



Derleme Makalesi

Büyük veri analitiği için yüksek performans hesaplama: çözüm ortamları ve kodlama

Hamza Erol*¹

¹Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye

Anahtar Kelimeler:

Küresel ekonomi
Ekonominin akıllı yönetimi
Akan büyük veri
Çok katmanlı büyük veri
Büyük verinin hacmi
Büyük verinin karmaşıklığı
Anaconda
NetworkX
PySpark
PyCuda
Google Colabratory

ÖZ

Bu çalışmada akan veya katmanlı büyük veri, hacmi ve karmaşıklığı tanımlanmıştır. Büyük veri analitiğinde kullanılan; Tek işlemci- tek çekirdek yapısı: standart hesaplama mimarisi; Tek işlemci- çok çekirdek yapısı: paralel hesaplama mimarisi ve Çok işlemci- çok çekirdek yapısı: dağıtık hesaplama mimarisi biçimlerdeki yüksek performans hesaplama sistemleri açıklanmıştır. Bu sistemlerde kullanılan hesaplama ortamları incelenmiştir. Donanım hesaplama ya da donanım hızlandırılmış hesaplama ve Yazılım hesaplama ya da yazılım optimize edilmiş hesaplama konuları vurgulanmıştır. Büyük veri analitiğinde uygulanan öğrenme yöntemleri: istatistiksel öğrenme, makine öğrenme ve derin öğrenme ifade edilmiştir. Bu öğrenme yöntemlerinin sonuç veya ürün odaklı uygulamaları olarak yapay zekâ açıklanmıştır. Akan veya katmanlı büyük veri analitiğinde web tabanlı çözüm ortamı olarak Google Colabratory ve bu ortamda kod geliştirmede Python uygulamaları verilmiştir.

High performance computing for big data analytics: solution environments and coding

Keywords:

Global economies,
Intelligent management of the economy
Stream big data
Multilayer big data
Size of big data
Complexity of big data
Anaconda
NetworkX
PySpark
PyCuda
Google Colabratory

ABSTRACT

Stream or multilayered big data, size and complexity is defined in this study. High performance computing systems used in big data analytics; Single processor - single core structure: standard computing architecture; Single processor - multi core structure: parallel computing architecture and Multi processor - multi core structure: distributed computing architecture are explained. Computing environments used in these systems have been examined. Hardware computing or hardware accelerated computing and software computing or software optimized computing are emphasized. Learning methods applied in big data analytics: statistical learning, machine learning and deep learning were expressed. Artificial intelligence has been explained as the result or product-oriented applications of these learning methods. Google Colabratory as a web-based solution environment in stream or multilayered big data analytics and Python applications for code development in this environment have been given.

*Sorumlu Yazar

(herol@mersin.edu.tr) ORCID ID 0000 - 0001 - 8983 - 4797

1. GİRİŞ

Dünya ekonomilerinde tüketim için yeterli miktarda hızlı ve fazla sayıda endüstriyel üretim amacıyla Endüstri 4.0 aşamasına geçmiştir. Endüstri 4.0 endüstriyel üretim için robot teknolojilerini, uydu teknolojilerini ve sensör teknolojilerini kullanır. Büyük veri endüstriyel veridir (Gökalp vd., 2016). Endüstri 4.0 ile büyük veri kavramı küresel dijital ekonominin akıllı yönetiminde kullanılmaya başlanmıştır. Günümüzde ekonominin akıllı yönetimi büyük anlam ve önem kazanmıştır. Büyük verinin ekonomik değeri hesaplanabilir. Büyük veri sensörler, robotlar, uydu teknolojileri ve akıllı teknolojiler kullanılarak toplanır. Analog olarak alınan veriler datalogger kullanılarak sayısallaştırılır. Büyük veri katmanlardan oluşur. Büyük veri katmanlı veri olarak tanımlanır.

