

Derleme

Nöroşirürji'de Yapay Zeka ve İnsan Zekası

Artificial Intelligence and Human Intelligence in Neurosurgery

Ömer Selçuk ŞAHİN¹, Melih ÜÇER², Burak ÖZAYDIN³, İhsan DOĞAN⁴

¹Sağlık Bilimleri Üniversitesi Dışkapı Yıldırım Beyazıt Eğitim ve Araştırma Hastanesi, Beyin ve Sinir Cerrahisi Kliniği, Ankara, Türkiye

²Kanuni Sultan Süleyman Eğitim ve Araştırma Hastanesi, Beyin ve Sinir Cerrahisi Kliniği, İstanbul, Türkiye

³University of Wisconsin Medical School and Public Health, Department of Neurosurgery, Madison, A.B.D.

⁴Ankara Üniversitesi Tıp Fakültesi, Beyin ve Sinir Cerrahisi Anabilim Dalı, Ankara, Türkiye

ÖZ

Yapay zeka ve bunun bir dalı olarak görülen makine öğrenimi son yıllarda giderek popülaritesini artırmış ve bu alanda yapılan çalışmalar bütün tıp literatürünü olduğu gibi nöroşirürjiyi de etkiler hale gelmiştir. Derlemede nöroşirürji alanındaki yapılan çalışmalar ve özellikle insan ve bu akıllı yazılımların performanslarının karşılaştırıldığı çalışmalar incelenmiş ve nöroşirürjinin teknolojik geleceğine ışık tutulmak istenmiştir. Yapay zekanın nöroşirürji alanında gelecekte sağlayabileceği avantajlar, ve her geçen gün daha da ilgi çeken bu alandaki yenilikler okuyucunun huzurunda tartışılmıştır. Okuyucunun bu alandaki muhtemel endişeleri göz önünde bulundurularak mevcut literatür incelenmiştir. Aynı zamanda makine öğrenmesi hakkında da bilgi verilmiş olup, bu konudaki uygulamalar çeşitli yönleriyle karşılaştırılmıştır.

ANAHTAR SÖZCÜKLER: Nöroşirürji, Yapay zeka, Makine öğrenimi

ABSTRACT

Artificial intelligence and its branch of machine learning have become increasingly popular in recent years with many new publications in neurosurgery as well as in the general medical literature. In this review, studies in the field of neurosurgery and especially those comparing human performance and the performance of intelligent software have been reviewed to try to elucidate the technological future of neurosurgery. Artificial intelligence can provide several advantages in the field of neurosurgery, and the innovations in this area, which are getting more and more interesting each day, have been discussed. The current literature has been reviewed in view of the reader's probable concerns in this area. In addition, information about machine learning has been provided, and the applications in this area have been compared from various aspects.

KEYWORDS: Neurosurgery, Artificial intelligence, Machine learning

■ GİRİŞ

Son yıllarda tıp alanında artan verilerin büyüklüğü ve karmaşıklığı insan kabiliyetinin üzerine çıkmıştır. Hastanelerde her gün devamlı suretle yeni veriler üretilmektedir. Birçok istatistiksel test doğrusal olmayan değişkenlerde ve çok büyük hacimdeki verileri sınıflandırmada yetersiz kalmaktadır. Nöroşirürjide günümüzde uygulanan karar alma mekanizmaları ise klinik kılavuzlar, eski olguların analizleri ve istatistiklerine dayanmaktadır.

Yapay zeka olarak bilinen akıllı bilgisayar algoritmaları, programlanmaya ihtiyaç duymadan büyük veriler üzerinde çalışıp anlamlı sonuçlar elde edilmesini olanaklı hale getirmeye başlamıştır (16,17). Bu uygulamalar birbirlerinden farklı matematiksel algoritmalar ile farklı çözümler sunabilmektedir ve istatistiksel analizlerin çözemeyeceği problemlerin üstesinden gelerek bu gibi durumlarda bizlere yardımcı olmaktadır. Klinisyenlerin bu verileri özelliklerine göre kategori etmesine gerek kalmadan yalın haliyle bu veriler yapay zeka tarafından



Yazışma adresi: İhsan DOĞAN

E-posta: Ihsan.Dogan@ankara.edu.tr

işlendiğinde anlamlı birer çıktıya dönüşebilmektedir. Bu durum özellikle patoloji kesitleri, radyoloji görüntüleri, serbest metinler ve monitörize edilmiş grafik verileri sınıflamada oldukça yararlı olabilmektedir. Ayrıca yapay zeka uygulamaları günümüz istatistiksel yöntemlerden farklı olarak insan faktörünü ekarte edebildikleri için “bias” yani yan tutma riskini de minimuma indirebilirler (7), ve bu sayede klinik çalışmalarda daha etkili ve doğru sonuçlar elde etmek mümkün olabilir.

