

# Geleceğin Doktoru Şimdilerde Komada

## IBM Dr.Watson Projesi Neden Başarısız Oldu? Sağlıkta Yapay Zekanın Zorlukları Neler?

Oğuz Dicle

**B**ilişim devlerinden IBM'e ait Dr.Watson projesi büyük vaatlerle çıktığı yolculuğunu bu yıl başında sessiz sedasız bir şekilde sonlandırdı. Sağlık alanının en önemli yapay zeka projesi olarak da bilinen bu proje niye başarısız oldu? Yapay zekanın sağlık alanındaki gidişatı bu kararla nasıl etkilenebilir?

IBM Watson projesi 2011 yılında duyurulduğunda sağlık alanında devrim yaratacak bir proje müjdesi verilmişti. Proje ile bilgisayar yazılımı Dr.Watson, hastanın kendisine verdiği yakınmaları ve laboratuvar verilerini değerlendirerek bir hekim gibi tanıyı koymayı ve tedavi planlaması yapmayı vaat etmişti. Bu tıp camiasında büyük bir heyecan yaratmış, mesleğin sonuna mı gelindi sorularına neden olmuştu. Ne var ki gelinen bu aşamada bu korku-

nun bir süre ertelendiği anlaşılmakta. Peki ama Dr.Watson'a bu geri adımı attıran nedenler nelerdi?

Watson başlangıçta bir soru yanıt sistemi olarak düşünüldü. Akla gelebilecek her türlü yazılı metni işleyerek adeta yenilmez bir bilgi yarışmacısı olması hedeflendi. Bu hedeflerinde önemli başarılar da elde edildi. Kampanyalarında Serena Williams (<https://vimeo.com/151899150>) ve Bob Dylan gibi ünlüleri kullandılar. Elde edilen başarılar firmayı sağlık alanında benzer bir başarı için motive etti ve proje sağlık alanına yönlendirildi. Projenin temel hedeflerinin verdikleri kararları doğal dil işleme yöntemleri ile çözümlenerek yapay zeka ile öğrenen ve benzer durumlarda önceki başarılı kararları yeni hastalara önermeye dayanıyor idi. Watson, hasta-

ların sađlık kayıtları yanı sıra tüm tıp literatürünü okuyabiliyordu: ders kitapları, hakemli dergi makaleleri, onaylanmış ilaç listeleri vb. Tüm bu verilere erişimle Watson, hiçbir insanın asla fark edemeyeceđi kalıpları ayırt eden bir süperdoktor olabilirdi. Umut, güçlü bilgi işlem gücüyle Watson'ın bu kayıtlardaki demografi, tümör özellikleri, tedaviler ve sonuçlar dahil olmak üzere yüzlerce deđişkeni incelemesi ve keşifler yapması ve hatasız öneriler vermesiydi. Ayrıca, her gün yayınlanan kanser tedavileri hakkında çok sayıda dergi makalesi ile güncel kalacaktı. Projede önemli bir sađlık kuruluşu olan MD Anderson ile işbirliđi yapıldı. Firmanın hedefi en geç iki yıl içinde süper doktoru gerçekleştirmektir.

Ancak, ünlü deyimdeki gibi evdeki hesap çarşıya uymadı. İş sađlık alanına gelince, mühendislerin öngöremediđi çok sayıda sorunla karşılaşıldı. Öncelikle sađlık verileri yapay zeka çalışmalarında tercih edilen ve yapılandırılmış veri denilen, verinin hangi bağlamda, hangi nedenle, ne zaman, ne için, kim tarafından toplandıđı gibi bilgileri hatasız olarak sunan verilerden oluşmuyordu. Verilerin %80'e yakın oranda yapılandırılmamış veriler olduđu görüldü. Serbest anlatıma dayalı, dilbilgisi kurallarına uymayan, kısaltmalarla dolu, çoklu anlamlar çıkarılabilen ifadeler içeren metinsel kayıtlar baş edilmesi olanaksız bir karmaşaya sahipti. Bir anlamda yapay zeka doktorların dađınlık okunması güç ve fıkralara konu olan yazılarına yenik düştü. Dil işlemeyi zorlaştıran bu teknik sorunlar dışında, kayıtlardaki tıbbi notların yetersizliđi ve eksik verilerden kaynaklı beklenmedik sorunlar sorunu büyüttü. Bu yetmezmiş gibi, işbirliđi yapılan ve sađlık kayıtlarının kul-

lanıldıđı MD Anderson, proje yürürken hastane kayıt sisteminde yaptıđı bir deđişiklik ile Watson yazılımının büyük yara almasına yol açtı. Kullanılan klinik bilgilerin eski olgulardan elde edilmesi de doğruluk oranlarında düşmelere neden oldu.

