

Proje Ana Alanı : Fiziki Bilimler

Proje Tematik Alanı : Meteoroloji, İklim, Atmosfer ve Uzay

Proje Adı (Başlığı) : İklim Verilerini Baz Alan ve Buzulların Erime Kütlesini Tahmin Eden Yapay Zekâ Temelli Algoritma Önerisi

Özet

Buzulların erimesi, olumsuz etkileri bakımından, günümüzün en önemli problemlerinden biri olarak değerlendirilmektedir. Buzullar eriyince deniz seviyesi yükselmekte, içindeki sera gazları salınmakta ve ekosistemi kötü yönde etkileyen birçok durum ortaya çıkmaktadır. Bu sebeple buzulların erimesine sebep olan faktörleri ve bu faktörlerin etki kuvvetini hesaplamak, böylece gelecekte ne kadarlık bir kütlenin eriyeceğini tahmin etmek önemli bir araştırma alanıdır. Sıcaklık değişimi, yağış rejimi, CO₂ salınımı gibi birçok faktör kütle değişimine sebep olabilir. Literatürde buzulların erimeye katkısı olan faktörleri inceleyen ve gelecek yıllarda oluşacak erime miktarını tahmin eden çalışmalar yer almaktadır. Fakat bu yöntemlerin başarı düzeyi çoğunlukla istenilen düzeyde değildir. Bu çalışmada makine öğrenmesi teknikleriyle çeşitli faktörlerin buzulların erimesine olan katkısı araştırılmış ve Antarktika'nın önümüzdeki 10 yıllık kütle değişimine dair bazı tahminler verilmiştir. Yapay zekâ modeli belirlenirken 10 farklı regresyon yöntemi test edilmiş, birbirlerine göre üstünlükleri değerlendirilmiş ve en uygun yöntem ile gelecek tahminleri hesaplanmıştır. Eğitim ve test aşamasında uydu verileri ile bölgede kurulu olan 18 farklı araştırma istasyonunun verileri kullanılmıştır. Bağımsız değişkenleri tahmin etmek için Prophet prosedürü ile TimeSeries Forecasting yapılmış, Cross-Validation kullanılmıştır. Elde ettiğimiz sonuçlara göre 10 yıl sonra, buzulların 2021 yılında eriyen buzul miktarınının 1.5 katı kadar daha fazla eriyeceği, su seviyesinin yaklaşık 0.6 kat artacağı sonucuna varılmıştır. Buna göre 2021 yılında eriyen buzul miktarı 2.64 gt (gigaton) iken 2032 yılında eriyecek buzul miktarınının 4.398 gt olacağı ve deniz seviyesinin 12.97 mm yükseleceği tahmin edilmiştir. Elde ettiğimiz çıktuların meydana getireceği çevresel problemlerin gündeme getirilmesiyle, kutuplardaki erimelerle ilgili olarak toplum bilincinin artırılmasına da katkı sağlanacağını düşünmekteyiz.

Anahtar kelimeler: Antarktika, Uydu verileri, Regresyon, Küresel Isınma, Buzulların Erimesi

Amaç

Bu çalışmada amacımız, çeşitli iklim verilerinden hareketle buzullardaki erime miktarını yüksek doğruluk payı ile tahmin eden yapay zekâ tabanlı bir algoritma geliştirmek, geliştirilen algoritma ile erime miktarını hesaplamaktır. Literatürde bu amaca yönelik çeşitli çalışmalar incelenmiştir. Fakat bu çalışmalarda anlaşılmıştır ki, tahminlerin doğruluk oranları çoğu zaman istenen düzeyde gerçekleşmemektedir. Literatürdeki çalışmaların birçoğunda bağımsız değişken sayılarının sınırlı tutulduğu, bu nedenle başarı oranlarının düşük olduğu öngörülmüştür. Bu sebeple bu çalışmada fazla sayıda bağımsız değişken kullanılarak 10 farklı regresyon yöntemi eğitilecek, geçmiş yılların verilerine göre doğruluk oranları test edilecektir. Ardından elde edilen sonuçlar karşılaştırılacak, en verimli yöntem belirlenip bu yönteme göre önümüzdeki 10 yıllık kütle değişim oranı hesaplanacaktır. Elde edeceğimiz verilerle toplumda farkındalığın oluşmasına dolaylı bir katkı sağlanacağı da değerlendirilmiştir.

