

PYTHON İLE TİROİT KANSER TEŞHİSİ: DERİN ÖĞRENME VE VERİ MODELLEME

Selin ŞENOCAK (21040301043)
Mühendislik Fakültesi
Yazılım Mühendisliği Öğrencisi
İstanbul Topkapı Üniversitesi, Türkiye
selinsenocak@stu.topkapi.edu.tr

Doç. Dr. Cevat RAHEBİ
Mühendislik Fakültesi
İstanbul Topkapı Üniversitesi, Türkiye
cevatrahebi@topkapi.edu.tr

Önsöz

Türkiye’de Çernobil Nükleer Santral kazasının insan sağlığına yönelik etkilerini inceleyen çok az sayıda araştırma yapılmıştır. Yapılan araştırmaların ortak noktası konuya ilişkin istatistiksel verilerin yetersizliği, karıştırıcı faktörlerin fazlalığı ile bu konunun daha ciddi bir şekilde araştırılmasının gerekliliğine yapılan vurgudur. Nükleer kazalarda yaşanan panik ve endişenin en önemli nedeni toplumun konu hakkında yeterince bilgi sahibi olmamasıdır. Gelecekte yaşanacak olası nükleer kirlilik durumlarında endişe ve paniği önlemenin yolu toplumun eğitilmesinden ve doğru biçimde bilgilendirilmesinden geçmektedir.

Amaç

26 Nisan 1986 tarihinde Sovyetler Birliği’ne bağlı Ukrayna Sovyet Sosyalist Cumhuriyeti’nin Pripyat şehri yakınlarındaki Çernobil Nükleer Santrali’nin 4 numaralı reaktöründe gerçekleşen nükleer kazanın ardından bıraktığı tiroit kanserinin, farklılaşmış tiroid kanseri tanısı alan hastalarda nüks olasılığını tahmin etmek için makine öğrenimi modellerini eğitmek [1].

Giriş

26 Nisan 1986’da Rusya Federasyonu’nda bulunan Çernobil Nükleer Santrali patlamıştır. Kaza sonrası radyoaktif saçılım başta Ukrayna, Belarus ve Rusya Federasyonu olmak üzere tüm Kuzey Yarı Küre’yi etkilemiştir. Süreçte, radyasyonla doğrudan karşı karşıya kalan halka uzun süre, kaza ile ilgili olarak resmi bir açıklama yapılmamış ve bu durum korku, panik, yalan haber ve kimi kez de radyasyona ilişkin duyarsızlığa neden olmuştur. Bu durum radyasyonun sağlık üzerine olumsuz etkilerine ek olarak, gerekli önlemlerin alınmasını geciktirdiği/engellediğinden radyasyonun olumsuz etkilerinin daha da artmasına yol açmıştır.

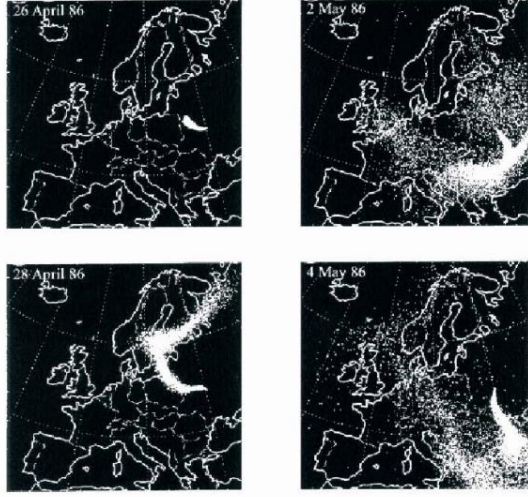
Radyoaktif maddeler ve bu maddelerin canlıların yaşam alanı üzerindeki etkilerini inceleyen Radyoekoloji bilimi, Çernobil kazası öncesinde de radyasyonla ilgili çalışmalarını sürdürmekteydi. Uzun süreli olarak dünya yüzeyinde biriken radyoaktif tortu, radyoaktif maddelerin besin zincirine giriş mekanizması ve insanlar üzerinde giderek artan etkileri bulunmaktadır [2].