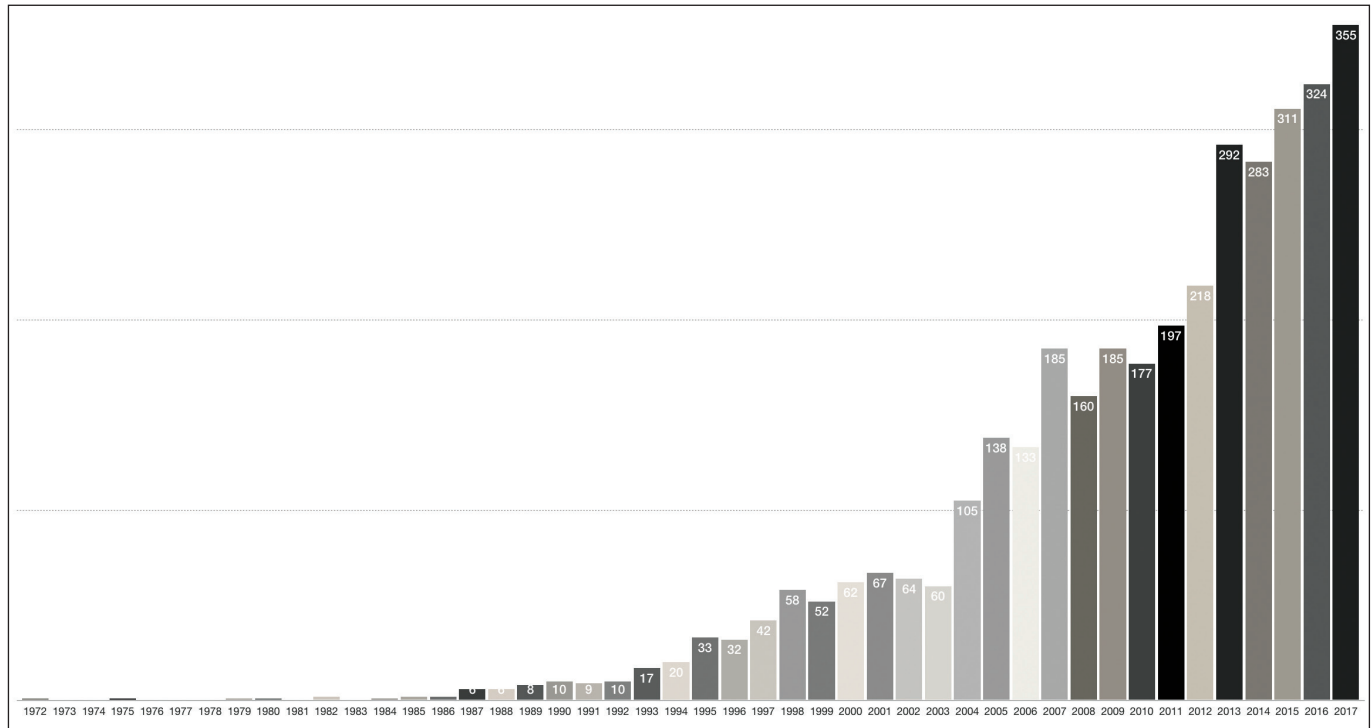
Bu yazıda ismi sıklıkla geçecek olan “makine öğrenmesi” terimi yapay zekanın bir türevidir olup genel manada bilgisayar algoritmalarının süreç aşamasında programlanmaya gerek duymadan verileri kullanarak istenen çıktılara ulaşması durumunu ifade eder (24). Son yıllarda makinelerin öğrenmesini olanaklı hale getiren çeşitli bilgisayar algoritmaları kullanılmaya başlandı. Ayrıca makine öğrenmesi teknikleri hali hazırda günlük hayatımıza fazlasıyla girmiş bulunmaktadır. Makine öğrenmesinin günümüzde, e-mail spam filtreleri, arama motorlarında gördüğümüz arama önerileri listesi, sosyal medyadaki arkadaş önerileri, internet tarayıcılarında arama geçmişimizi kullanan reklamlar, akıllı telefonlardaki ses tanıma ve fotoğraf makinele- rindeki yüz tanıma sistemleri gibi birçok popüler kullanım alanı mevcuttur (17). Son yıllarda yapay zeka tıp literatüründe de kendine önemli bir yer edinmiş olup her geçen gün tanıdan tedaviye kullanım alanının boyutlarını daha da büyütmektedir.

Bütün tıp uzmanlık alanları içerisinde nöroşirurji, teknoloji ve bilimden her zaman ilk etkilenenlerden biri olmuştur. Giderek artan bir ivmeyle bütün tıp literatüründe olduğu gibi nöroşirurjide de yapay zeka uygulamalarına olan ilgi artmıştır (25) (Şekil 1). Yapay zeka uygulamalarının bu denli

ilerlemesi her yeni teknolojik gelişmede olduğu gibi umut ve endişeyi bir arada getirmektedir. Son yıllarda yapay zekanın nöroşirurji alanındaki verimliliğini ortaya koyan çok sayıda yeni literatür bulunmakta olup bu yazıda özellikle okurlara makine öğrenmesi yöntemlerini kullanan yazılımlar ile klinik uzmanların performanslarını karşılaştıran çalışmalara dikkat çekilmek istenmiştir. Bu literatür derlemesinin temel amacı yapay zekanın, nöroşirurji içerisindeki potansiyel kullanım alanlarını ve çeşitlerini anlatarak bu popüler alanda okuyucunun merakını gidermektir. İkincil amacı ise insan ve makinenin birbirleri ile kıyasını ve performansını analiz etmek ve son olarak da yapay zekanın geleceğini ve beraberinde getirdiği endişeleri okuyucunun huzurunda tartışmaktır.

Makine Öğrenmesi

Yapay zeka uygulamaları farklı problemlere değişik matematiksel algoritmalar ile farklı çözümler sunabilmektedir. Yapay zeka algoritmaları denetimli ve denetimsiz olarak sınıflayabileceğimiz iki temel mekanizma yardımı ile işlem görmektedirler (4,11). Denetimli makine eğitimi uygulamalarında yapay zekanın bir eğitim ve gelişim süreci mevcuttur. Bu eğitim sürecinde çıktılar bilgisayara verilerle, bilgisayar yazılımının girdiler ve çıktılar arasındaki ilişkiyi öğrenmesi beklenir (11). Yazılım bu süreçte her belirli girdi için tek tek kodlanmak yerine neden-sonuç arasındaki ilişkiyi öğrenecek şekilde kodlanıp, kendi işleyiş algoritmasını oluşturması sağlanır. Bu eğitim sürecinin sonunda yazılım, çıktıları sadece deneten kişinin bildiği bir teste tabi tutulur (24). Eğer bu testin sonunda yapay zeka uygulaması ham verileri kullanarak yüksek bir isabet oranında doğru çıktıları tahmin edebilecek hale gelirse bu yazılım gerçek



