

SAYISAL İKİZ VE VERİ BİLİMİ

Çağrı Aksu¹

1. GİRİŞ

Günümüzde koskocaman bir üretim bandına dönmüş olan dünyamızı anlamak ve bu üretim bandında yer alabilmek için yeni teknolojileri ve bu yeni teknolojilerin işimize kattığı değerleri içselleştirmek oldukça önemlidir diye düşünmekteyim.

Yakın gelecekte bu üretim bandının otonomlaşması hedefi ile verimliliği arttırmayı, kısıtlı kaynakları daha iyi yönetebilmeyi uman iş çevrelerinin üretim alanlarını otonomlaştırma ve sonrasında yapay zekâ tabanlı pazar yerleri aracılığıyla alışveriş döngülerini de otonomlaştırmak istemeleri, bu hedefler doğrultusunda ilerleyen üretim noktalarını diğerlerine göre daha çok çalışılabilir olarak işaretleme nedeniyle, üretim ekosistemini tümüyle değiştirecektir.

Bu yeni üretim düzeni kuşkusuz kendisine ayak uyduran üreticiler yaratmak konusunu ön plana çıkarmaktadır. Bu yeni ekosistemde yer alabilmek için bize düşen görev ise işimizi sayısallaştırma ve işimizin matematiğini bulmak olacaktır.

2. SAYISALLAŞMA

Sayısallaşma (Dijitalleşme); kurumsal deneyimin kayıt altına alınmasıdır.

Kurumsal deneyim ise bir üretim süreci içerisinde yer alan her bir adım ve bu adımların farklı ortam ya da hammadde şartlarına göre farklı işleyiş, işletiliş biçimlerinin toplamı ve hammaddenin bir ürün olarak (belki de başka bir üretim sürecinde kullanılacak bir yarı bitmiş ürün) ortaya çıkarılıp taşıma işlemine devredilinceye kadar geçen sürede olan her şeyin bilinmesi ve kayıt altına alınmasıdır.

Üretim alanlarını sayısallaşma bağlamında dört gruba ayırabiliriz:

- **İlkel sayısallaşma:** Kurumsal deneyimin çalışanlar tarafından tutulması.
- **Gelişmekte olan sayısallaşma:** Kurumsal deneyimin yazılı belge üzerinde tutulması, bu belgelerin zamanla parça parça sayısallaştırılması.
- **Gelişmiş sayısallaşma:** Kurumsal deneyimin elektronik aygıt ve gereçler ile bilgisayar ortamına aktarılması.
- **Tam sayısallaşma:** Tüm üretim sürecinin birbiri ile ilişkili olarak elektronik aygıt ve gereçler ile bir bütünlük veri biriktirme ortamında irdelenip değerlendirilmeye hazır durumda kopyalanması.

Yukarıda yapılan sınıflandırmalar ölçülebilir değildir ancak, her bir sınıf, kurumların kusursuz sayısallaşmaya

¹ INTERTECH Bilgi İşlem, Master product owner AI, MLOps, - cagri@datascienceearth.com

geçmeleri için gereken aşamaları gösterir. Bu aşamaların sonucunda tüm üretim tesisinin sayısal bir ikizinin oluşturulması hedeflenmektedir.

Burada bir örnek olması bağlamında içinde bulunduğu günlerin popüler yapay zekâ aracı Chat-GPT-4'e "Sayısal ikiz nedir ve aşamaları nelerdir?" sorusunu sorarak, gelen yanıt üzerinden yukarıda anlatılanların daha anlaşılır olması sağlanmaya çalışılacaktır.

- "Sayısal ikiz nedir ve aşamaları nelerdir?"
- "Sayısal ikiz, fiziksel dünyanın sayısal bir kopyasıdır. Bu kopya, gerçek dünya nesnelere tasarımını, özelliklerini ve davranışlarını ayrıntılı bir şekilde temsil eder. Sayısal ikizler, endüstriyel tasarımdan otomasyona kadar birçok alanda kullanılır.