Çernobil Nükleer Santral Kazasının Türkiye’ye Etkisi

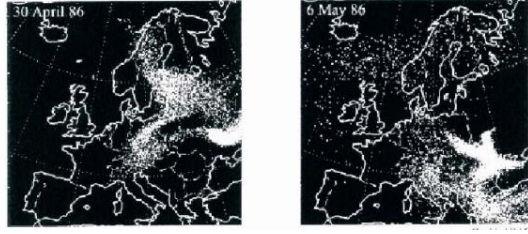
Çernobil kazasından 10 geçmişti ki radyoaktif parçacıklar tüm Türkiye’ye yayılmıştı. Radyoaktif bulutun geçtiği sırada Trakya ve Doğu Karadeniz’de; özellikle fındık, tütün ve çay üretimi yapılan tarım alanlarında yağış olması bu bölgelerdeki radyoaktif bulaşı artırdı. Kaza’dan bir hafta sonra Batı Trakya’daki ve Doğu Karadeniz Bölgesi’ndeki sağanak yağış ile bulaş tüm Türkiye’de kendini gösterdi.



Şekil 1. Dünya haritasından bir kesit.



Şekil 2. Radyoaktif bulut yayılımını gösteren Dünya haritasından bir kesit (1986).



Şekil 3. Radyoaktif bulut yayılımını gösteren Dünya haritasından bir kesit. (1986)

Bu yağışların etkisi, çevresel radyasyon ölçümlerindeki yükseklik ilk kez 30 Nisan 1986 tarihinde fark edildi.

Trakya ve Karadeniz kıyılarında normal koşullarda 8–10 microröntgen/saat olan radyasyon düzeyi; kazadan 10 gün sonra 30–40 microröntgen/saat düzeyinde ölçüldü. Yüksek seviyedeki bu ölçümler radyasyondan korunmak için zaman yitirmeden yeni önlemlerin alınmasını zorunlu hale soktu[3].

Metodoloji

Bu prospektif çalışmada veri kümesi olarak 383 hastadan oluşan bir kohort kullanıldı. Bu kohort çalışması 15 yıllık bir zaman dilimi içerisinde minimum 10 yıllık bir süre boyunca gözlemlendi. Nüks potansiyelini tahmin etmek için on üç klinikopatolojik özellik değerlendirildi.

Klasik (K-en yakın komşular, destek vektör makineleri (SVM), ağaç tabanlı modeller) ve yapay sinir ağları (ANN), üç farklı özellik kombinasyonu ile eğitildi: Amerikan Tiroid Birliği (ATA) risk puanı hariç tüm özellikleri içeren bir veri seti (12 özellik), bir diğeri yalnızca ATA riskini içerir ve üçüncüsü tüm özelliklerin birleşimidir (13 özellik). Eğitim süreci için 283 hasta, evre doğrulaması için ise 100 hasta ayrıldı [4]. Veri kümesi eksik değer içermemektedir.

Tablo.1. Veri kümesinde kullanılan değişkenler tablosu.

Değişkenler Tablosu

Değişken ismi	Rol	Tip
Yaş	Özellik	Tamsayı
Cinsiyet	Özellik	Kategorik
Sigara içmek	Özellik	Kategorik
Hx Sigara İçme	Özellik	Kategorik
Hx Radyoterapi	Özellik	Kategorik
Tiroid fonksiyonu	Özellik	Kategorik
Fiziksel Muayene	Özellik	Kategorik
Adenopati	Özellik	Kategorik
Patoloji	Özellik	Kategorik
Odaklanma	Özellik	Kategorik
Risk	Özellik	Kategorik
T	Özellik	Kategorik
N	Özellik	Kategorik
M	Özellik	Kategorik
Sahne	Özellik	Kategorik
Cevap	Özellik	Kategorik
Tekrarlanan	Hedef	Kategorik

Bu veri kümesi python ile jupyter notebook üzerinden eğitilen yapay zeka doğruluk testine tabi tutulmuştur. Test doğrultusunda sunulan grafik ve sonuçlar [web sitesinde](#) [5] detaylıca gösterilmektedir.

Gereç ve Yöntem

Bu araştırma bilgisayar aracılığıyla elektronik ortamda literatür taraması ile başlamıştır ve Andrew Leatherbarrow'ın "Chernobyl: A Stalkers' Guide" [6] kitabı ile devam etmiştir. Çok sayıda araştırmadan konuyla ilgisi en yakın olanlar seçilmeye çalışılmıştır. Araştırma, veri toplama, makalelerin incelenmesi ve tasnifi, çeviri ve özet çıkarma, derleme ve rapor yazımı aşamalarıyla yaklaşık iki ay sürmüştür.

Bu sırada kullanılan temel anahtar sözcükler "Çernobil", "Nükleer kaza", "Radyasyon ve Sağlık", "Tiroit kanseri" ve "kanser"dir. Elde edilen yazıların kaynakları metin içinde geçiş sırasına göre numaralanarak gösterilmiştir. Kaynaklar; yazarlar, soyadı, adının ilk harfi, makale başlığı, yayınlandığı kitap-dergi bilgileri ve yayın yılı olarak belirtilmiştir.