Büyük verinin özellikleri: (i) Genişliğine büyük veri: veride değişken sayısının fazla olması durumudur. (ii) Derinliğine büyük veri: veride gözlem sayısının fazla olması durumudur. (iii) Hem genişliğine hem de derinliğine büyük veri. (iv) Akan büyük veri: sürekli derinliğine gözlem değeri alınması durumudur. (v) Katmanlı büyük veri: değişken özellik gruplarının birleşiminden oluşan veridir. Büyük verinin katmanlı olması, hacminin büyüklüğü ve karmaşıklığı önemlidir. Büyük veri analitiğinde karmaşıklık: (i) Veri karmaşıklığı (ii) Hesaplama karmaşıklığı ve (iii) Bilgi karmaşıklığı olmak üzere incelenir (Erol ve Erol, 2018). Büyük veri analitiği; veri analizini ve veri modellemeyi kapsar. Büyük veri analitiği: (1) Verinin analitik için hazırlanması, (2) Verideki değişkenlerin ilişkilendirilmesi, (3) Verinin şekillendirilmesi, (4) Veriye anlam kazandırılması ve (5) Verinin bilgiye dönüştürülmesi aşamalarından oluşur (Erol ve Korkmaz, 2020).

Büyük veri analitiği, küresel ekonomilerde şirketlerin daha iyi kararlar alması için önemli ilişkisel yapıları, modelleri, eğilimleri ve tercihleri keşfetmek amacıyla büyük veri kümelerinde gelişmiş bilişim teknolojilerinin kullanılmasıdır. Endüstri 4.0'da büyük veri analitiği, bakım ve onarım işlemlerine ne zaman ihtiyaç duyulacağını tahmin etmek için üretim makinelerinden gelen sensör verilerinin analiz edildiği akıllı fabrikalar dahil birçok alanda rol oynar (Vassakis vd., 2018). Küresel ekonomilerinde üreticiler, bulut bilişim ve nesnelerin interneti platformları aracılığıyla sensörlerden topladıkları büyük miktardaki veriyi tedarik zinciri yönetiminin verimliliğini arttıracak yapısal kalıpları ortaya çıkarmak için büyük veri analitiğini kullanır (Thelin, 2020).

Büyük verinin önemini ve değerini Yuval Noah Harari yaşamın geleceği olarak ifade etmiştir. Bunu: "Geçtiğimiz yüzyıllarda dünyanın geçer akçesi toprak ve araziydi; fakat artık büyük veri ve verinin geleceği sadece insanın geleceğini değil, yaşamın geleceğini belirleyecek" şeklinde ifade etmiştir (İnternet Kaynağı 1, 2018).

2. YÖNTEM

2.1. Büyük Verinin Analitiğinin Küresel Ekonomilerde Kullanımı

Küresel ekonomilerde ekonominin akıllı yönetimini büyük veri analitiği kullanarak gerçekleştirenlerden biri de Avrupa Birliği ülkeleridir. Avrupa komisyonunun belirlediği ekonominin sektörlerinde 27 sektör bulunmaktadır. Tüm sektörlerde büyük veri analitiği kullanılmaktadır (İnternet Kaynağı 2, 2020). Avrupa komisyonunun hazırladığı ekonominin sektörleri listesinde 1 ile 11 maddelerinin belirlenmesinde paydaşlar; 12'den 27'ye kadar olan sektörlerin belirlenmesinde paydaşların danışmanları katkı sağlamıştır.

