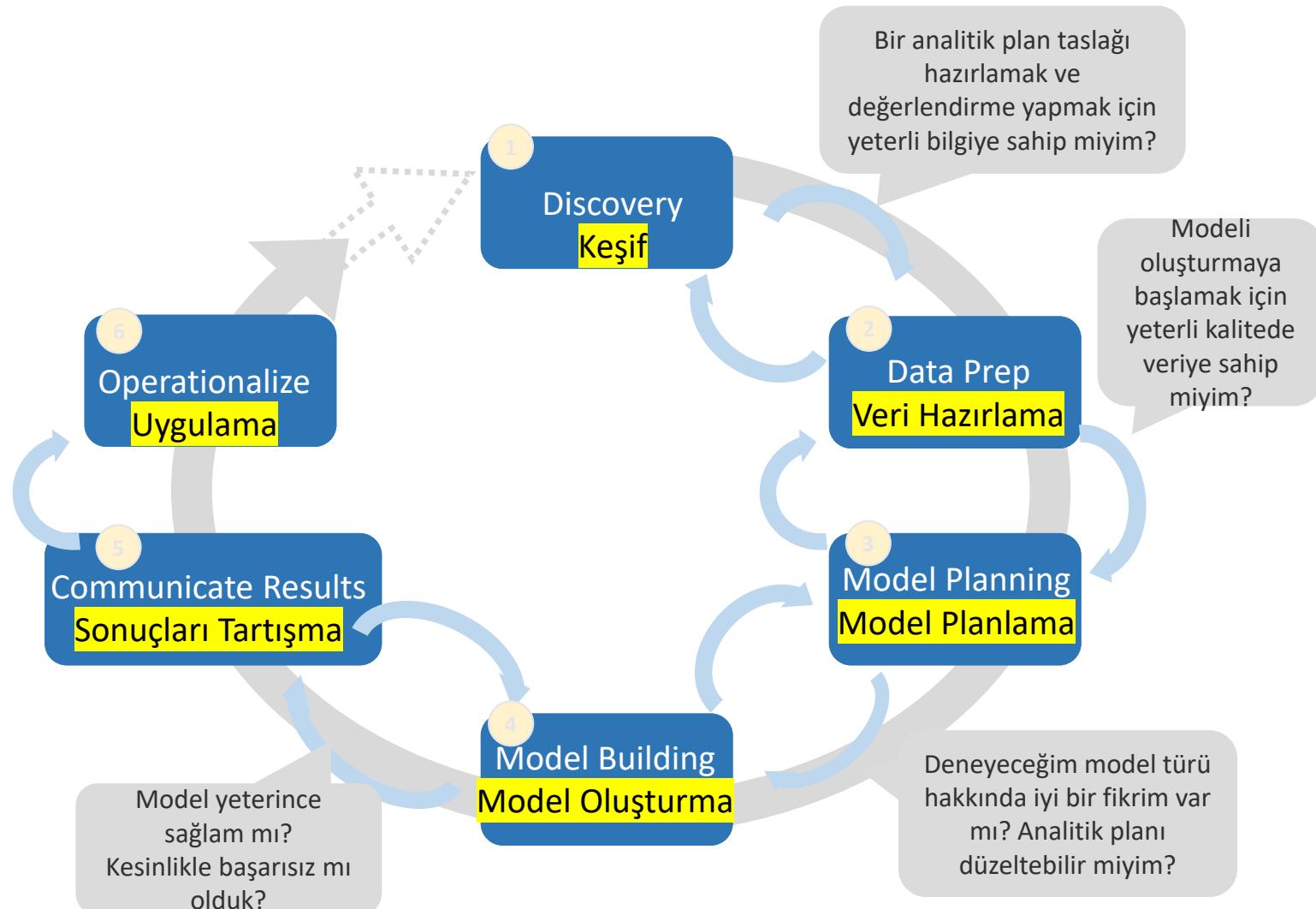
Veri Bilimi – İş Zekası

Business Intelligence vs. Data Science



Predictive Analytics & Data Mining (Data Science)	
Teknikler & Veri Tipleri	<ul style="list-style-type: none">Optimizasyon, tahmine dayalı modelleme, tahmin, istatistiksel analizYapilandırılmış/yapilandırılmamış veriler, birçok kaynak türü, çok büyük veri kümeleri
Sorular	<ul style="list-style-type: none">İşletmemiz için en uygun senaryo nedir?Bundan sonra ne olacak?Peki ya bu eğilimler devam ederse?Bu neden oluyor?
Business Intelligence – İş Zekası	
Teknikler & Veri Tipleri	<ul style="list-style-type: none">Standart ve özel raporlama, göstergeler tabloları, uyarılar, sorgular, isteğe bağlı ayrıntılarYapilandırılmış veriler, geleneksel kaynaklar, yönetilebilir veri setleri
Sorular	<ul style="list-style-type: none">Geçen çeyrekte ne oldu?Kaç tane sattık?Sorun nerede? Hangi durumlarda?

Data Analytics Lifecycle – Veri Analitiği Yaşam Döngüsü



Role Göre Başarılı Bir Analitik Projenin Temel Çıktıları

Rol	Tanım	Nihai Çıktılarda Rolün İhtiyacı Olanlar
Şirket Çalışanı	Nihai sonuçlardan yararlanan ve proje ekibine nihai sonuçların değeri ve bunların nasıl operasyonel hale getirileceği konusunda danışmanlık yapabilecek ve tavsiyelerde bulunabilecek kişi.	<ul style="list-style-type: none"> Sponsor Sunumunun işaret ettiğleri: Sonuçlar benim için iyi mi? Bulguların faydaları nelerdir? Bunun benim için sonuçları nelerdir?