Giriş

İklim değışikliđi, dünya apında biyoeřitliliđin kalıcılıđına ynelik en byk tehditlerden birini oluřturmaktadır. Antarktika Yarımadası ise son yıllarda Gney Yarımkuredeki en hızlı sıcaklık artıřlarından birini yařamaktadır. İklim değışikliđi, Antarktika'daki buzsuz alanların nemli lde geniřlemesine sebep olabilir, bu da muhtemel olarak bazı biyolojik trlerin geniřlemesine hatta buzsuz alanlar birleřmeye bařladıđında antik biyocođrafik sınırları gemelerine izin vererek Antarktika kıtasının daha da fazla ısınmasına neden olabilir. Yapılan arařtırmalara gre yzyılın sonunda Antarktika kıtasında 2.100 ile 17.267 km² arasında buzsuz alanın oluřacađı tahmin edilmektedir. Eđer ki global karbon emisyonları azalmazsa bu Antarktika Yarımadasında buzsuz alanın yaklaşık olarak 3 kat artacađı anlamına gelmektedir (Lee ve ark., 2017).

Bu sebeplerden dolayı Antarktika'nın ısınması son yıllarda oldukça dikkat ekmeye bařlamıřtır. Yapılan alıřmalar Antarktika'nın hızlı bir Őekilde ktle kaybettiđini ortaya koymaktadır. Batı Antarktika Buz rtsnn kritik eřik deđerini ařtıđı ve Dođu Antarktika buzullarının da eđer belirli buz tıkaları kaldırılsa istikrarsızlařacađı grlmřtr. Yapılan alıřmalar sonucunda Antarktika buzullarının erimesinin deniz seviyesi miktarlarında ykselmeye yol atıđını ve artıřın her geen yıl ile hızlanacađı ve zamanla deniz seviyesi ykselmesine sebep olan en nemli faktr olacađı bulunmuřtur. řu an deniz seviyesine etki eden en nemli faktrler termal geniřlemeler (%30) ve karasal buzların erimesidir (%~55) (Ludescher ve ark., 2016). Kutuplardaki ve yksek dađlardaki buzullar kresel ısınmanın artmasına bađlı olarak erimeye devam etmektedir. Deniz seviyesi son 20 yılda 15-20 cm ykselmiřtir (Akın, 2013).

Kresel ısınmanın ve bunun tetiklediđi buzulların erimesi, deniz seviyelerinin ykselmesi ve kuraklık gibi dođal felaketlerin artarak devam ettiđini ve halen zlmesi gereken en nemli kresel sorun olduđunu gstermektedir (Akın, 2009; Kadıođlu, 2007). Ayrıca Kutup Blgelerindeki donmuř topraklar, kresel ısınma sonucu zlmeye bařlayarak binlerce yıl boyunca bnyelerinde bulunan sera gazlarını atmosfere bırakacaktır (Akın, 2013).

Antarktika Yarımadası'nda buz sahanlıđı istikrarsızlıđı ve kaybının sebebi yazın daha yksek hava sıcaklıkları nedeniyle yzeyin erimesindedir. Radar saınımler verileri, 2000 ile 2009 yılları arasında Antarktika'nın Yarımada erimesinin, Antarktika'nın toplam yzey erime yođunluđunun %50'sinden fazlasını oluřturduđu gstermektedir (Abram ve ark., 2013).

Kutupların ilerleyen yıllardaki erime durumunun gidiřatının Őimdiden tahmin edilmesi ve buna gre nlem alınması hayati deđer tařımaktadır. Bu sebeple kutupların gelecek yıllardaki ktle deđiřiminin tahmininin yapılması iin eřitli alıřmalar bulunmaktadır. Buzul ktle dengesinin, bir blgedeki kmlatif yađıř ve sıcaklık deđiřimlerini yansıttıđı ve bu nedenle blgesel ve sinoptik lekli iklimin temel bir ifadesi olduđu belirtilmiřtir. alıřmalarda genellikle sıcaklık deđiřimi, yađıř rejimi, CO₂ (Karbon dioksit) deđiřiklikleri ile buzulların erime durumu tahmin edilmeye alıřılmıřtır (Shea ve ark., 2007; Radić ve ark., 2006; Hall ve ark., 2003).