2.2. Büyük Verinin Toplanması

Küresel ekonomilerinde ekonominin akıllı yönetiminde veri toplama önemli kavramdır. Büyük veri sensörler, robotlar, uydu teknolojileri ve akıllı teknolojiler kullanılarak toplanır. Analog olarak sensörler vasıtasıyla alınan veriler dataloggerlar kullanılarak sayısallaştırılır ve fiziksel disk üzerine aktarılır. Tarım sektöründe akıllı tarım uygulamaları çerçevesinde toplanan çok bantlı uydu görüntü verileri Erol vd. (2018) tarafından çalışılmıştır. Bu uydu görüntüsü yer istasyonunda bulunan görüntüleme sistemi kullanılarak alınmış ve daha sonra sayısallaştırılmıştır. Sayısallaştırılan uydu verisi akan büyük veridir. Büyük verinin analitiği yapılarak tarımsal bölgede tarımsal ürünlerin rekolte tahminleri yapılmıştır. Tarım sektöründe akıllı tarım uygulamaları çerçevesinde yine toplanan çok bantlı SpectroRadiometer sensör verileri Erol vd. (2017) tarafından çalışılmıştır. SpectroRadiometer sensör verisi bağlı olduğu bilgisayardaki yazılımla sayısallaştırılmıştır. Sayısallaştırılan SpectroRadiometer sensör verisi akan büyük veridir. Büyük verinin analitiği yapılarak tarımsal bölge yetiştirilen farklı mısır çeşitlerinde nişasta, yağ gibi farklı ürünlerin verimliliklerinin performans karşılaştırmaları yapılmış ve bu ürünlerin rekolte tahminleri gerçekleştirilmiştir.

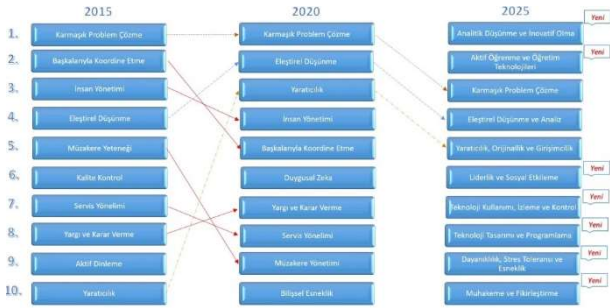
2.3. Büyük Verinin Analitiğinin Yapılmasında Mesleklerin Önemi

2019 yılında gerçekleştirilen 49. Dünya Ekonomik Formunda 2022 yılı için belirlenen en iyi 10 meslek Tablo 1'de verilmiştir (İnternet Kaynağı 3, 2019), (İnternet Kaynağı 4, 2018). Tablo 1'den de görüleceği gibi veri analitikçi ve büyük veri uzmanı meslekleri önem kazanan ve en iyi meslekler arasında yer almaktadır.

Tablo 1. 2019 yılında gerçekleştirilen 49. Dünya Ekonomik Formunda 2022 yılı için belirlenen en iyi 10 meslek.

Sıra No.	Meslek
1	Veri analitikçi / Veri bilimci
2	Yapay zekâ Makine öğrenme / Derin öğrenme uzmanı
3	Genel ve operasyonel yönetici
4	Yazılım ve uygulama geliştirici ve analizci
5	Satış ve pazarlama uzmanı
6	Büyük veri uzmanı
7	Dijital dönüşüm uzmanı
8	Yeni / yenilikçi teknoloji uzmanı
9	Organizasyonel gelişim uzmanı
10	Bilgi teknoloji servisleri uzmanı

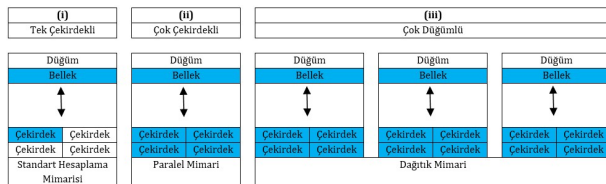
Veri analitikçi ve büyük veri uzmanı meslekleri büyük veri analitiği çerçevesinde küresel dünya ekonomilerinde ekonominin akıllı yönetimi için önemlidir. İş hayatında en etkin 10 yetenek Şekil 1'de gösterilmiştir (Internet Kaynağı 5, 2020)



Şekil 1. İş hayatında en etkin 10 yetenek.