Sayısal ikizlerin oluşturulması aşağıdaki aşamalardan oluşur:

- 1. Veri Toplama:** Sayısal ikiz oluşturulması için gerçek dünya nesnelere tasarım ve davranış hakkında ayrıntılı veriler toplanması gereklidir. Bu veriler, sensörler, kameralar ve diğer aygıtlar aracılığıyla toplanabilir.
- 2. Veri İrdelenmesi:** Toplanan veriler irdelenerek, nesnelere tasarım ve davranış hakkında daha fazla bilgi edinilir. Bu aşamada, veriler işlenir, sınıflandırılır ve yorumlanır.
- 3. Sayısal Modelleme:** Toplanan ve irdelenen veriler kullanılarak sayısal ikiz modeli oluşturulur. Bu model, gerçek dünya nesnelere tasarım ve davranışını ayrıntılı bir şekilde yansıtır.
- 4. Benzetim (Simülasyon):** Sayısal ikiz, gerçek dünya nesnelere davranışını benzetmek için kullanılabilir. Bu benzetimler, gerçek dünya nesnelere davranışını sınamak ve iyileştirmek için kullanılabilir.
- 5. Gerçek Zamanlı İzleme:** Sayısal ikiz, gerçek dünya nesnelere davranışını gerçek zamanlı olarak izlemek için kullanılabilir. Bu izleme, gerçek dünya nesnelere başarımlarını ölçmek ve iyileştirmek için kullanılabilir.

Sayısal ikizler, gerçek dünya nesnelere tasarım ve davranış hakkında daha ayrıntılı bilgi sağlayarak, ürün geliştirme sürecini hızlandırabilir ve maliyetleri azaltabilir. Ayrıca, sayısal ikizler, gerçek dünya nesnelere başarımlarını ölçmek ve iyileştirmek için de kullanılabilir." OpenAI-Chat-GPT-4

Yanıt üç aşağı beş yukarı, bu soru bana sorulsaydı vereceğim yanıt çok yakındır. Peki bu yanıt nasıl üretildi? Elbette ki son yıllarda dünyamızdaki sayısal ortamın artma-

sının sonucunda ortaya çıkan birbiri ile ilişkili inanılmaz büyük bir veri uzayının yapay bir zekâ modelinin eğitilmesi için kullanılması sonucunda elde edildi. Yani insanlığın deneyiminin sayısal olarak kayıt altına alınması ve bu sayısal verilerden yararlanılarak sorulan sorunun anlaşılır ve uygun yanıtın aranması sonucunda ortaya çıkmıştır.

Aynı senaryoyu bir üretim tesisi için de düşünebiliriz. Tam sayısal bir üretim tesisinin verileri ile üretilen bir yapay zekânın, üretim maliyetinin nasıl düşürüleceği ya da ortaya çıkan ürünün nasıl daha nitelikli olabileceğini söyleyebilmesi olasıdır.

Burada Sayısal İkiz kavramını biraz daha açmak isterim. Bir üretim tesisi için sayısal ikizden bahsettiğimizde, aslında tek bir sayısal ikizi anlayamayız. Çünkü, her bir üretim elemanının nicelik bağlamında bir sayısal ikizinin olması ve aşama aşama bunların birleşmesi ile daha karmaşık bir sayısal ikiz oluşturulur. Nitelik bağlamında kastettiğimiz ise, farklı görevleri yapabilmesi bağlamında sınıflandırılabilen bir yapıdır.

2.1 Sayısal İkiz Düzeyleri

Sayısal ikizler, kapsam ve işlevsellikleri bakımından da sınıflandırılmaktadırlar. Çok farklı sınıflandırmalar olmakla birlikte, anlaşılması açısından aşağıdaki gibi bir sınıflandırma paylaşılmaktadır.