Yerel istasyonlardan alınan sıcaklık ve yağış verileri ile doğrusal regresyon geçmişte buzul kütle dengesi değerlerini tahmin etmek için yaygın olarak kullanılmıştır (Hopkinson, 1981; Letreguilly, 1988; Moore ve Demuth, 2001). Ancak çalışmaların çoğunda görülen ortak sorun kullanılabilirlik oranının yani değişkenlerin, sonuçla bağlantısının düşük olmasıdır. Örneğin bölgesel iklimsel anomalilerin kullanıldığı ve Peyto buzunun kütle dengesinin tahmin edilmesi için yapılan bir araştırmada R2 değeri 0.68 çıkmıştır (Shea ve ark., 2007). Bu durum veri setinin, regresyon sonuçlarıyla %68 alakalı olması ve doğruluk oranının düşük olmasına yol açmaktadır. Bir başka sorun ise bu çalışmaların kutupların genelinde değil, bazı buzullar üzerinde yapılması ve sonucunda elimize kısıtlı veri geçmesidir. İsveç'te yapılan bir iklim modeli duyarlılık çalışmasında da sıcaklık ve yağış verileri kullanılmış, model sonucunda ise tahminlerin R2 değeri 0.8 civarında olmasına rağmen hata oranı yüksek çıkmıştır (Radić ve ark., 2006).

Bu amaçla yaptığımız çalışmada; Antarktika'nın 1992-2021 yılları arasındaki çeşitli hava durumu verilerinden oluşan veri seti ile, Python ve R dillerinde kodlanmış ve regresyon yöntemleri kullanılarak oluşturulmuş bir yapay zekâ modeli eğitilmiştir. Yapay zekâ çıktı olarak Antarktika'nın önümüzdeki 10 yıllık kütle değişimini yüksek kullanılabilirlik ve düşük hata oranında vermektedir. Böylece kutuplar hakkında yapılan bilimsel çalışmalara katkı sağlamanın yanında buzulların durumu ile ilgili toplumsal bilinç oluşturulması hedeflenmiştir.

Çalışmanın veri setinde NASA-GISS'dan alınan 1972-2021 yılları arası sıcaklık, Kaggle'dan alınan 1991 – 2021 yılları arası hava basıncı ve rüzgâr hızı, Our World In Data'dan alınan 1984-2021 arası karbondioksit emisyonu, IMBIE'den alınan 1992-2021 Antarktika kümülatif buzul kütle ve deniz seviyesi değişimi verileri kullanılmıştır. Bu verilerin tamamının buzulların erimesiyle bir bağlantısı olduğu tespit edilmiştir.

Veri biliminde öngörü (forecasting), tahmin yapmak ve stratejik karar vermeyi bilgilendirmek için istatistik ve modelleme kullanarak zaman serisi (Time Series) verilerini analiz etme sürecidir. Bu öngörü kesin bir tahmin değildir ve tahminlerin olasılığı değişebilir. Öngörü bize olayların yaşanma ihtimalinin diğer olayların yaşanma ihtimaline kıyasla daha fazla veya az olduğunu gösterir. Çoğu zaman, işlenen veriler ne kadar kapsamlı olursa, tahminler o kadar doğru olur. Öngörü sıklıkla hava durumu, iklim, finans, sağlık gibi alanlarda kullanılmaktadır (Tableau, 2022).