Şekil 1: Pubmed medikal literatür arama motorunda, “artificial intelligence” ve “neurosurgery” kelimeleri birlikte aratıldığında yıllara göre çıkan sonuçların karşılaştırmalı grafiği. Son yıllarda bu alana olan ilgi dikkat çekmektedir.

hasta modelleri üzerinde de denenebilir olarak kabul edilir (11). Denetimli modellerin eğitim süreci oldukça büyük bir önem arz eder. Veri sayısının hasta sayısına oranının büyük olduğu durumlarda rastgele hatalar veya anlamsız ilişkiler gelişmesi olağandır. Bu gibi durumların önlenmesi için bu algoritmaların belirli sayıda hastanın bulunduğu eğitim havuzlarında belirli sürelerde çalıştırılması gerekmektedir (25).

Denetilmeyen yapay zeka uygulamalarında ise amaç bir neden-sonuç ilişkisini ortaya koymak değildir. Bu yöntemde ise amaç karmaşık ve büyük bir veri havuzundaki veriler arasında insan gözünün kolaylıkla belirleyemeyeceği nesnelere arasındaki benzerlikleri ve farklılıkların belirlenmesidir (7,24,25). Tıp'ta kullanımına değinecek olursak bu yöntem özellikle karmaşık hasta verilerinin içerisinden hastalıkların altında yatan patofizyolojiyi anlamamızda yardımcı olacaktır. Bu duruma örnek vermek gerekirse, hastanın görüntülerinde görülen tümörün radyolojik boyutu, hastanın kan elektrolit düzeyleri, EKG verileri, hastanın ebeveynlerinden birisinin genetik bir hastalığının olması bilgisi gibi birbiriyle tamamen alakasız gibi görülen veriler arasında bir ilişki kurmayı başarıp hastalıkların tanı ve tedavisinde gelecekte yeni yöntemler keşfetmemize yardımcı olabilirler (7). Bu yöntem örnek vermek gerekirse beyin tümörlü hastaların genetik verileri arasında bir ilişki kurmayı ve hatta tümörlerin genetik altyapısını keşfetmemizi olanaklı hale getirebilir. Denetilmeyen algoritmalarda, eğitim süreci veya örnek boyutu gibi bir sınır veya kısıtlama bulunmamaktadır.

Geleneksel programlamalarda bir yazılımcı kodları yani kuralları belirler ve istenen çıktılara çeşitli girdiler vasıtasıyla ulaşılmasını sağlar. Makine öğrenmesinde ise öncelikli hedef yazılımın çıktılarına değil çıktılarına ulaşmayı sağlayan bağlantı yollarının program tarafından öğrenilmesidir. Beyin ve sinir cerrahisi biliminde bu çıktılarına örnek olarak manyetik ve rezonans görüntüleme (MRG) görüntülerindeki tümör boyutunda preoperatif ve postoperatif çekimler arasındaki değişimler, yoğun bakım hastalarının intrakranial basınç (ICP) değerlerinin takip sürecindeki değişimi ve gelecek günlerdeki değişimlerinin tahmin edilmesi, spektroskopik incelemelerde tümörün doğası hakkında ön bilginin elde edilmesi gösterilebilir (4).

Yukarıda da bahsedildiği gibi makine öğrenmesi için çok sayıda matematiksel algoritmalar geliştirilmiştir. Artificial neural network (ANN), support vector machine (SVM), decision tree, K-nearest neighbors (KNN), Naive Bayes, Fuzzy C-means (FCM) ve logistic regression bunlar arasında sınıflandırılabilir (7,17).

Artificial neural network, yani yapay nöral ağlar tıpkı canlı sinir hücrelerinin çalışma prensibine benzer bir şekilde katmanlar halinde organize edilmiş kontrol noktalarından oluşan bir ağ sistemini oluşturur. Girdiler ve çıktılar arasında gizli bir katman bulunmaktadır ve bu katmandaki kontrol noktaları her yeni veri için değişkenlik gösterir. Özellikle eğitim süreci sonunda kompleks ilişkileri kolaylıkla ortaya koyabilmesi ve her yeni veri ile tekrar eğitilebilir olmaları sağladığı avantajlardandır (7,17,24). Ancak, yüksek miktarda bellek ve işlem gücü gerektirir ve küçük boyutlardaki veri havuzlarıyla çalışırken yüksek hata oranına sahip olabilmektedirler.

Support vector machine, yani destek vektör makinesinde ise karmaşık veri havuzları içerisinde birbiri ile ilişkisi olmayan verileri ayırarak kümeleştirilmesi ve iki boyutlu veri havuzlarında kernel matematik fonksiyonlarını kullanarak onları üç boyutlu hale getirebilmesi, bu sayede ise aralarında farklılık bulunmuyor olarak gözüken verileri birbirinden ayırabilmesi ile diğer algoritmalara göre avantaj sağlar. Çok boyutlu ve karışık, düzensiz verilerle çalışırken kolaylık sağlanmasına karşın bu yöntem de yine yüksek miktarda bellek ve işlem gücü gerektirmektedir (24,25).

Decision tree, yani karar ağacında ise diğerlerinden farklı olarak aşamalı bir şekilde seçimler yapacak şekilde algoritma eğitilir ve her aşamada mevcut seçimlerin ihtimallerini göz önünde bulundurarak algoritma yeni olasılıklara ulaşır. Bu yöntemde ise teorik olarak her yeni veri girişi ile yeni olasılıklar ve seçimler oluşacak ve bu seçimler arasından en muhtemel doğru sonucu algoritma seçecektir. Uygulaması görece hızlıdır, kolaydır ancak karmaşık ve büyük verilerle çalıştırılmaya uygun değildir. Random forest, yani rastgele orman yöntemi ise karar ağacı yönteminin geliştirilmiş bir versiyonu olup çok sayıda karar ağacı algoritmasının bir arada çalışarak daha büyük verilerle çalışmasını olanaklı kılmaktadır. Son yıllarda popülerliği giderek artmaktadır (7,24,25).