- 1. Basit Sayısal İkizler:** Basit sayısal ikizler, bir nesnenin sadece temel özelliklerini ve davranışlarını yansıtır. Yani burada veri akışı tek yönlüdür. Toplanan veri görselleştirilebilir, raporlamalar için kullanılabilir. Fakat bu düzeyde sayısallaştırılma yapıldığında ortama karışma söz konusu değildir.
- 2. Orta Düzey Sayısal İkizler:** Orta düzey sayısal ikizlerde veri akışı iki yönlü olmaktadır. Gerçek ortamın verileri alınır, görselleştirilir, raporlanır. Ortaya çıkan bilgiler ışığında bir işletmen, sayısal ikiz üzerinden gerçek üretim ortamını değiştirebilir. Örneğin üretim bandının hızını değiştirebilir.
- 3. Zeki Sayısal İkizler:** Zeki sayısal ikizler, bir işletmen yardımına gerek duymadan üretim bandına üreteceği bilgiler doğrultusunda karışabilir. Örneğin ortam sıcaklığının ürün niteliğini etkileyecek bir düzeye gelmesi sonrasında bir iklimlendirme aygıtını çalıştırıp üretim alanını en uygun sıcaklık değerine getirebilir.
- 4. Çoklu Zeki Sayısal İkizler:** Bu düzeyde artık sayısal ikiz bütün üretim elemanlarını ve üretim sürecini eş zamanlı bir şekilde yönetebilmektedir. Bu düzey, artık bir "Karanlık Fabrika" olarak tanımlanabilir.

Yukarıdaki sınıflandırma elbette ki farklı konu ve içerikteki üretim alanlarında farklı şekilleriyle uygulanabilmektedir. Burada aklımıza gelen önemli bir soru "Bir sayısal ikiz nasıl zeki olabilir?" olacaktır.

3. VERİ BİLİMİ

Veri Bilimini Kullanmak: İşinizin matematiğini keşfedin.

Veri bilimi; veri biriktirme-işleme, istatistik, bilgisayar bilimleri, sosyoloji, işletme gibi çeşitli alanların birlikte çalışması ile birlikte bilgi üretme sürecinin tamamını kapsayan bir çatı kavramdır.

Veri bilim bileşenleri, öncelikle tarih içerisinde bilimsel araştırmalarda kullanılmak için geliştirilen yöntemler ile ortaya çıkmıştır.

Bilgisayar teknolojisinin gelişimi ile veri biriktirme gereksinimi artmış ve bu biriken verinin ticari ve bilimsel alanlarda kullanımı gündeme gelmiştir.

2000'li yıllara kadar farklı disiplinler olarak devam eden bileşenler, Büyük Veri ile birlikte Veri Bilim çatısı altında toplanmıştır. 'Veri Bilimi' deyimini ilk kez, istatistikçi William Cleveland tarafından 2001 yılında kullanılmıştır.

3.1 Veri Bilimini Kullanmak: Bilgi Fabrikaları

Bilgi fabrikası, temel üretim alanı ile birlikte bu üretim sürecinin matematiğinin bulunması için veri ve bilgi de üretmeyi hedefleyen tesislerdir. Bu süreç içerisinde şirketler Veri Bilimi ile bütünleşmeyi hedeflerler.

Veri Bilimi ile bütünleşen şirketler, kendilerini geliştiren Veri Bilimi süreçlerinin giderek artan başarımı ile kendilerini tam olarak anlatabilen matematiğin giderek daha kesin sonuçlar verdiğine tanık olurlar. Bu sürecin sonu kuşkusuz ki, yapay zekanın şirketin geleceğinde oynayacağı rolün habercisidir.

Bilgi fabrikası olmak yolunda şirketlerin izlemesi gereken yol ve aşamalar aşağıdaki gibi özetlenebilir:

- 1. İlkel şirket:** Ürettiği verinin farkında olmayan, veriyi işleyemeyen şirket.
- 2. Gelişime açık şirket:** Verinin gücünü fark etmiş ve veriyi irdelemeye başlamış şirket.
- 3. Veri odaklı şirket:** Yaptığı irdeleme çalışmalarının olumlu sonuçlarını almış ve veri kalitesini artırma yolunda adımlar atmış şirket.
- 4. Veri Bilim ile bütünleşmiş şirket:** Yaptığı işi matematiksel formüllerle anlatabilen şirket.

5. Yapay zekâlı şirket: Yapay zekânın etkin olarak yönetim kademesine katıldığı şirket.

Herhangi bir şirketin Veri Bilimi ile olan macerasına başlaması için kendilerini öncelikle veri odaklı şirket konumuna çekebilmesi gerekmektedir.