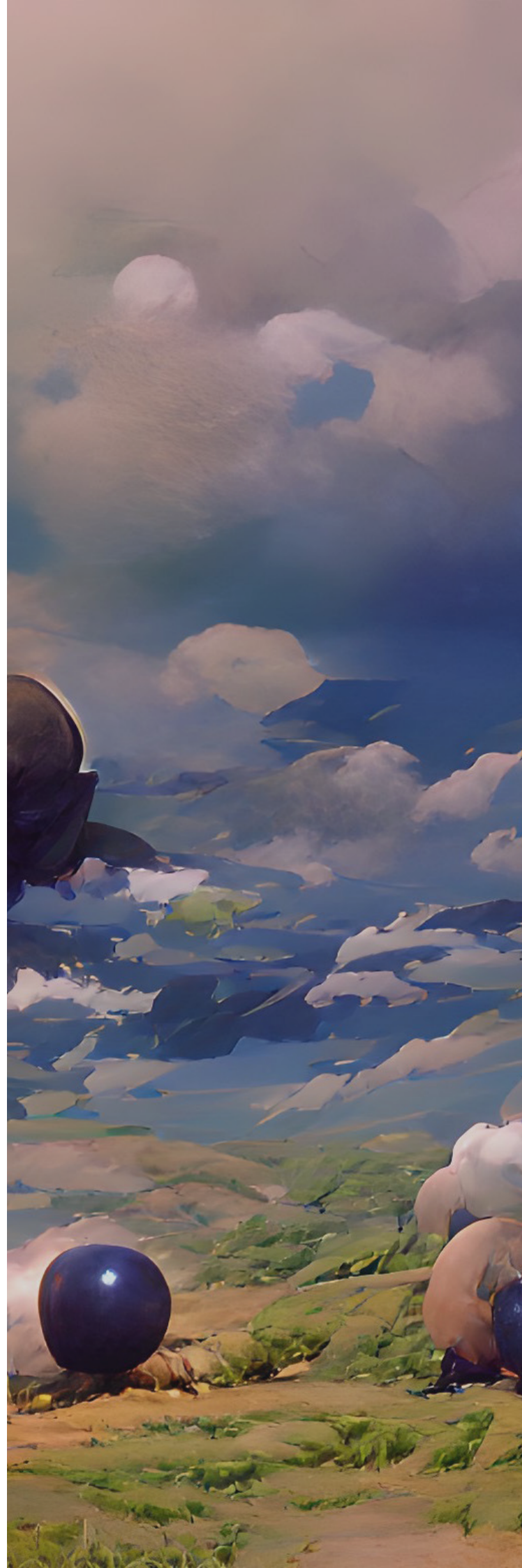
Dr.Watson projesinin başarısız şekilde sonlanması sadece unsurlara bağlamak elbette yetersiz olur. Konuyla ilgili yapılmış deđerlendirmelere bakıldıđında şirketin stratejik hatalarının büyük rol oynamış olduđu görülüyor. Uzmanların dile getirdiđi nedenlerin başında IBM'in gerçekleştiremeyecek vaatlerde bulunarak hayal kırıklıkları yaratması, sorunları tek tek çözmek yerine bir seferde tümünden çözme arayışı, kaliteli veri toplamaya yeterli zaman ayırmaması, önceliđi pazarlamaya vermesi, Google ve Apple gibi devlerle rekabette olması nedeniyle hızlı hareket etmek zorunda kalması ve hedefine ulaşmaması geliyor. Medyada sıkça çıkan güven kırıcı haberlerin de bu süreçte etkili olduđu belirtiliyor.

Dr.Watson projesi, sađlıkta yapay zeka çalışmalarında karşılaşılan en önemli engellerinden biri olan klinisyen ve mühendisler arasındaki sorunlu ilişkilere ait derslerle de doludur. Klinisyenler yoğun iş gücü altında yıllar içinde geliştirdikleri tempolu hizmet pratiklerini etkileyecek en küçük bir aksaklıđa karşı tahammülsüzdürler. İşlem basamaklarındaki fazladan bir tıklama bile direnç nedeni olabilmektedir. Watson'un klinik iş yükünü azaltacađı yönündeki vaadine rağmen bunu sağlayamaması klinisyen tarafındaki desteđi zaman içinde azaltmıştır. Uygulamanın günlük pratiđe ek yük getirmemesi gerekiyordu. Program bunu çok

dikkate almadığından klinikte kabul görmesi istenilen hızda olmadı. Hekimler yazılımı kullanmaya çalışırken hastalarla ilgilenmeye zamanları kalmadı.