K-nearest neighbor, yani en yakın komşu yöntemi olarak bilinen bu yöntemde birbirlerine en yakın değerlerdeki veriler kümeleştirilerek çok sayıda farklı grup oluşması sağlanmaktadır. Bu algoritma birbirine en yakın iki değeri birbiriyle eşleyerek girdileri kendi aralarında sınıflandırmaktadır. Doğrusal olmayan verileri sınıflamada başarılı ve uygulama yöntemi olarak diğerlerine göre görece kolay kabul edilir (17,24,25).

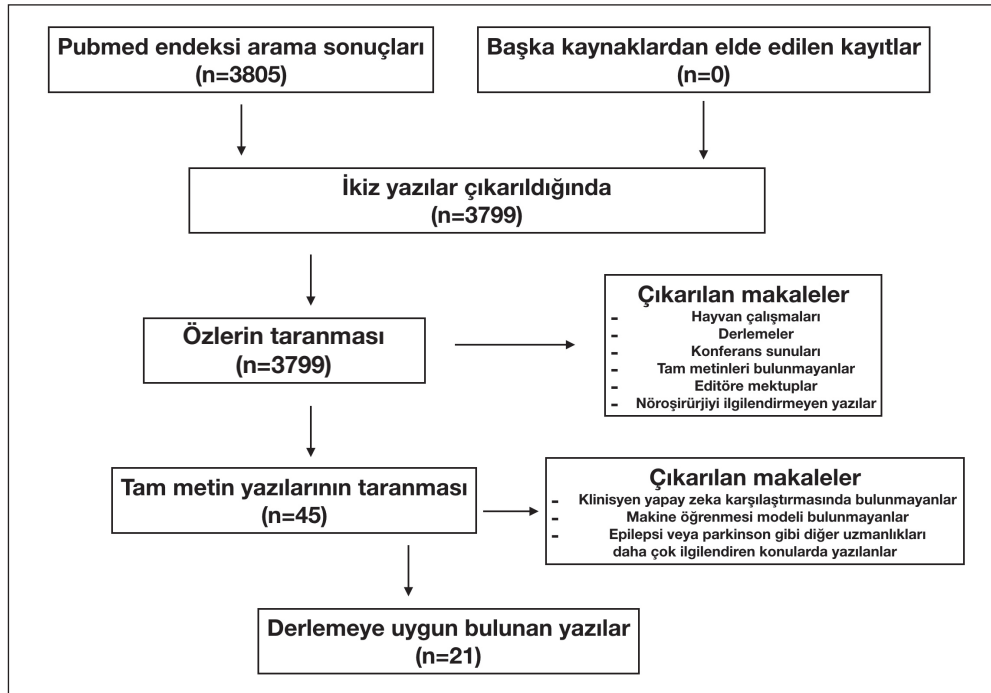
Naive Bayes algoritmaları girdileri kullanarak en olası çıktıları listeler. Kullanıcıya olasılıksal bir çözüm sunar. Uygulaması kolay ve hızlıdır ancak küçük hacimdeki veriler ile uyum sağlar (5,6).

Fuzzy C-means, denetimsiz tip bir yapay zeka algoritması olup çoklu veri kümeleri arasında bir bağlantı olup olmadığını inceler. Eğitim sürecine gerek duymaz ve gizli bağlantıları kullanıcıların farketmesine olanak sağlar (7,24).

Lojistik regresyon modellerinde ise bütün veriler doğrusal bir düzlem içerisinde değerlendirilir. Ancak bu yöntem kullanım alanı olarak yalnızca doğrusal girdiler içeren örnekler ile sınırlıdır ve bir kez kurulduğunda yeni girdilerin sisteme girmesine olanak sağlamaz (7,17).

■ YÖNTEM

Pubmed veri tabanında sistematik bir tarama yapılarak Preferred Reporting Items for Systematic Reviews (PRISMA) kılavuzuna uygun olarak Nisan 2018 tarihi itibarı ile "neurosurgery" ve "artificial intelligence" ana kelimelerini de içeren geniş bir arama yelpazesinde bütün makaleler tarandı (Şekil 2) (21). Bu alanda önceden yapılmış derlemeler ve meta-analizler incelendi (7,24,25). Histolojik tanı, radyolojik tanı ve prognosis tahminleri adı altında bu yapılan çalışmalar kategorize edildi. İncelemeye sadece İngilizce dilinde yapılmış çalışmalar dahil edildi. Yapılan çalışmalar içerisinde klinisyen ve yapay zeka



Şekil 2: PRISMA akış diyagramı kullanılarak yapılan literatür taraması.

yazılımlarının karşılaştırıldığı makaleler seçilerek istatistiksel olarak yeterli büyüklükteki ve uygun istatistiksel karşılaştırma ve analiz yöntemlerinin kullanıldığı makaleler kullanıldı. Dışlanma kriterleri; epilepsi ve elektroensefalografi (EEG) üzerine yapılan çalışmalar nöroşirurji pratiğinden çok nöroloji pratiğinde faydalı görüldüğünden çalışma dışında bırakıldı. Ayrıca sadece yapay zeka algoritmalarının performansının birbirleriyle karşılaştırıldığı çalışmalar da bu derlemenin konusu dışında kaldığından değerlendirmeye alınmadı.

Tam metinlerine erişilemeyen, konferans yazıları, poster yazıları, editöre mektuplar, yazım dili İngilizce olmayan makaleler ve hayvan çalışmaları hariç tutuldu.