2.4. Büyük Verinin Analitiğinde Kullanılan Yüksek Performans Hesaplama Sistemleri

Büyük veri analitiğinde kullanılan yüksek performans hesaplama sistemleri Şekil 2'de gösterilmiştir. Büyük verinin katman sayısı, hacminin büyüklüğü ve karmaşıklığı büyük veri analitiğinde önemlidir. Bunlar dikkate alınarak büyük veri analitiğinde 3 farklı mimaride yüksek performans hesaplama sistemi kullanılır. Bu sistemler: (i) Tek işlemci - tek çekirdek yapısı: standart hesaplama mimarisi; (ii) Tek işlemci- çok çekirdek yapısı: paralel hesaplama mimarisi ve (iii) Çok işlemci- çok çekirdek yapısı: dağıtık hesaplama mimarisi şeklindedir. Bu sistemlerde (i) Donanım hesaplama ya da donanım hızlandırılmış hesaplama ve (ii) Yazılım hesaplama ya da yazılım optimize edilmiş hesaplama kullanılır.



Şekil 2. Büyük veri analitiğinde kullanılan yüksek performans hesaplama sistemleri.

2.5. Yüksek Performans Hesaplama Sistemlerinde Kullanılan Yazılım Araçları ve Kod Geliştirme Ortamları

Büyük veri analitiğinde çalışan meslek grupları: (1) Veri bilimi, (2) Veri Mühendisliği, (3) Bilgisayar Mühendisliği, (4) Yazılım Mühendisliği, (5) Makine Öğrenme Mühendisliği, (6) Derin Öğrenme Mühendisliği, (7) Yapay Zekâ Mühendisliği, (8) Yapay Beyin Mühendisliği. Büyük veri analitiğinde kullanılan ücretsiz ve açık kaynak kodlu; Linux ve Windows tabanlı bazı yazılım araçları Tablo 2'de verilmiştir. Büyük veri analitiğinde kullanılan bu ücretsiz yazılım araçları veriden bilgi üretmek amacıyla kullanılır.

Tablo 2. Büyük veri analitiğinde kullanılan ücretsiz ve açık kaynak kodlu; Linux ve Windows tabanlı bazı yazılım araçları.

Linux Tabanlı		Windows Tabanlı	
Sıra No	Ücretsiz Yazılım Aracı	Sıra No	Ücretsiz Yazılım Aracı
1	R	1	R
2	RapidMiner	2	RapidMiner
3	MOA	3	Orange
4	Orange	4	Tanagra
5	ROOT	5	NeoNeuro Data Mining
6	Rattle	6	khcoder
7	ELKI	7	Weka
8	DataMelt		
9	KNIME		
10	Weka		
11	KEEL		

Büyük veri analitiğinde kullanılan ticari yazılım araçları Tablo 3'te verilmiştir. Büyük veri analitiğinde kullanılan bu ticari yazılım araçları veriden bilgi üretmek amacıyla kullanılır.

Tablo 3. Büyük veri analitiğinde kullanılan ticari bazı yazılım araçları.

Sıra No	Ticari Yazılım Aracı
1	Oracle Big Data Analytics
2	IBM Big Data Analytics
3	SAP Big Data Analytics
4	Alteryx Analytics
5	MicroStrategy Enterprise Big Data Analytics
6	SAS High-Performance Big Data Analytics
7	BIRT Analytics
8	Tableau Big Data Analysis
9	Qlik Big Data
10	TIBCO Spotfire

Büyük veri analitiğinde kullanılan bazı kod geliştirme ortamları Tablo 4'de verilmiştir. Büyük veri analitiğinde kullanılan bu kod geliştirme ortamları veriden bilgi üretmek amacıyla kullanılır.

Tablo 4. Büyük veri analitiğinde kullanılan bazı kod geliştirme ortamları.