Yenilikçi yazılımların tümü için geçerli olan güven oluşturma konusunda da IBM Watson proje yöneticilerinin hatalar yaptıkları söylenebilir. Yapay zeka yazılımlarının temel sorunlarından biri olan “kara kutu sorunu” güven oluşturmaya engelledi ve ürünün yaygın klinik uygulamasına engel oluşturdu. Kara kutu, yapay zeka algoritmalarının sonucu nasıl ulaştıklarının bilinmemesi durumuna verilen isimdir. Örüntü tanıma üzerine kurulu algoritmalar bu örüntüyü çıkarabilmek için binlerce özellikten yararlanmaktadır. Bu özelliklerin bize göre anlamlı olmaları da gerekmemektedir. Ancak bu durum kanıta dayalı, neden -sonuç ilişkileri içinde karar veren hekimler için çoğu zaman kabul edilebilir değildir. Bu gerçeğe rağmen şirket, kanaat önderlerini ikna etme gibi güven sağlayıcı olabilecek yollara da baş vurmadı. Hatta klinisyenleri işin içine sokmadan sadece bize inanın demekle yetinildiği eleştirileri bile bulunuyor. Watson belki de bu nedenle üretildiği ABD’den çok bunu bir pazarlama unsuru olarak kullanan Uzak Asya ülkelerinde tercih edilen bir ürün olmuştur.

Yaşanan gelişmeler IBM’in, makinelerin öğrenme şekli ile doktorların çalışma şekli arasındaki temel farkları göz ardı ettiğini ortaya koymaktadır. Makineler tanımlanan algoritmaları ve bu algoritmalara uyum gösteren yeni verileri kullanırken belirsizlik durumlarına karşı oldukça hazırlıksızdır. Buna karşın tıbbın en başat özelliklerinden bir olan belirsizlik ve hızla değişen durumlarda hekimlerin hızlı karar verebilme becerileri onlara eşsiz bir



üstünlük vermektedir. Watson'ın düşüncesi ise istatistiklere dayanıyor ve verebileceği karar istatistiksel bir yalnızlığa sahip. IBM sağıktaki belirsizlik sorununu önemsemedi ve hekimlerin önceki kararlarını algoritmaları işletmede yeterli buldu. Yapay zeka ile çalışan makineler öğrenme sürecinde kullanılan eğitim setlerine bağımlı durumdadır. Bir başka şekilde ifade edecek olursak makinanın doğruluğu, belirli bir evrende gözlenen ve o sınırlar içerisinde doğru olarak işaretlenmiş verilerle sınırlıdır. Evren değiştiğinde algoritma çalışmayabilir. Eğitim setleri ile test edilen verinin aynı kaynaktan gelmesi Dr.Watson projesinin bir başka kısıtı olmuştur. Bunun bir sonucu olarak farklı merkezlerden yapılan yayınlarda hekimlerin kararları ile Watson'nun kararları arasındaki uyum %43 ile 83 arasında kalmaktadır.

Şimdilik rafa kaldırılmış olan IBM Watson projesine ne kadarlık bir kaynak ayrıldığı bilinmemektedir. Texas Üniversitesi tarafından 2016 yılında yapılan bir denetim, kanser merkezinin projeyi iptal etmeden önce projeye 62 milyon dolar harcadığını belirlemiştir. Bu büyüklüklerde harcamalar yapılmış ve sağıktaki yapay zeka çalışmaları açısından büyük bir hayal kırıklığı olan projenin kazandığı birkaç başarıya da değinmek yerinde olacaktır. Yazılım ekibinin genetik alanda geliştirdiği Watson for Genomics projesi önemli bir başarı öyküsüdür. Belirsizlik içermeyen veri setlerinin kullanıldığı alanda yapay zeka algoritmaları öngörülerinde yüksek başarılar elde etmiş, günlük pratiği hızlandıran yararlı araçlar olarak kabul görmüşlerdir.

Dr.Watson'un öyküsü bize gelecekteki yapay zeka çalışmaları için önemli dersler içeriyor. Kilidi aç-

cağ anahtarın klinisyen ve mühendislerin birbirlerini anlamaya daha fazla zaman ayırmaları olduğunu düşünüyorum. Hekimler nasıl düşünür ve karar verir sorusu daha iyi araştırılması gereken bir soru. Belirsizlik yönetimi belki de geleceğin bilimi olacak.

#### Kaynaklar

- <https://www.nytimes.com/2021/07/16/technology/what-happened-ibm-watson.html>
- <https://medcitynews.com/2021/04/why-ibm-watson-health-could-never-live-up-to-the-promises/>
- <https://blog.sepstream.com/reasons-why-ibm-watson-flopped-in-the-healthcare-industry/>
- <https://spectrum.ieee.org/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>
- <https://www.quora.com/Why-did-IBM-Watson-fail>
- The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age (McGraw-Hill) 2015, 1st ed.
- Zhao X et al. Concordance between treatment recommendations provided by IBM Watson for Oncology and a multidisciplinary tumor board for breast cancer in China Jpn J Clin Oncol. 2020 Aug 4;50(8):852-858.