Sonuç

Elektronik ortamda basit bir arama motoru kullanıldığında Çernobil ve Çernobil ile ilişkili sağlık olaylarını içeren bir birinden farklı yorumlarla milyonlarca yazı, literatür tarandığında ise farklı sonuçlara ulaşılmış binlerce araştırma olduğu görülmektedir. Bu kaynakların, yorumların ve sonuçların hepsinin doğru olduğunu düşünmek olanaklı değildir.

Bu araştırma için incelenen literatür ışığında, konuya ilişkin olarak elde edilen sonuçların araştırma zamanına, araştırma evrenine ve araştırmanın yöntemine göre sonuçların değişebildiği izlenmiştir. Burada dikkat edilmesi gereken temel konunun, yapılan epidemiyolojik çalışmada sunulan kanıtın kalitesi, olduğu temel bilgisinden yola çıkılarak, araştırma sonuçlarının güvenilirliği tartışılmalıdır [7].

Bu çalışmada kullanılan bilimsel araştırmaların incelenmesi sonucu şu çıkarımlar yapılır:

%94.81 doğruluk oranı, modelin test veri kümesi üzerindeki performansının oldukça yüksek olduğunu göstermektedir. Bu yüksek doğruluk oranı, modelin veri setindeki örüntüleri başarılı bir şekilde öğrendiğini ve bu örüntüleri genelleme yeteneğiyle test verilerine de uygulayabildiğini işaret etmektedir.

Yüksek doğruluk oranı, modelin karmaşıklığı ve modelleme stratejisinin etkinliği hakkında önemli bilgiler verir. %94.81 gibi yüksek bir doğruluk oranı, modelin verilerdeki gizli yapıları tanımlama yeteneğinin güçlü olduğunu gösterir. Bu, modelin eğitim verilerindeki örüntüleri

ezberlemek yerine, genellemek için temel özellikleri anladığını gösterir.

Ayrıca, bu doğruluk oranı modelin gerçek dünya verileriyle etkili bir şekilde çalışabileceğini gösterir. Modelin test veri kümesi üzerindeki performansı, gerçek dünya uygulamalarında ne kadar etkili olabileceğini yansıtır. Bu, modelin gerçek dünya problemlerini çözmek için güvenilir bir araç olduğunu gösterir.

Bununla birlikte, %94.81 doğruluk oranıyla elde edilen sonuçların, modelin performansını değerlendirmek için tek kriter olmadığını belirtmek önemlidir. Modelin diğer performans metrikleri ve hiperparametrelerle birlikte değerlendirilmesi gerekir. Ayrıca, veri setinin doğası, dengesizlik, gürültü veya diğer özel durumlar da dikkate alınmalıdır.

Sonuç olarak, %94.81 doğruluk oranı, modelin başarılı bir şekilde eğitildiğini ve test veri kümesi üzerinde etkili bir şekilde çalışabildiğini gösterir. Bu yüksek doğruluk oranı, modelin güvenilirliğini ve uygulanabilirliğini vurgular, ancak daha geniş bir bağlamda değerlendirilmesi ve anlaşılması önemlidir[5].

Kaynaklar

[1] K. Sonat, Çernobil nükleer felaketi: 37 yıl önce neler yaşandı, riskler sürüyor mu?, BBC NEWS, <https://www.bbc.com/turkce/haberler-dunya-52426424> [05.03.2024]

[2] Bell JNB, Shaw G Ecological lessons from the Chernobyl accident. Environmental International 31(2005) 771-777.

[3]Özemre A.Y., Çernobil Komplosu, Bilge Yayıncılık, 2004.

[4] Borzooei, S., Briganti, G., Golparian, M. ve diğerleri. Tiroid kanseri hastalarının risk sınıflandırması için makine öğrenimi: 15 yıllık bir kohort çalışması. Eur Arch Otorinolaringol 281 , 2095–2104 (2024). <https://doi.org/10.1007/s00405-023-08299-w>

[5] <https://selinsenocak.com.tr/works.html>

[6] A. Leatherbarrow, Chernobyl: A Stalkers' Guide, 1, İstanbul, Indigo, 2019

[7] David AG, Kenneth FS. An overview of clinical research: the lay of the land. The Lancet Epidemiology Series. Lancet 2002; 359: 57-61.