Proje Sponsoru	Projenin doğusundan sorumlu olan, projeye ve temel iş sorununa ivme kazandıran kişi, genellikle finansmanı sağlar ve çalışma ekibinin nihai çıktılarının değer derecesini ölçecektir.	<ul style="list-style-type: none"> Sponsor Sunumunun işaret ettiğleri: Bunu yapmanın ticari etkisi nedir? Riskler nelerdir? Yatırım getirisi? Bu, organizasyon içinde (ve ötesinde) nasıl yaygınlaştırılabilir?
Proje Yürüttücsü	Önemli kilometre taşılarının ve hedeflerin zamanında ve beklenen kalitede olmasını sağlar.	
İş Zekası Analisti	Verilerin, key performance indicator KPI'ların, temel ölçümllerin ve iş zekasının raporlama perspektifinden derinlemesine anlaşılmasılığını sağlayan uzman	<ul style="list-style-type: none"> Analist sunumunu göster Raporların değişip değimeyeceğini belirleyin
Veri Mühendisi	Veri yönetimi, çıkarma ve analitik sanal alana veri alımını desteklemek için SQL sorgularının ayarlanmasına yardımcı olacak derin teknik becerilere sahip olan mühendis	<ul style="list-style-type: none"> Analitik projedeki kodu paylaşın. Nasıl uygulanacağına ilişkin teknik belge oluşturun.
Veritabanı Yöneticisi (DBA)	Database Administrator who provisions and configures database environment to support the analytical needs of the working team	<ul style="list-style-type: none"> Analitik projedeki kodu paylaşın. Nasıl uygulanacağına ilişkin teknik belge oluşturun.
Veri Bilimcisi	Analitik teknikler, veri modelleme, geçerli analitik tekniklerin belirli iş sorunlarına uygulanması ve genel analitik hedeflerin karşılanmasılığını sağlamak için konu uzmanlığı sağlamak.	<ul style="list-style-type: none"> Analist sunumunu göster Kodu paylaş

Teşekkürler...

Doç. Dr. Taner ARSAN
arsan@khas.edu.tr
0532 651 65 87