Veri odaklı şirket olabilmek ise yine veriyi iyi işleyip irdelemeye bağlıdır. Çünkü veri irdelemesi aslında, geri dönüp iş süreçlerini en uygun duruma getirmeye yarayan sihirli bir sopadır. Verisini iyi işleyip kullanabilen şirketler, bu aşamada kazandıkları deneyimlerden yola çıkarak iş süreçlerini en uygun duruma getirip, elde ettikleri bilgilerle de geleceklerini şekillendirirler.

İş süreçlerinin en uygun duruma getirilmesinin en önemli sonucu, şirketlerin ürettikleri ürünün ve elde ettikleri verinin niteliğinin artmasıdır. Verinin daha iyi ve doğru olması kuşkusuz ki, başarıyı giderek artıran bir döngüdür.

Şirketlerin bu döngüye girebilmesi için ise öncelikle kendilerini veri bakımından gelişime açık şirket konumuna çekebilmeleri gerekir.

Veri bakımından gelişime açık şirket olmak, veri irdelenip işlenmesinin hem bir amaç hem de araç olduğunun fark edilmesi ile başlayan bir süreçtir. Bu gerçeğin farkına varan şirket, öncelikle veri biriktirme teknolojilerine yaptıkları yatırımı artırır. Yapılan yeni yatırımlar şirket çalışanlarının yaratılan yeni süreçlere uyum sağlamasına yönelik eğitimlerle gelişir ve şirket yarattığı veriyi nasıl kullanabileceği konusunda çalışmalara başlar. Şirketlerin geçirdikleri bu süreçler, hem kârlılıklarını hem de iş yaptıkları pazarları geliştirir. Bu gelişim süreci doğal olarak daha fazla, daha çeşitli veri üretilmesine yol açar. Yeni veri üretim süreçleri ise yeni veri analiz süreçleri gerekliliğini ortaya koyar.

3.2 Veri Bilimi Takımı

Veriden bilgiye giden süreç temel olarak, şirketin gereksinimleri doğrultusunda başlamakta ve bu gereksinimlere göre şekillenmektedir. Ancak şirket gereksinimlerinin belirlenmesi, bunların karşılanması için gerekli verinin saptanarak kayıt altına alınması, kayıt altına alınan verinin tanımlanması, sonrasında bu sınıflandırılmış verinin irdelenip işlenerek bilgiye dönüştürülmesi ve ortaya çıkan bu bilgilerin kullanıma uygun şekilde düzenlenmesi bir takım işidir.

Takım üyeleri yukarıda anlatılan süreçlerin içerisinde aldıkları roller bakımından belli yeterliliklere sahip olmalıdırlar. Bu yeterlilikler öncelikle yapılan işin işleyişini, kurlarlarını bilmektir.

Bu anlamda, Veri Bilim ekibi, şu soruların yanıtlanması için çalışma yapar:

- Veri nedir?
- Nasıl tanımlanır?
- Nasıl biriktirilir?
- Hangi araçlar ile irdelenip işlenir ?
- Hangi yöntemler kullanılır?
- Elde edilen sonuçlar nasıl sınılanır?
- Bilgi nasıl kullanılır?

3.2.1 Veri Bilimi Takımı Veriyi Anlamalı

Veriyi anlamak için öncelikle veri tiplerini anlamak gerekir çünkü veri irdeme süreci veri tipleri ile başlar. Hangi algoritmayı hangi amaçla kullanabileceğine aslında veri karar verir.

Veri bilimi içerisinde kullanılan veri kaynakları Düzenli (Structured) ve Düzenlenmemiş (Unstructured) olarak ikiye ayrılır.

Veri tabanları, veri ambarları, ERP, CRM gibi verinin düzenli tutulduğu veri kaynakları, Sosyal Medya verisi, Resim, Video, Belge, Sensör gibi düzenlenmemiş veri kaynakları ile birleştirilerek kullanılır.

3.2.2 Veri Bilimi Takımı Veriyi Biriktirmeli

Kesinlikle veri biriktirme teknolojilerini çok iyi bilmelisiniz. Unutmayın ki, veri ile çalışabilmek için önce veriye erişebilmelisiniz.