Yöntem

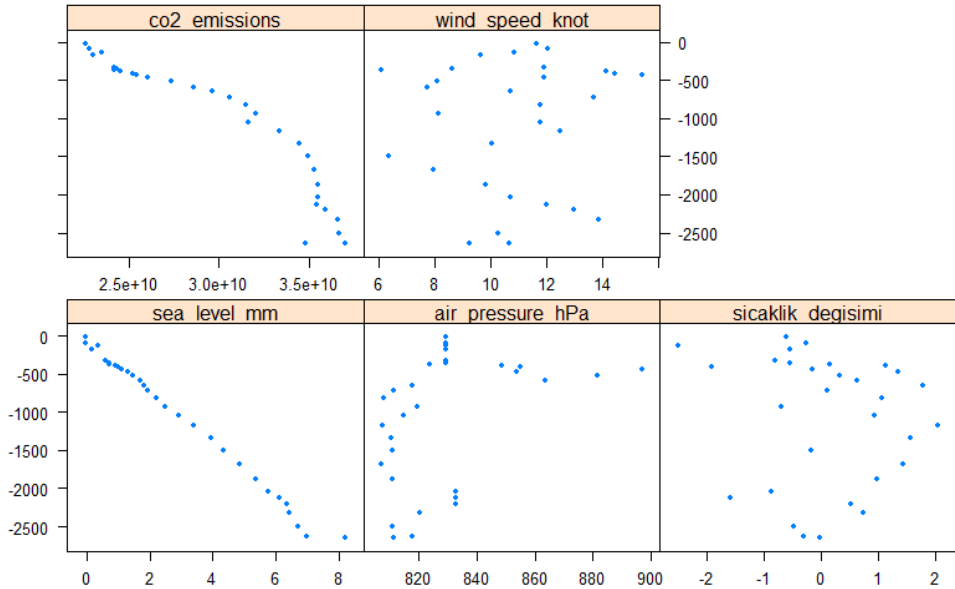
Araştırmamızı yapmak için veri bilimi ve yapay zekada en çok kullanılan diller arasında olan Python ve R tercih edilmiştir. Python'ı, veri bilimi ve yapay zekâ için kendisini zenginleştiren bir sürü kütüphane içermesi sebebiyle araştırmamızda kullanmaya uygun gördük. R dili ise ortaya çıkış nedeni istatistiksel hesaplama, analiz ve veri bilimi için özel olarak çıkarılmış bir dildir. KDnuggets'ta (veri analizi, veri madenciliği ve haberler gibi birçok veri bilimi alanını işleyen bir web sitesi) yapılan bir araştırmaya göre 2018-2019 yıllarında veri biliminde en çok tercih edilen diller yayınlanmıştır.

Daha iyi sonuçlar alınması için ise iki dilin de güçlü yanlarından faydalanılmıştır. Araştırmamız için kullanılan kütüphaneler Tablo 1’deki gibidir.

R	Python
e1071 rpart pgmm caret tidyverse AppliedPredictiveModeling	numpy pandas pytorch
dslabs rpart.plot partykit ipred broom FNN glmnet MASS ISLR	scikit-learn darts
PerformanceAnalytics funModeling Matrix kernlab randomForest	matplotlib tensorflow
gbm nnet neuralnet GGally prophet pls elasticnet	keras theano scipy
NeuralNetTools	jupyter ipython dask

Tablo 1. Araştırmamızda kullandığımız kütüphaneler

Araştırmamızda en doğru ve hata paysız sonuca ulaşabilmek için regresyon kullanılmıştır. Regresyon, günümüzde veri analizinin en önemli türlerinden biridir. Özellikle finans ve ekonomi alanında çok kullanılan regresyonlar; bağımlı değişkenlerin bağımsız değişkenlerdeki değişikliklere göre, bağımsız değişkenin değer değişimini inceleyerek bağımlı değişkenler için bir katsayı hazırlar. Bu katsayılar, gelecekteki bağımlı değişkenleri ilişkilendirerek bağımsız değişkenin yaklaşık değerini sunmaktadır. Regresyonun neredeyse kusursuz bir sonuç verebilmesi için güçlü ve doğru bir veri setine ihtiyaç duymaktadır. Bu veri setlerini Antarktika’da bulunan 18 gözlem evinden ve British Antarctic Survey ve UK Polar Data Centre’den alınmıştır. Yaptığımız literatür taramasında buzul erimesinin sıcaklık(°C), karbon emiysonu, yağış, deniz seviyesi, hava basıncı(hPa), ve rüzgâr hızına(knot) bağlı olduğu sonucu alınmıştır. Veri seti, tablo olarak gösterimi fazlaca yer kaplayacağından görsel olarak Şekil 1’de gösterilmiştir.