Elde edilen verilerde 1) yayınlanma yılı 2) ilk yazarın ismi 3) kullanılan yapay zeka modeli 4) girdiler 5) çıktılar 6) kullanılan istatistiksel yöntem ve 7) yazarın vardığı sonuç, olarak yedi farklı kategoride değerlendirildi.

Nöroşirurji alanında çok çeşitli konularda farklı yöntem ve büyüklükte çalışmalar olduğu ve bunları birbirleriyle doğrusal olarak kıyaslamak mümkün olmadığı için her bir makalede kullanılan kendi istatistiksel yöntemine göre ortalama gelişim eğrileri, isabet oranları ve alıcı çalışma eğrisi altındaki alan gibi farklı değerler kendi içlerinde değerlendirildi. Bu yüzdeler içerisinde yazılımların ve klinisyenlerin birbirlerine olan performansları yüksek veya eşit olarak değerlendirildi. Yazarların istatistiksel analizlerin sonucunda vardığı anlamlı veya anlamsız sonuçlar analiz edildi. Bu sonuçlara göre yüksek performans ($p \leq 0,5$) veya eşit performans ($p > 0,5$) olarak değerlendirildi.

■ BULGULAR

Arama sonuçlarında toplamda 3805 makaleye ulaşıldı. Birbirleriyle aynı olan yazılar çıkarıldığında geriye kalan makalelerin özleri tarandı ve bu tarama neticesinde 45

adet makale değerlendirildi. Bu derlemenin amacına uygun olmayan 24 makale elendi ve klinisyenler ile yapay zekanın karşılaştırıldığı 21 farklı çalışma ayrıntılı olarak incelendi (Tablo I) (1-3,5-6,8-10,12-15,18-20,22,23,26-29).

On dokuz çalışmada denetlenen tip algoritma kullanılmış olup 2 çalışma denetlenmeyen algoritmalar yoluyla yapılmıştır (9,13). En sık uygulanan algoritmalar ise sırası ile 9 çalışma ile ANN ve 7 çalışma ile SVM olarak belirlenmiştir (1-3,5,6,8,10,12,14,15,18,23,26,28,29). Diğer algoritmalar 3 random forest, 2 fuzzy C-means, ve 1 naive bayes ve 1 lojistik regresyon olarak göze çarpmıştır (8-10,13,19,22,27). Çalışmalarda yer alan klinik uzmanlar; beyin ve sinir cerrahları, nörologlar, radyologlar ve acil tıp uzmanları ve diğer uzmanlar olarak gözlenmiştir.

Makine eğitimi modelleri ile klinisyenlerin performanslarının karşılaştırıldığı bu çalışmaların her birinde farklı boyutlarda örnekler kullanılmış olup her birinde kıyaslanan istatistiksel yöntem farklıdır. Bu heterojen yapı bu istatistiksel veriler arasında bir korelasyon yapmayı olanaksız hale getirmektedir. Ancak bu çalışmalar incelendiğinde en sık kullanılan istatistiksel yöntemler sırası ile alıcı çalışma eğrisi altındaki alan (AUC) ve isabet oranı olarak göze çarpmaktadır (1-3,8,13-15,18-20,23,26-29). ($p \leq 0,5$) olduğu durumlarda klinisyenler ve makine öğrenmesi modelleri arasında yazarlar tarafından performans farkı olduğu ortaya koyulan yazılar ile ($p > 0,5$) kabul edilen birbirlerine karşı anlamlı farklılık bulunmadığını ortaya koyan yazarların yazdığı makaleler farklı kategorilerde değerlendirilmiştir. Buna göre çalışma sonuçlarında %29 (6/21) oranında klinisyenler ile yapay zeka uygulamaları arasında performans farklılığı bulunmamaktadır (2,3,8,13,20,22); %71 (15/21) oranında ise yapay zeka yazılımları klinik uzmanlardan daha yüksek performans göstermiştir (1,5,6,9,10,12,14,15,18,19,23,26-29). Analiz edilen 21 çalışmanın herhangi birinde klinik uzmanlar

Tablo I: Yirmi bir Farklı Makalede Kullanılan Yapay Zeka Modeli, Girdileri, Çıktıları, İstatistiksel Karşılaştırma Yöntemi ve Makale Yazarlarının Vardığı Sonuçların Karşılaştırılması