İlişkisel veri tabanlarını (SQL), NoSql ve doğaldır ki büyük veri araçlarını kullanabilmelisiniz.

3.2.3 Veri Bilimi Takımı Veriyi İşleyebilmeli

Veri tabanlarında, diğer veri kaynaklarınızda inanılmaz çok sayıda yapılandırılmış ya da yapılandırılmamış veriler olacaktır. Fakat siz, irdeme yapabilmek için gerek duydüğünüz veriyi hazır olarak bulamayacaksınız.

İrdeleyeceğiniz veri setini, veriyi işleyerek siz oluşturmalısınız. Ve bu veri işleme süreci gerçekten çok zorlu geçecektir. Bunu başarabilmek için, kesinlikle bu konuda eğitim almış ve çok fazla uygulama yapmış olmak gerekir.

Veri işleme için hazır birçok araç bulunduğundan, bunları kullanmayı öğrenmelisiniz.

3.2.4 Veri Bilimi Takımı Veriyi İrdeleyebilmeli

İrdeleme aşamasına geldiyseniz, bunu yapabilmek için şunları yapmalısınız:

Öncelikle sağlam bir istatistiksel alt yapı oluşturmalısınız. İstatistik, Veri Bilimi kavramının çok önemli bir parçası ve olmaz ise olmaz unsurudur.

Makina öğrenmesi tekniklerini de öğrenmelisiniz. Hangi algoritma hangi işe yarar, hangi algoritma hangi veri ile kullanılır? sorularının yanıtlarını, deneyerek ve araştırarak bulmalısınız. Zamanla algoritmaların teorik altyapılarını da bilmek zorunda olduğunuzu fark edersiniz.

3.2.5 Veri Bilimi Takımı Veriyi Ürüne Çevirebilmeli

Örneğin; Google arama çubuğuna yanlış bir şey yazdığınızda hemen size "Şunu mu demek istedin?" diye sorar, işte o bir veri bilimi ürünüdür.

Bu konuda kendi kendinizi geliştirmeniz çok kolay değildir. Fakat ürün denemeleri yaparak bunları izleyerek, yararlı deneyimler elde edebilirsiniz.

3.2.6 Veri Bilimi Takımı Gelişime Açık Olmalı ve Farklı Düşünebilmeli

Her sabah yeni bir yöntemin uygulandığını görebildiğimiz günümüz iş dünyası içerisinde iyi bir Veri Bilimi takımı, güncel süreçleri eksiksiz bir şekilde izleyip yönetebilmeli ve gelişime açık olmalıdır.

Gelişime açık olmak, kurum içi gelişime yönelik yeni aksiyonlar alma ve kurum dışındaki yenilikleri kurum içerisine uyarlayabilme becerisidir.

Veri Bilimi takımı, farklı becerileri olan farklı deneyim düzeylerinden ve farklı düşünebilen insanlardan oluşturulur.

Buradaki en önemli olgu, farklı düşünebilme yetisidir. Farklı düşünebilmek; olaylara, nesnelere ve bunlar arasındaki ilişkilere farklı açılardan bakabilmektir. Bu farklı bakma yeteneği, zihinde uzun mantık zincirleri kurarak oluşturulabilir.

Veri bilimi, elbette ki bir takım oyunudur. Bu oyunun içinde yer alabilmek için çalışanların oyun içindeki rollerden birini seçmeleri gerekmektedir. Ve doğal olarak, rol seçimleri yapılırken, çalışanların yetenekleri göz önüne alınmalıdır.

4. SONUÇ

Sayısal İkizler ve Veri Bilimi, ilk bakışta çok karmaşık bir yol ya da bazen sihirli bir çözüm gibi görünebilir. Fakat bu ikisi de doğru değildir. Bunlar doğru bir yol haritası ile zaman içerisinde ilerlenecek birer yoldur. Bu yollara çıkmadan önce iyi bir planlama yapmak, plana sadık kalmak çok önemlidir. Çünkü bu yollar, zorlu parkurlar ve bazı tuzaklar içerir. Bu süreçten kazançlı çıkmak sabır ve yatırım gerektirir. ⚡