

Module 3 - Applications in Data Science

Bu notlar, IBM Data Science Professional Certificate sertifika programının **What is Data Science** eğitimi üzerinden alınan bilgiler temel alınarak hazırlanmıştır. Notların geliştirilmesinde internet kaynaklarından ve yapay zeka araçlarından da yararlanılmıştır.

1. Veri Biliminin Gerçek Dünyadaki Uygulamaları

1.1 Temel Amaç ve Kazanımlar

Veri biliminin temel amacı, farklı sektörlerde karşılaşılan sorunlara optimum çözümler üretilmesinin sağlanması olarak tanımlanmaktadır.

- **Ortak hedef** olarak, geniş veri kaynaklarından toplanan ve temizlenen veriler üzerinde gerçekleştirilen analizler yoluyla yenilikçi yöntemlerin geliştirilmesi öngörülmektedir.
- **Veri stratejisi** benimseyen kuruluşlarda maliyet ve zaman açısından önemli avantajların elde edilmesi hedeflenmektedir.
- Büyük hacimli verilerden elde edilen içgörülerle karmaşık problemlerin çözülmesi sağlanmakta, bu durum kurumların rekabet üstünlüğü elde etmesi için temel bir araç olarak değerlendirilmektedir.

1.2 Ulaşım Sektöründe Örnek Uygulamalar

Veri biliminin ulaşım sektöründeki uygulamaları, hem bireysel hem de toplu taşıma alanlarında kritik faydalar sağlamak amacıyla çözüm odaklı olarak kullanılmıştır.

- **Uber Örneği:** Gerçek zamanlı kullanıcı verileri kullanılarak şehirdeki sürücü ve yolcu yoğunluğunun izlenmesi, anlık fiyatlandırma (surge pricing) stratejisinin

uygulanması ve doğru zaman–mekan dengesinin sağlanması mümkün kılınmıştır.

- **Toronto Transportation Commission Örneği:** Sokak tramvaylarının çalışma düzenine ilişkin detaylı verilerin; müşteri şikâyetleri, sensör verileri (probe data) ve GPS bilgileri aracılığıyla analiz edilmesi sonucunda, yoğun saatlerde trafik sıkışıklığının azaltılmasına yönelik düzenlemeler uygulanmış ve şehir içi kayıp zaman önemli oranda düşürülmüştür.

1.3 Çevre Sorunlarında Örnek Uygulamalar

Veri biliminin çevre alanındaki uygulamaları, göllerdeki zararlı yosun (cyanobacteria) patlamalarının önceden tahmin edilmesine yönelik stratejiler geliştirmede kullanıldığı görülmektedir.

- Robotik tekneler, şamandıralar ve insansız hava araçları aracılığıyla fiziksel, kimyasal ve biyolojik veriler yüksek hacimde toplanarak;
- Makine öğrenmesi ve istatistiksel modelleme tekniklerinin entegrasyonu ile zararlı patlamaların “ne zaman ve nerede” gerçekleşeceği öngörülmekte;
- Böylece halk sağlığının, içme suyu kaynaklarının ve rekreasyonel faaliyetlerin korunmasına yönelik proaktif önlemler geliştirilmektedir.

2. Etkili Veri Stratejisi ve Proje Süreci

Başarılı veri bilimi projelerinde etkili bir veri stratejisinin benimsenmesi ve projelerin planlama aşamasından sonuçların değerlendirilmesine kadar tüm süreçlerin sistematik bir biçimde yürütülmesi büyük önem arz etmektedir.

2.1 Temel Adımlar

- **Problemin Doğru Tanımlanması:** Sorunun kapsamı, paydaş ihtiyaçları ve performans göstergeleri netleştirilmekte; kurum içi veya toplumsal problemlerin ölçülebilir hedeflerle belirlenmesi sağlanmaktadır.
- **Veri Toplama:** Çeşitli kaynaklardan (sensör, anket, mobil uygulama, CRM vb.) veri temin edilmekte; verinin kalitesi ve güvenilirliği için standartlar oluşturulmaktadır.

- **Araç ve Teknoloji Seçimi:** İhtiyaç duyulan analiz türlerine (zaman serisi, mekânsal, makine öğrenmesi vb.) göre büyük veri işleme teknolojileri (Hadoop, Spark) veya veri ambarı çözümleri (Hive) belirlenmekte, uygun algoritmalar seçilmektedir.
- **Model Geliştirme ve Uygulama:** Makine öğrenmesi ve diğer analiz teknikleri ile modelin eğitim, test ve doğrulama aşamaları titizlikle yürütülmekte; proje koşullarına göre özelleştirme sağlanmaktadır.
- **Sürekli İyileştirme:** Geri bildirim döngüsü oluşturularak model ve strateji düzenli aralıklarla güncellenmekte; performans metrikleri doğrultusunda iyileştirme alanları belirlenmektedir.