Yıl	İlk yazar	Kullanılan yapay zeka modeli	Girdiler	Çıktılar	Kullanılan istatistiksel kıyaslama yöntemi	Yazarların vardığı sonuç
1992	Floyd (15)	ANN	Spektroskop	Lezyonun tanımlanması	AUC	YZ daha iyi performans göstermiştir
1995	Christy (8)	ANN, LR	MRG	Glial tümörlerin sınıflaması	İsabet oranı, sensitivite, spesivite	Eşit performans göstermişlerdir
1997	Arle (3)	ANN	Yaş, Cinsiyet, MRG	Pediyatrik posterior fossa tümörlerinin ayrılması	İsabet oranı	Eşit performans göstermişlerdir
1997	Abdolmaleki (1)	ANN	MRG	Glial tümörlerin sınıflaması	İsabet oranı, AUC	YZ daha iyi performans göstermiştir
1998	Clarke (9)	FCM	MRG	Glial tümörlerin segmentasyonu	Spearmen korelasyonu	YZ belirgin daha iyi performans göstermiştir
2001	Sinha (26)	ANN	BT	Travmatik beyin hasarlı çocuklarda CT anormalliklerini saptama	İsabet oranı, sensitivite, spesivite, AUC	YZ daha iyi performans göstermiştir
2004	Bidiwala (5)	ANN	Yaş, Cinsiyet, Semptomlar, Belirtiler, BT, MRG	Pediyatrik posterior fossa tümörlerinin ayrılması	Sensitivite, spesivite, pozitif prediktif değer	YZ daha iyi performans göstermiştir
2008	Yamashita (28)	ANN	Yaş, hikaye, MRG	Glial tümörlerin sınıflaması	AUC	YZ daha iyi performans göstermiştir
2009	Emblem (13)	FCM	MRG	Glial tümörlerin sınıflaması ve yaşam süresi tahmini	AUC	Eşit performans göstermişlerdir
2010	Juntu (18)	SVM, ANN	MRG	Benign ve malign yumuşak doku tümörlerinin sınıflaması	İsabet oranı, sensitivite, spesivite, AUC	YZ daha iyi performans göstermiştir
2010	Zhao (29)	SVM	Yaş, MRG	Glial tümörlerin sınıflaması	İsabet oranı, AUC	YZ belirgin daha iyi performans göstermiştir
2010	Rughani (23)	ANN	MRG	Glial tümör hastalarının hayatta kalma süresi	AUC	YZ belirgin daha iyi performans göstermiştir
2013	Campillo-Gimenez (6)	SVM	Serbest metin	Cerrahi yara yeri enfeksiyonunun tanınması	Sensitivite, pozitif prediktif değer	YZ daha iyi performans göstermiştir
2014	Emblem (14)	SVM	MRG	Glial tümör hastalarının hayatta kalma süresi	AUC	YZ daha iyi performans göstermiştir
2016	Dolz (12)	SVM	MRG	Trigeminal nevraljide beyin sapının segmentasyonu, beyin metastazları ve beyin sapı kavernom hastalarının sınıflandırılması	Hacim farkı yüzdesi, hız	YZ daha iyi performans göstermiştir
2016	Cohen (10)	SVM, NB	Serbest metin	Pediyatrik epilepsi hastalarında cerrahiye uygun adayların belirlenmesi	F-ölçümü	YZ daha iyi performans göstermiştir
2016	Porz (22)	Random forest	MRG	Glial tümörlerin segmentasyonu	Pozitif prediktif değer, sensitivite	Eşit performans göstermişlerdir
2017	Alcaide-Leon (2)	SVM	MRG	SSS lenfoması glial tümör ayırımı	AUC	Eşit performans göstermişlerdir
2017	Kellner-Weldon (20)	Belirtilmemiş	MRG	Glial tümör boyutunun hesaplanması	İsabet oranı	Eşit performans göstermişlerdir
2018	Kang (19)	Random forest	MRG	Tanı performansı	AUC	YZ daha iyi performans göstermiştir
2018	Suh (27)	Random forest	MRG	Beyin tümörünün boyutu, şekli ve doğası	AUC, pozitif prediktif değer, sensitivite	YZ daha iyi performans göstermiştir

ANN: Artificial Neural Network, **SVM:** Support Vector Machine, **NB:** Native Bayes, **MRG:** Manyetik Rezonans Görüntüleme, **AUC:** Alıcı Çalışma Eğrisi Altında Kalan Alan, **LR:** Lojistik Regresyon, **FCM:** Fuzzy C-Means.

makine eğitimi algoritmalarına karşı istatistiksel olarak anlamlı bir üstünlük gösterememiştir. Bu sonuçlar daha önceki bu alanda yapılan çalışmalar ile korelasyon gösterse de güncel verileri içermesi ve özellikle epilepsi ve parkinson alanlarında yapılan çalışmaların hariç tutulması ile diğerlerinden farklılık göstermektedir (24).

Makalelerin %90'ı (19/21) tümörlerin radyolojik sınıflamasına dayalı olup bunlar içerisinde glial tümörlerin sınıflaması, segmentasyonları, glial tümörler ile lenfomaların ayırımı ve lezyonların tanımlanması gibi farklı alanlardaki çalışmalar göze çarpmaktadır (1-3,5,8,9,12-15,18-20,22,23,26-29). Pediatrik epilepsi hastalarında cerrahiye uygun adayların belirlenmesi, cerrahi yara yeri enfeksiyonun tanımlanması gibi farklı alanlarda da makine eğitimi algoritmaları yazarlar tarafından sinanmıştır (6,10). Çalışmaların %81'inde (17/21) MRG verileri girdi olarak kullanılmış olup, 2 çalışmada serbest metin yazıları, 3 çalışmada yaş, cinsiyet ve semptom verileri,1 çalışmada ise MRG spektroskopisi verileri kullanılmıştır (1-3,5,6,8-10,12-15,18-20,22,23,27-29).

■ TARTIŞMA

Bildiğimiz kadarıyla bu sistematik analiz insan ve makine eğitiminin karşılaştırıldığı en kapsamlı ve güncel analizdir. Daha önce Senders ve ark. yaptığı çalışmadan farklı olarak daha güncel olması nedeniyle 4 yeni çalışma da dahil edilmiştir (2,19,20,24,27). Ayrıca epilepsi ve parkinson gibi öncelikli olarak nöroşirurjiyenlerden çok nörologların uzmanlığını da ilgilendiren çalışmalar hariç tutulmuştur. Bu özelliği ile en yakın gösterilebilecek diğer derlemelerden ayrılmaktadır (7,24,25). Bu sistematik analizin amacı nöroşirurji alanında son yıllarda ivmeli bir gelişim gösteren yapay zekanın klinik uzmanlar ile performanslarını karşılaştırabilmektir. Bu çalışmalar incelendiğinde makine eğitimi algoritmaları klinik uzmanlara kıyasla belirgin bir avantaj sağlamaktadırlar. Ancak çalışmalar incelendiğinde mevcut teknolojik eğilim ve olanaklar sıklıkla MRG görüntülerini anlama ve yorumlama üzerine yapılmıştır. Klinik karar verme, sonuçları tahmin etme ve uygulama alanlarında henüz yeterli sayıda makale bulunmamaktadır. Belirli sayıda ve sınırlı çeşitlilikte kullanılan veri girdileri nöroşirurjiyenlerin günlük pratiğinde oldukça küçük bir yer kaplamaktadır. Mevcut sonuçlar cerrahın karar verme sürecine yardımcı olabilecek seviyeye ulaşmamıştır. Bilgisayar yazılımlarının insan beynine oranla hesaplama kapasitesi ve hızının daha üstün olduğu hali hazırda bilinen bir gerçektir. Ancak makinelerin klinik uygulamalarda insanların yerini alabilmesi, onların hastalarla empati kurabilen, hastaların sayılardan oluşmadığını anlayabilen ve yalnızca mantıksal değil duygusal kararlar almayı başarabilen birer organizma haline dönüşmeden başarabilmesi mümkün değildir.