Şekil 1. Veri Seti

Verileri ortak bir yıl itibariyle almak adına 1992 – 2021 aralığı alınmıştır. Alınan verilerin ilerideki 10 yılı tahmin edebilmesi TimeSeries Forecasting yöntemi uygulanmıştır. Bunun için Facebook veri analistleri tarafından geliştirilen “Prophet” prosedürü kullanılmıştır. Prophet 2017 yılında Facebook tarafından geliştirilmiş Python ve R dillerinde kullanılabilen açık kaynaklı bir öngörü kütüphanesidir (Lewis, 2021). Şirket içi reklam, ziyaretçi sayısı gibi veriler kullanılarak bu verilerin gelecekteki değerlerini tahmin etmek için yazılmıştır. Bu prosedür; yıllık, aylık ya da günlük şekilde ölçülen ve doğrusal olmayan veri setleri için çıkarılmıştır. Aynı zamanda bu prosedürün en iyi şekilde sonuç verebilmesi için mevsimlik veri gerekmektedir. Elimizdeki veri setleri tek tek “Prophet” prosedürüne tabii tutulmuştur.

Bu değerlerin buzul erimesini ne kadar etki ettiğini bulmak için doğrusal olan ve doğrusal olmayan regresyonlar kullanılmıştır. Verilen sonuçlar sonunda hata paylarına bakarak test verileri karşılaştırılacak ve en iyi modele karar verilecektir. Kullanılan regresyonlar aşağıda gösterilmiştir. Aynı zamanda RMSE, R2 ve MAE formülü verilmiştir. RMSE (Kök Ortalama Kare Hatası) değeri tahmin hatalarının standart sapmasıdır. Spesifik olarak tahmin aralığındaki gerçek değerlerin yoğunluğu da denebilir. 0 ile ∞ arasında bir değer alır ve 0 değerine yaklaşması hatanın azaldığı anlamına gelmektedir. R2 (R-Kare) çıkan bağımlı değişkenin bağımsız değişkenlerle ne kadar alakalı ve orantılı olduğunu belirlemektedir. 0 ile 1 arasında değer alır. Çıkan değeri 100 ile çarpıldığında ise yüzdelik dilim olarak bulunabilir. Örnek olarak 0.6 R-Kare değeri verilmiş bir regresyonun bağımlı değişkeninin bağımsız değişkenlerle ilgisi ve kullanılabilirliği %60’tır. MAE (Ortalama Mutlak Hatası) gerçek değer ile veriye en iyi uyan tahminin arasındaki dikey veya yatay mesafedir. 0 ile ∞ arasında değer almaktadır. Düşük MAE değerli bir eğitilmiş model daha iyi performans göstermektedir. Bahsettiğimiz formüllerin matematiksel açıklamaları Şekil 2’de verilmiştir.

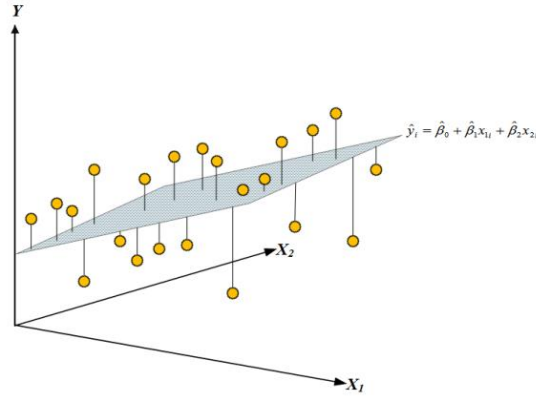
$$RMSE = \sqrt{MSE} \quad RMSE = \sqrt{\frac{\sum_{j=1}^n e_j^2}{n}} \quad MSE = \frac{1}{n} \sum_{j=1}^n e_j^2 \quad R^2 = 1 - \frac{RSS}{TSS}$$

Şekil 2. Hata payı ve güvenilirlik testi için kullanılan formüller

Cross-Validation: İstatiksel bir yeniden örnekleme yöntemidir. Modeller eğitilirken manuel olarak girilen parametreleri otonom bir şekilde mümkün olduğunca objektif, güvenilirliği yüksek ve doğru bir şekilde hesaplandırılmasında yardımcı olmaktadır. Araştırmamızda alpha, lambda gibi değerler için kullanılmıştır. Formülü aşağıdaki şekilde verilmiştir.