2.2 Sonuç ve Faydalar

Veri bilimi projelerinin etkili sonuçlar vermesi için;

- Doğru veri stratejisi, operasyonel maliyetlerde azalma, verimlilik artışı ve proaktif müdahale imkânları sağlanmakta;
- Kurumların rekabet avantajı elde etmesi, yenilikçi çözümler geliştirilmesi ve toplumsal faydanın maksimize edilmesi mümkün kılınmaktadır.

3. Sağlık ve Afet Yönetiminde Veri Biliminin Rolü

Veri biliminin sağlık ve afet yönetimi alanlarındaki uygulamalarının, hem bireysel hem de kurumsal düzeyde kritik faydalar sağladığı gözlemlenmektedir.

3.1 Sağlık Sektöründe Örnek Uygulamalar

- **Tahmine Dayalı Analitik ve Kişiselleştirilmiş Tedavi:** Genetik veriler, tıbbi geçmiş, çevresel faktörler gibi veriler kullanılarak en uygun testlerin ve tedavi önerilerinin belirlenmesi sağlanmaktadır.
- **Örnek Vaka – Yaşam Kurtarıcı Teşhis Testleri:** Belirli kanser türleriyle ilgili tanı testlerinin kullanımında, veri bilimi tabanlı araçların hekimlerin farkındalığını

artırması yoluyla daha fazla hastanın yaşamsal değeri olan testlere erişimi mümkün kılınmaktadır.

- **Elektronik Tıbbi Kayıtlar (EMR):** North Shore University Health System örneğinde, anonimleştirilmiş devasa hasta verileri kullanılarak tanımlayıcı analizlerden öngörü sağlayan analizlere geçiş yapılması, hizmet kalitesinin artırılmasına katkı sağlamaktadır.

3.2 Afet Hazırlığı Uygulamaları

- **Yüksek Kaliteli Veri Setleri ve Tahmin Modelleri:** Deprem, kasırga, sel gibi doğal afetlerin gerçekleşme zamanı ve yerini öngörmeye yönelik modeller geliştirilmektedir.
- **Sosyal Medya İçerik Analizi:** University of Warwick örneğinde, sosyal medya paylaşımlarından elde edilen veriler, meteoroloji istasyonlarından gelen bilgilerle birleştirilerek yerel hava tahminlerinin iyileştirilmesine katkı sağlamaktadır.
- **Akademik Yaklaşımlar:** University of Chicago Graham School gibi programlar, afet yönetiminde veri bilimi uygulamalarını derinlemesine ele alarak, dağıtık veri kaynaklarının entegrasyonu ve model doğruluğu konularında çalışmalar yürütmektedir.

4. Veri Biliminin İşletme ve Toplum Üzerindeki Etkisi

Veri biliminin işletmeler ve toplum üzerindeki etkisi, hem tüketici davranışlarının izlenmesi hem de kurumsal stratejilerin belirlenmesinde önemli rol oynamaktadır.

4.1 Tüketicilerin Günlük Veri Üretimi

- **Dijital İzler:** İnternet aramaları, alışverişler ve sosyal medya etkileşimleri sonucunda oluşan büyük veri havuzları, kullanıcı davranışlarının analizine olanak tanımaktadır.
- **Öneri Motorları ve Konum Bazlı Öneriler:** Amazon, Netflix, Spotify ve Google gibi şirketlerin; kullanıcı geçmişleri ve konum verileri kullanarak

kişiselleştirilmiş öneriler sunulması sağlanmakta; giyilebilir teknolojiler de kişisel sağlık analizlerinde kullanılmaktadır.

Öneri motorları (recommendation engines), kullanıcıların önceki etkileşimleri, davranış geçmişleri ve benzer kullanıcı profilleri gibi verileri analiz ederek kişiselleştirilmiş içerik veya ürün önerileri sunmayı amaçlayan sistemlerdir. Bu sistemlerin temelinde genellikle iki ana yaklaşım bulunur:

- **İşbirlikçi Filtreleme (Collaborative Filtering):** Benzer ilgi alanlarına veya beğenilere sahip kullanıcı gruplarının tercihlerini karşılaştırarak öneriler sunar. Örneğin, bir kullanıcının beğendiği ürünleri başka benzer kullanıcıların da beğenip beğenmediğine bakarak tahmin yapar. Netflix ve Amazon gibi platformlar sıklıkla bu yaklaşımı kullanır.
- **İçerik Tabanlı Filtreleme (Content-Based Filtering):** Kullanıcıya önerilecek içerik veya ürünün özelliklerini (örneğin tür, kategori, içerik etiketleri) kullanıcının önceki tercihleriyle eşleştirerek öneri sunar. Spotify ve müzik uygulamaları, şarkıların türleri, sanatçı özellikleri veya dinleme geçmişini dikkate alarak kullanıcıya yeni şarkılar önerebilir.