Yapay zeka, fotoğraflardaki yüz tanıma sistemleri, yol navigasyonları vb. çeşitli örneklerde olduğu gibi günlük kullanıma çok hızlı bir şekilde uyum sağlamıştır. Gelecekteki on yıllar içerisinde tıp alanındaki günlük kullanım sıklığı da artacak ve öncelikle triyaj, radyoloji, patoloji gibi alanlardaki hekim ve yardımcı sağlık personelinin iş alanında bir alternatif çözüm olarak kullanılmaya başlanacaktır. Sağlık harcamalarının fazlalığı bütün dünyada hükümetlerin çözüm

aradığı önde gelen problemlerden biri haline gelmiştir. Yapay zeka sayesinde personel giderlerinin azaltılabilmesi bu süreci hızlandıracaktır. Günümüzde bu alanda yapılan çalışmalar her ne kadar yeni görünsün de, mevcut ivmelenme göz önünde bulundurulduğunda akıllı bilgisayar yazılımlarıyla idare edilen hastanelerin çok yakın bir gelecekte var olacağını göz önünde bulundurmamız gerekmektedir. Yirmibirinci yüzyıldaki bu akıllı bilgisayar yazılımları yarışında yerimizi almamız ve bu teknolojik devrimi elimizde tutan taraflardan biri haline gelmemiz gerekmektedir.

Yakın gelecekte bütün tıp literatüründe olduğu gibi beyin ve sinir cerrahisinde de bu alanda yapılan çalışmaların daha da hız kazanacağı aşikardır. Yeni algoritmalar, yeni uygulama yöntemleri daha isabetli ve doğru kararlar alınmasını sağlayacaktır. Veri havuzları büyüdükçe insan zihni bu devasa büyüklükteki verileri işlemede ve onları anlamlandırmada yetersiz kalacaktır. Her geçen gün yapay zekaya ve diğer bilgisayar yazılımlarına olan ihtiyacımız daha da artacaktır. Yapay zeka uygulamaları karmaşık hastalıkların altında yatan sebepleri anlamamızı ve onların tedavilerini keşfetmemizi kolaylaştıracaktır. Ayrıca yapay zeka uygulamalarının getireceği bir diğer yenilik klinisyenlerin yeni ve daha isabetli sınıflama yöntemlerini bulması ve kullanması olacaktır. Günlük rutinlerimizde kullandığımız birkaç parametreden oluşan tanı ve tedavi algoritmaları yerini daha karmaşık ve hastaya özel tanı algoritmalarına bırakacaktır.

■ Kısıtlamalar

Çok olasıdır ki burada yer alan çalışmalar makine öğrenimi yazılımlarının üstün olduğu alanlarda uygulanmıştır. Bu sebeple yapay zeka normalde olduğundan daha abartılmış sonuçlar göstermiş olabilir. Ayrıca ancak yapay zeka ve klinisyenler karşılaştırılabildikleri alanlarda değerlendirilmiştir. Çalışmalar tek tek incelendiğinde herbirinin p değeri yoktur ve bu sorun bütün bu çalışmaların birbirleriyle karşılaştırılmasını zorlaştırmaktadır ve güvenilirliğini sorgulamaktadır. Yapay zeka ve insanları objektif bir ölçekte birbirleriyle kıyaslamak mümkün değildir. Buna rağmen bu incelemenin amacı mevcut çalışmaların yönüne ışık tutmak ve kıyaslayabilmenin bir derece mümkün olduğu alanlardaki performansları analiz edebilmektir.

■ SONUÇ

Bütün bu veriler ışığında yapay zekanın hekimlere alternatif bir çözüm değil onlara yardımcı birer araç görevi göreceğini bilmek gerekir. Yapılan çalışmaların klinisyen, yazılım ve klinisyen + yazılım olarak üç farklı kategoride değerlendirenlerin her birinde klinisyenlerin yapay zeka kullanarak karar verdiği çözümlerin diğerlerinden daha başarılı olduğu görülmüştür. Yaşadığımız yüzyıl yeni teknolojik gelişmeler ile birlikte onlara ayak uydurabilen hekimlerin başarılarının konuşulduğu bir yüzyıl olacaktır. Beyin ve sinir cerrahisinin geleceğini de yapay zeka kavramından bağımsız bir şekilde ele almak mümkün değildir.

■ KAYNAKLAR

1. Abdolmaleki P, Mihara F, Masuda K, Buadu LD: Neural networks analysis of astrocytic gliomas from MRI appearances. Cancer Lett 118(1):69-78, 1997