Sıra No	Kod Geliştirme Alanı	Kod Geliştirme Ortamı	Kod Geliştirme Dili
1	Veri Analitiği Tek işlemci- çok çekirdek yapısı (Paralel hesaplama mimarisi)	Anaconda Jupyter Notebook	Python
2	Web Analitiği Tek işlemci- çok çekirdek yapısı (Paralel hesaplama mimarisi)	NetworkX	Python
3	Sosyal Medya Analitiği Tek işlemci- çok çekirdek yapısı (Paralel hesaplama mimarisi)	NetworkX	Python
4	Büyük Veri Analitiği Tek işlemci- çok çekirdek yapısı (Paralel hesaplama mimarisi)	PySpark	Python
5	Büyük Veri Analitiği Çok işlemci- çok çekirdek yapısı (Dağıtık hesaplama mimarisi)	PyCuda	Python
5	Büyük Veri Analitiği Çok işlemci- çok çekirdek yapısı (Dağıtık hesaplama mimarisi)	Google Colabratory	Python

3. BÜYÜK VERİ ANALİTİĞİ YÖNTEMLERİ VE UYGULAMALARI

Büyük veri analitiğinde yapılan her proje için (i) Strateji oluşturulur, (ii) Yol haritası belirlenir, (iii) Planlama yapılır ve (iv) Modelleme gerçekleştirilir. Büyük veri analitiğinde yapılacak işlemler: (1) Büyük verinin kümelenmesi, (2) Büyük verinin sınıflandırılması, (3) Büyük veride ilişkisel yapı modellerinin oluşturulması ve (4) Büyük verinin görselleştirilmesi.

Büyük veri analitiğinde bir proje gerçekleştirilirken takip edilmesi gereken adımlar: (1) Büyük verinin hazırlanması, (2) Büyük verinin katmanlı yapısına, hacmine ve karmaşıklığına göre 3 farklı hesaplama sisteminden hangisinin kullanılacağına ve 2 farklı programlamadan hangisinin uygulanacağına karar verilir, (3) Büyük veri analitiğinde hangi yazılımın kullanılacağına karar verilir, (4) Büyük veriye uygulanacak işleme göre 2 farklı öğrenme yönteminden: Makine Öğrenme / Derin öğrenme seçeneklerinden hangisinin uygulanacağına karar verilir, (5) Büyük veriye 4 farklı işlemten biri ya da birkaçı uygulandıktan sonra Yapay Zeka uygulaması geliştirilip geliştirilmeyeceğine karar verilir. (6) Büyük veri için bilgi oluşturulur. Bu adımlarla (i) Büyük veri analitik için hazırlanır, (2) Verideki

değişkenler ilişkilendirilir, (3) Büyük veri şekillendirilir, (4) Şekillendirilmiş büyük veriye anlam kazandırılır, (5) Büyük veri bilgiye dönüştürülür.

Yüksek performans sistemleri kullanılarak büyük veri analitiği uygulamaları: (1) Büyük verinin kümelenmesi, (2) Büyük verinin sınıflandırılması (Korkmaz ve Erol, 2020), (Gülgün ve Erol, 2020(a)), (Gülgün ve Erol 2020(a)), (3) Büyük veride ilişkisel yapı modellerinin oluşturulması (Erol ve Erol, 2020) ve (4) Büyük verinin görselleştirilmesi (Erol ve Erol, 2016) aşağıda alt başlıklar altında açıklanmıştır.

3.1. Google Colabratory Ortamında Büyük Verinin Kümelenmesi

Bu alt bölümde yüksek performans hesaplama sistemi kullanılarak Google Colabratory ortamında büyük verinin kümelenmesi incelenecektir (İnternet Kaynağı 6, 2021). Bu çalışmada eğitimsiz makine öğrenimi modellerinin bir kategorisi açıklanmıştır: boyut indirgeme. Burada, eğitimsiz makine öğrenim modellerinin başka bir sınıfı ele alınmıştır: kümeleme algoritmaları. Kümeleme algoritmaları, verilerin özelliklerinden içerdiği gruplarının optimal bir şekilde bölünmesini veya ayrı ayrı etiketlenmesini öğrenmeye çalışır. Birçok kümeleme algoritması mevcuttur, ancak belki de anlaşılması en basit olanı k-ortalama kümeleme olarak bilinen bir algoritmadır. Bu çalışmada Google Colabratory ortamında büyük verinin kümelenmesi uygulaması yapılmıştır.