Basit Doğrusal Regresyon: Tek bir bağımsız değişkenin bağımlı değişkenle arasındaki ilişkiyi bulmak amaçlanmıştır.

Çoklu Doğrusal Regresyon: Birden çok bağımsız değişkenin bağımlı değişkenle arasındaki ilişkiyi bulmak amaçlanmıştır. Bunun için her bir değere kat sayı atar. Çoklu doğrusal regresyonun uygulanabilmesi için ise bağımsız değişkenler arasında multicollinearity (çoklu bağımlılık) bulunmalıdır. Örnek bir çıktı Şekil 3’te verilmiştir. Çoklu doğrusal regresyon yönteminde kullandığımız bağımsız değişkenler Şekil 4’te verilmiştir.

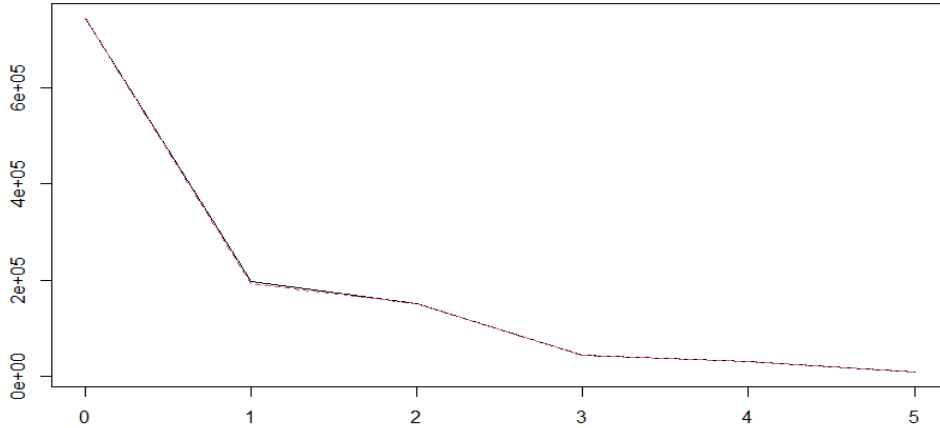


Şekil 3. Çoklu Doğrusal Regresyon Örneği

Coefficients:				
(Intercept)	sea_level_mm	air_pressure_hPa	sicaklik_degisimi	
-5.608e+02	-3.454e+02	4.492e-01	6.831e+00	
co2_emissions	wind_speed_knot			
5.832e-09	-2.732e+00			

Şekil 4. Çoklu Doğrusal Regresyon Örneği

PCR: Bağımsız değişkenlerin N adet olduğunu varsayarsak, içerdiği bilgiden daha azını maksimum şekilde kullanarak veri kaybına sebep olmaktadır. Böylelikle bağımsız değişkenler arasındaki birbirleriyle korelasyonlarını ortadan kaldırmayı amaçlamıştır. Veri tahmini yaparken benzer tahmin yapma ve hata olasılığını en aza indirir. Çalışmamızda N değerinin katsayısı sıfıra çok yakın olmuştur. Bu süreç Şekil 5'te verilmiştir.