2. Alcaide-Leon P, Dufort P, Geraldo AF, Alshafai L, Maralani PJ, Spears J, Bharatha A: Differentiation of enhancing glioma and primary central nervous system lymphoma by texture-based machine learning. *AJNR Am J Neuroradiol* 38(6):1145-1150, 2017
3. Arle JE, Morriss C, Wang ZJ, Zimmerman RA, Phillips PG, Sutton LN: Prediction of posterior fossa tumor type in children by means of magnetic resonance image properties, spectroscopy, and neural networks. *J Neurosurg* 86(5):755-761, 1997
4. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S, Montazeri A: Artificial neural networks in neurosurgery. *J Neurol Neurosurg Psychiatry* 86(3):251-256, 2015
5. Bidiwala S, Pittman T: Neural network classification of pediatric posterior fossa tumors using clinical and imaging data. *Pediatr Neurosurg* 40(1):8-15, 2004
6. Campillo-Gimenez B, Garcelon N, Jarno P, Chaplain JM, Cuggia M: Full-text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. *Stud Health Technol Inform* 192:572-575, 2013
7. Celtikci E: A systematic review on machine learning in neurosurgery: The future of decision-making in patient care. *Turk Neurosurg* 28(2):167-173, 2018
8. Christy PS, Tervonen O, Scheithauer BW, Forbes GS: Use of a neural network and a multiple regression model to predict histologic grade of astrocytoma from MRI appearances. *Neuroradiology* 37(2):89-93, 1995
9. Clarke LP, Velthuis RP, Clark M, Gaviria J, Hall L, Goldgof D, Murtagh R, Phuphanich S, Brem S: MRI measurement of brain tumor response: Comparison of visual metric and automatic segmentation. *Magn Reson Imaging* 16(3):271-279, 1998
10. Cohen KB, Glass B, Greiner HM, Holland-Bouley K, Standridge S, Arya R, Faist R, Morita D, Mangano F, Connolly B, Glauser T, Pestian J: Methodological issues in predicting pediatric epilepsy surgery candidates through natural language processing and machine learning. *Biomed Inform Insights* 8:11-18, 2016
11. Deo RC: Machine learning in medicine. *Circulation* 132(20):1920-1930, 2015
12. Dolz J, Betrouni N, Quidet M, Kharroubi D, Leroy HA, Reyns N, Massoptier L, Vermandel M: Stacking denoising auto-encoders in a deep network to segment the brainstem on MRI in brain cancer patients: A clinical study. *Comput Med Imaging Graph* 52:8-18, 2016
13. Emblem KE, Nedregaard B, Hald JK, Nome T, Due-Tonnessen P, Bjornerud A: Automatic glioma characterization from dynamic susceptibility contrast imaging: Brain tumor segmentation using knowledge-based fuzzy clustering. *J Magn Reson Imaging* 30(1):1-10, 2009
14. Emblem KE, Pinho MC, Zöllner FG, Due-Tonnessen P, Hald JK, Schad LR, Meling TR, Rapalino O, Bjornerud A: A generic support vector machine model for preoperative glioma survival associations. *Radiology* 275(1):228-234, 2014
15. Floyd CE Jr, Tourassi GD: An artificial neural network for lesion detection on single-photon emission computed tomographic images. *Invest Radiol* 27(9):667-672, 1992
16. Ghahramani Z: Probabilistic machine learning and artificial intelligence. *Nature* 521:452-459, 2015
17. Jordan MI, Mitchell TM: Machine learning: Trends, perspectives, and prospects. *Science* 349(6245):255-260, 2015
18. Juntu J, Sijbers J, De Backer S, Rajan J, Van Dyck D: Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. *J Magn Reson Imaging* 31(3):680-689, 2010
19. Kang D, Park JE, Kim YH, Kim JH, Oh JY, Kim J, Kim Y, Kim ST, Kim HS: Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma: Development and multicenter external validation. *Neuro Oncol* 2018 (Epub ahead of print)
20. Kellner-Weldon F, Stippich C, Wiest R, Lehmann V, Meier R, Beck J, Schucht P, Raabe A, Reyes M, Bink A: Comparison of perioperative automated versus manual two-dimensional tumor analysis in glioblastoma patients. *Eur J Radiol* 95:75-81, 2017
21. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 6(7):e1000097, 2009
22. Porz N, Habegger S, Meier R, Verma R, Jilch A, Fichtner J, Knecht U, Radina C, Schucht P, Beck J, Raabe A, Slotboom J, Reyes M, Wiest R: Fully automated enhanced tumor compartmentalization: Man vs. machine reloaded. *PLoS One* 11(11):e0165302, 2016
23. Rughani AI, Dumont TM, Lu Z, Bongard J, Horgan MA, Penar PL, Tranmer BI: Use of an artificial neural network to predict head injury outcome. *J Neurosurg* 113(3):585-590, 2010
24. Senders JT, Arnaout O, Karhade AV, Dasenbrock HH, Gormley WB, Broekman ML, Smith TR: Natural and artificial intelligence in neurosurgery: a systematic review. *Neurosurgery* 83(2):181-192, 2018
25. Senders JT, Zaki MM, Karhade AV, Chang B, Gormley WB, Broekman ML, Smith TR, Arnaout O: An introduction and overview of machine learning in neurosurgical care. *Acta Neurochir (Wien)* 160(1):29-38, 2018
26. Sinha M, Kennedy CS, Ramundo ML: Artificial neural network predicts CT scan abnormalities in pediatric patients with closed head injury. *J Trauma* 50(2):308-312, 2001
27. Suh HB, Choi YS, Bae S, Ahn SS, Chang JH, Kang SG, Kim EH, Kim SH, Lee SK: Primary central nervous system lymphoma and atypical glioblastoma: Differentiation using radiomics approach. *Eur Radiol* 28(9):3832-3839, 2018
28. Yamashita K, Yoshiura T, Arimura H, Mihara F, Noguchi T, Hiwatashi A, Togao O, Yamashita Y, Shono T, Kumazawa S, Higashida Y, Honda H: Performance evaluation of radiologists with artificial neural network for differential diagnosis of intra-axial cerebral tumors on MR images. *AJNR Am J Neuroradiol* 29(6):1153-1158, 2008
29. Zhao ZX, Lan K, Xiao JH, Zhang Y, Xu P, Jia L, He M: A new method to classify pathologic grades of astrocytomas based on magnetic resonance imaging appearances. *Neurol India* 58(5):685-690, 2010