3.2. Google Colabratory Ortamında Büyük Verinin Sınıflandırılması

Bu alt bölümde yüksek performans hesaplama sistemi kullanılarak Google Colabratory ortamında büyük verinin sınıflandırılması amaçlı iki çalışma incelenecektir.

Birinci çalışmada tıbbi görüntülerin sınıflandırılmasını içermektedir (İnternet Kaynağı 7, 2021). Birinci sınıflandırma uygulamasında tıbbi görüntüleme için derin öğrenme kullanılmıştır. Akciğer X-ışın görüntülerinin sınıflandırılması yapılmıştır. Bu, tıbbi görüntü sınıflandırması için makine öğrenimine üst düzey bir giriş niteliğindedir. Sınıflandırmada derin öğrenme yöntemi uygulanmıştır. Tıbbi görüntüler görüntülenmiş ve görüntü düzeyinde açıklamalar oluşturulmuştur. Sınıflandırma modelini eğitmek için veri kümeleri oluşturulmuştur. Bu çalışmada Google Colabratory ortamında büyük verinin sınıflandırılması uygulaması yapılmıştır.

İkinci çalışma hazır görüntülerin (İnternet Kaynağı 8, 2009) sınıflandırılmasını içermektedir (İnternet Kaynağı 9, 2019). Sınıflandırmada evrişimli sinir ağı (CNN) kullanılmıştır. Görüntülerin evrişimli sinir ağı (CNN) kullanılarak sınıflandırılmasında hem model oluşturulmuş hem de model eğitilmiştir. Bu çalışmada kullanılan veri seti, her sınıfta 6.000 resim olmak üzere 10 sınıfta

60.000 renkli görüntü içerir. Veri kümesi 50.000 eğitim görüntüsüne ve 10.000 test görüntüsüne bölünmüştür. Sınıflar birbirini dışlar ve aralarında hiçbir örtüşme yoktur. Bu çalışmada da Google Colabratory ortamında büyük verinin sınıflandırılması uygulaması yapılmıştır.

3.3. Google Colabratory Ortamında Büyük Veride İlişkisel Yapı Modellerinin Oluşturulması

Bu alt bölümde yüksek performans hesaplama sistemi kullanılarak Google Colabratory ortamında büyük veride ilişkisel yapı modellerinin oluşturulması incelenecektir (İnternet Kaynağı 10, 2019). Bu çalışmada ele alınan örnek uygulama TensorFlow kullanarak zaman serisi tahminine giriş niteliğindedir. Bu çalışmada evrişimli sinir ağları (CNN) ve tekrarlayan sinir ağları (RNN) dahil olmak üzere birkaç farklı model stili oluşturulmuştur. Bu çalışma iki ana bölümde ele alınmıştır. (1) Tek bir zaman adımı için tahmin: (i) Tek bir özellik, (ii) Tüm özellikler. (2) Birden çok adım için tahmin: (i) Tek seferlik: Tahminleri tek seferde yapılır. (ii) Oto regresif: Her seferinde bir tahmin yapılır ve çıktı modeli geri beslemek için kullanılır.

3.4. Google Colabratory Ortamında Büyük Verinin Görselleştirilmesi

Bu alt bölümde yüksek performans hesaplama sistemi kullanılarak Google Colabratory ortamında büyük verinin hem kümelenmesi hem sınıflandırılması hem de ilişkisel görselleştirilmesi incelenecektir (İnternet Kaynağı 11, 2016). Bu çalışmada ele alınan konu güzel uygulama örneğidir.