Şekil 5 – PCR ile Hata Oranlarının Azalması

PLS: Model oluşturulurken gizli bağımsız değişkenler optimum şekilde kullanılır. Bağımsız değişkenler bağımlı değişkenlerin arasındaki kovaryansı en yüksek değeriyle döndürür. Gizli değişkenler, bağımlı değişkenler bir değişken bloğu olarak alınarak birbirine dik, karşılıklı bağımsız ve orijinal değişkenlerin doğrusal kombinasyonları olacak şekilde gizli değişkenler elde edilir. (Ekonometri ve İstatistik Sayı:15 2011 38-52) Çalışmamızda hesaplanan optimum değeri Şekil 6'da ve optimum değer süreci Şekil 7'de verilmiştir.

```

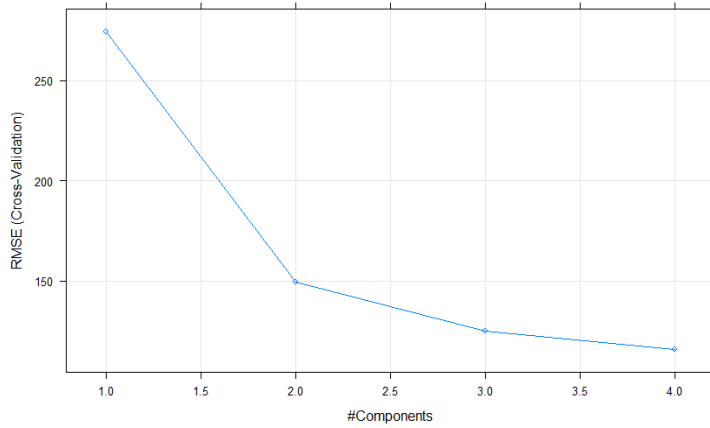
Pre-processing: centered (5), scaled (5)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 24, 24, 23, 23, 24, 24, ...
Resampling results across tuning parameters:

ncomp  RMSE      Rsquared  MAE
1      274.4041  0.9520497 244.6897
2      149.6399  0.9852918 136.5130
3      124.9835  0.9877248 109.8054
4      115.7717  0.9956513 105.7702

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was ncomp = 4.

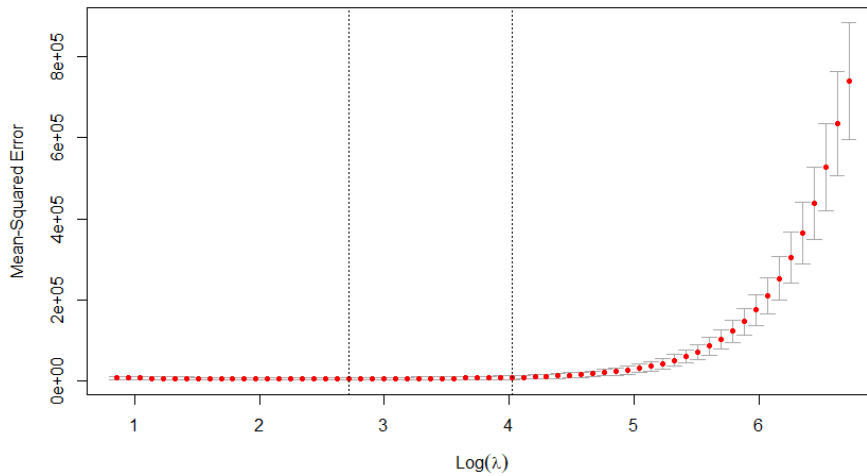
```