Bu iki yaklaşımın birlikte kullanıldığı **Hibrit Sistemler** de mevcuttur. Hibrit sistemler, hem işbirlikçi filtrelemenin hem de içerik tabanlı filtrelemenin avantajlarını bir araya getirerek daha doğru ve kapsamlı öneriler sunabilir.

4.2 İşletmelerde Rekabet Avantajı

- **Veri Temelli Karar Alma:** McKinsey ve UPS örnekleri ile; veriye dayalı rota optimizasyonu, maliyet ve zaman tasarrufu sağlanmakta; bu durum işletmelerin rekabet gücünü artırmaktadır.
- **İçerik Üretimi ve Eğlence Sektörü:** Netflix örneğinde, kullanıcı davranışlarının detaylı analizi ile "House of Cards" gibi içeriklerin başarısının öngörülmesi ve doğru yatırımların yapılması sağlanmaktadır.

4.3 Genel Çıkarımlar

Veri bilimi, işletmelerin hem operasyonel verimliliğini artırmakta hem de müşteri deneyimini kişiselleştirerek rekabet avantajı sağlamaktadır. Doğru veri stratejileri sayesinde;

- Müşteri memnuniyeti yükseltilmekte,
- Operasyonel maliyetler düşürülmekte,
- Yeni ürün ve hizmetler geliştirilmekte; böylece sürdürülebilir büyüme desteklenmektedir.

5. Nihai Teslimat ve Veri Odaklı Hikâye Anlatımı

Veri bilimi projelerinin sonuçlarının etkili biçimde aktarılabilmesi için nihai teslimat (final deliverable) ve veriye dayalı hikâye anlatımının önemi büyüktür.

5.1 Nihai Teslimatın Amacı ve Biçimleri

- **Ana Hedef:** Analitik sonuçların paydaşlara aktarılması ve bu iç görülerin politika veya strateji oluşturulmasında kullanılmasının sağlanması amaçlanmaktadır.
- **Farklı Biçimler:** Akademik ortamda uzun makaleler veya raporlar, iş dünyasında ise daha kısa, görsel ağırlıklı raporlar şeklinde sunumlar tercih edilmektedir.

5.2 Deloitte Örneği

Deloitte tarafından hazırlanan ABD Ekonomik Görünümü Raporu örneğinde;

- Raporun, 24 sayfalık kapsamlı doküman şeklinde hazırlanarak, güçlü bir anlatı ile desteklendiği,
- Zaman serisi grafikleri, tüketici harcamalarındaki değişimleri ve çeşitli sektör verilerini içeren görsellerle, veriye dayalı senaryoların sunulduğu; bu yaklaşımın, raporun ikna ediciliğini artırdığı gösterilmiştir.

5.3 Hikâye Anlatımının Değeri

- **Anlatı ve Bağlamın Önemi:** Veriye dayalı analizlerin, tablolar ve grafiklerin ötesinde, metinsel açıklamalar, edebi referanslar ve tematik unsurlar kullanılarak sunulması, okuyucunun duygusal ve entelektüel bağ kurmasını sağlamaktadır.
- **Planlama Yaklaşımı – “Geriye Doğru Çalışma”:** Analiz sürecine başlanmadan önce nihai mesajın belirlenmesi, gerekli tablo ve grafiklerin önceden planlanması ile raporun bütünlüğünün ve kalitesinin artırıldığı gözlemlenmektedir.

6. Sonuç ve Anahtar Çıkarımlar

Veri bilimi, modern iş dünyasında ve toplumsal alanlarda;

- **Doğru veri toplama, sistematik analiz, stratejik sunum** prensiplerine dayalı olarak,
- Rekabet avantajı yaratma, operasyonel verimliliği artırma ve toplumsal fayda sağlama işlevi üstlenmektedir.

Özetle:

- **Veri stratejilerinin** doğru belirlenmesi,
- **Analitik sonuçların** etkili bir anlatı çerçevesinde sunulması,
- **Sürekli geri besleme** ile projelerin dinamik olarak iyileştirilmesi; bu temel yaklaşımların benimsenmesiyle, kurumların ve toplumun yararına değerli iç görüler elde edilmektedir.