Derin öğrenme, büyük miktarda veriyi analiz etmek için güçlü bir araçtır (İnternet Kaynağı 12, 2015). Uzaktan algılamada derin öğrenme özellikle anormallik tespitinde kullanılmaktadır (İnternet Kaynağı 13, 2017). Denetimsiz öğrenim kullanılarak çok zamanlı uzaktan algılama görüntülerinde birden fazla değişiklik türü belirlenebilmektedir (İnternet Kaynağı 14, 2018). Bu araştırma içinde görüntü sınıflandırmasına (İnternet Kaynağı 15, 2016) ek olarak biyofiziksel değişken regresyon modelleme de yapılmıştır (İnternet Kaynağı 16, 2019). Burada yapılan çalışma Landsat 8'i, eğitim verileri olarak Ulusal Arazi Örtüsü Veri Kümesini ve basit bir evrişimli sinir ağını kullanarak arazi örtüsü sınıflandırmasının temel bir örneğidir. Bu çalışmada Google Colabratory ortamında evrişimli sinir ağları kullanılarak arazi örtüsü haritalanmış ve sınıflandırması yapılmıştır. Çalışılan bölgenin arazi örtüsü büyük verinin kümelenmesi çerçevesinde farklı kümelere etiketlenmiş ve daha sonra evrişimli sinir ağları kullanılarak sınıflandırılmıştır. Sınıflandırma işlemiyle birlikte büyük verideki değişkenler için yapısal modeller oluşturulmuş ve büyük verinin görselleştirilmesi yapılmıştır.

4. SONUÇLAR VE ÖNERİLER

Akan veya katmanlı büyük veri analitiği söz konusu olduğunda büyük verinin hacmi ve karmaşıklığı dikkate alınmalıdır. Büyük veri analitiğinde kullanılan hesaplama sistemleri de önemlidir. Büyük veri olması durumunda analitiği Anaconda ortamında standart tek işlemci- tek çekirdek mimarisi kullanılarak Python kodlama ile yapılabilir. Büyük verinin hacminin büyük olması durumunda analitiği Anaconda ortamında tek işlemci- çok çekirdek mimarisi ile paralel hesaplama mimarisi kullanılarak Python kodlama ile yapılabilir. Büyük verinin hacminin büyük ve karmaşıklığının olması durumunda analitiği Google Colabratory ortamında çok işlemci- çok çekirdek mimarisi ile dağıtık hesaplama mimarisi kullanılarak Python kodlama ile yapılabilir.

Yüksek performans hesaplama sistemlerinde büyük veri analitiğinde donanım hesaplama ya da donanım hızlandırılmış hesaplama yapılabilir. Alternatif olarak büyük veri analitiğinde yazılım gerçekleştirilebilir.

Büyük verinin hacminin büyük olması durumunda analitiğinde makine öğrenme yöntemleri uygulanabilir. Büyük verinin hacminin büyük ve karmaşıklığının fazla olması durumunda derin öğrenme yöntemleri uygulanabilir.

KAYNAKÇA

Gökalp MO, Kayabay K, Akyol MA, Eren PE and Koçyiğit A. (2016). Big Data For Industry 4.0: A Conceptual Framework. 2016 International Conference on Computational Science and Computational Intelligence. 978-1-5090-5510-4/16 10.1109/CSCI.2016.87.

Hamza Erol ve Recep Erol (2018). Determining Big Data Complexity Using Hierarchical Structure of Groups and Clusters in Decision Tree. 2018 3rd International Conference on Computer Science and Engineering (UBMK).

Hamza EROL ve Timuçin KORKMAZ (2020). Büyük Veri Analitiği İçin Yüksek Performans Hesaplama Sistemi Gibi Davranan Bir Dağıtık Bilgisayar Sistemi Mimarisi. Bilgisayar Bilimleri ve Teknolojileri Dergisi - DergiPark. Yıl 2020, Cilt 1, Sayı 2, Sayfalar: 74-81.

Vassakis K, Petrakis E and Kopanakis I (2018). Big Data Analytics: Applications, Prospects and Challenges. Mobile Big Data, Lecture Notes on Data Engineering and Communications Technologies 10, Springer International Publishing AG, https://doi.org/10.1007/978-3-319-67925-9_1.

Thelin R (2020). What is Big Data? Characteristics, Types, and Technologies.

<https://www.educative.io/blog/what-is-big-data>.

İnternet Kaynağı 1 (2018). Yuval Noah HARARI 48. Dünya Ekonomik Formu. <https://www.weforum.org/events/world-economic-forum-annual-meeting-2018/sessions/will-the-future-be-human>.

İnternet Kaynağı 2 (2020). https://ec.europa.eu/esco/portal/escopedia/List_of_sectors_of_economic_activities_for_the_development_of_ESCO_v1.

Hamza Erol, Bala Mikat Tyoden, Recep Erol (2018). Classification Performances Of Data Mining Clustering Algorithms For Remotely Sensed Multispectral Image Data. 2018 Innovations in Intelligent Systems and Applications (INISTA).

Hamza Erol, Celaledin Barutçular, Ayman El Sabagh, Recep Erol (2017). Data Mining Models for Selection of the Best Spectral Reflectance Indices in Estimation of Crop Yields and Classification of Maize Hybrid Types Using SpectroRadiometer Data. 2017 European Conference on Electrical Engineering and Computer Science (EECS).

İnternet Kaynağı 3 (2019). <https://www.weforum.org/events/world-economic-forum-annual-meeting-2019>.

İnternet Kaynağı 4 (2018). <https://eab.com/insights/daily-briefing/workplace/the-top-10-emerging-jobs-for-2022/>

İnternet Kaynağı 5 (2020). <https://www.weforum.org/reports/the-future-of-jobs-report-2020>.

Timuçin Korkmaz, Hamza Erol (2020). Classification Of Human Facial Expressions For Emotion Recognition Using A Distributed Computer System. 2020 5th International Conference on Computer Science and Engineering (UBMK).

Osman Doğuş GÜLGÜN and Hamza EROL (2020) (a). Classification Performance Comparisons Of Deep Learning Models In Pneumonia Diagnosis Using Chest X-Ray Images. Yayın Bilgisi: 2020, Turkish Journal of Engineering. DOI: 10.31127/tuje.652358.

Osman Doğuş GÜLGÜN and Hamza EROL (2020) (b). Medical image classification with hybrid convolutional neural network models. Bilgisayar Bilimleri ve Teknolojileri Dergisi - DergiPark. Yıl 2020, Cilt 1, Sayı 1, Sayfalar: 28-41.

Hamza EROL, Recep EROL (2020). Reliability And Chaotic Risk Modeling For Real Time Data Driven Smart Systems. 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT).

Hamza Erol and Recep Erol (2016). Logical circuit design using orientations of clusters in multivariate data for decision making predictions: A data mining and artificial intelligence algorithm approach. 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA).

İnternet Kaynağı 6 (2021). <https://colab.research.google.com/github/jak-evdp/PythonDataScienceHandbook/blob/master/notebooks/05.11-K-Means.ipynb#scrollTo=0XcQWHWPCZpc>.

İnternet Kaynağı 7 (2021). <https://colab.research.google.com/github/mdai/ml-lessons/blob/master/lesson1-xray-images-classification.ipynb#scrollTo=Hk7H4FqrGCTM>.

İnternet Kaynağı 8 (2009). Learning Multiple Layers of Features from Tiny Images, Alex Krizhevsky, <https://www.cs.toronto.edu/~kriz/cifar.html>.

İnternet Kaynağı 9 (2019). <https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/images/cnn.ipynb#scrollTo=DSPCom-KmApV>.

İnternet Kaynağı 10. https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/structured_data/time_series.ipynb#scrollTo=G U8C5qm_4vZb.

İnternet Kaynağı 11 (2016). http://patrickgray.me/open-geotutorial/chapter_6_neural_networks.html.

İnternet Kaynağı 12 (2015). <https://doi.org/10.1038/nature14539>.

İnternet Kaynağı 13 (2017). <https://doi.org/10.1109/MGRS.2017.2762307>

İnternet Kaynağı 14 (2018). <https://doi.org/10.1109/TGRS.2018.2872509>.

İnternet Kaynağı 15 (2016). <https://doi.org/10.1109/TGRS.2016.2612821>.

İnternet Kaynağı 16 (2019). <https://doi.org/10.3390/rs11070768>.