Şekil 6 – Optimum Değer Seçimi



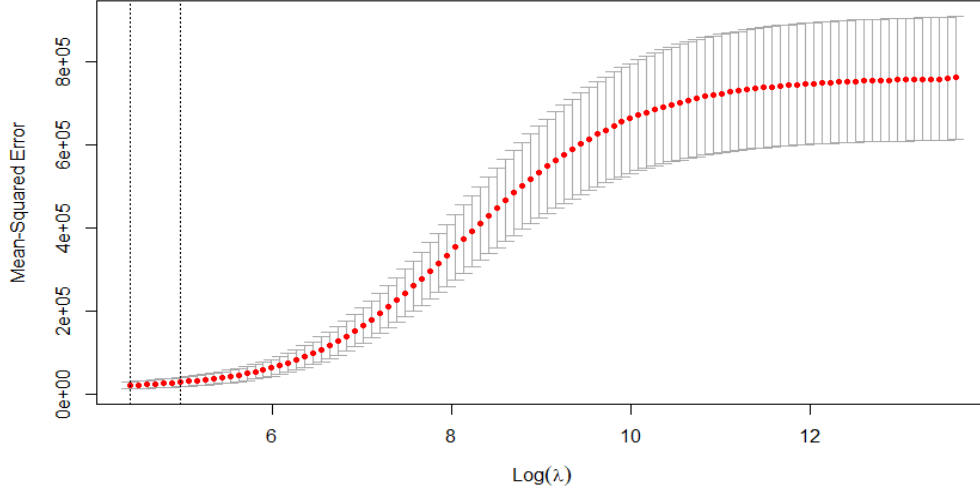
Şekil 7 – Optimum Değer Grafiği

Lasso: Hata kareler toplamını minimize eden katsayıları, bunlara ceza uygulayarak bulmaktadır. Bunun sebebiyle modelin aşırı öğrenmesi ve ezberlemesi durumuna karşı koymaktadır. Hiçbir kat sayı 0'a eşitlenmemektedir. Bağımlı değişkene etkisi olmasa bile minimal bir kat sayı değere atanır. Modelin doğruluk oranını yükseltmek için ceza sayısının iyi seçilmesi gerekmektedir. Bunun için Cross-Validation kullanılmıştır. Çalışmamızda elde edilen grafik Şekil 8'de verilmiştir.



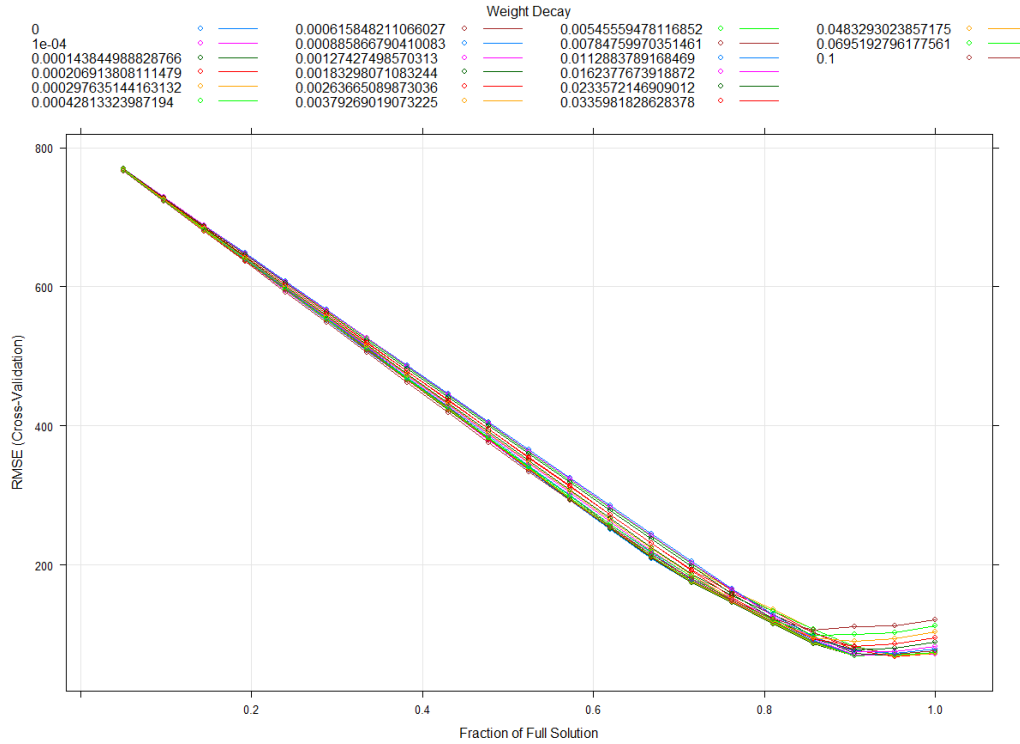
Şekil 8 – Modelin Eğitimi Sırasında Hata Kareler Toplamı ve Ceza Grafiği

Ridge: Genellikle bağımsız değişken miktarı fazla olan verileri analiz etmekte kullanılır. Hata kareler toplamını minimize eden katsayıları, bunlara ceza uygulayarak bulmaktadır. Bunun sebebiyle modelin aşırı öğrenmesi ve ezberlemesi durumuna karşı koymaktadır. Modelin doğruluk oranını yükseltmek için ceza sayısının iyi seçilmesi gerekmektedir. Bunun için Cross-Validation kullanılmıştır. Çalışmamızda elde ettiğimiz hata kareler toplamı Şekil 9’da verilmiştir. Özel olarak bu modelin kodları sayfa 10’da Şekil 13, 14, 15’te verilmiştir.



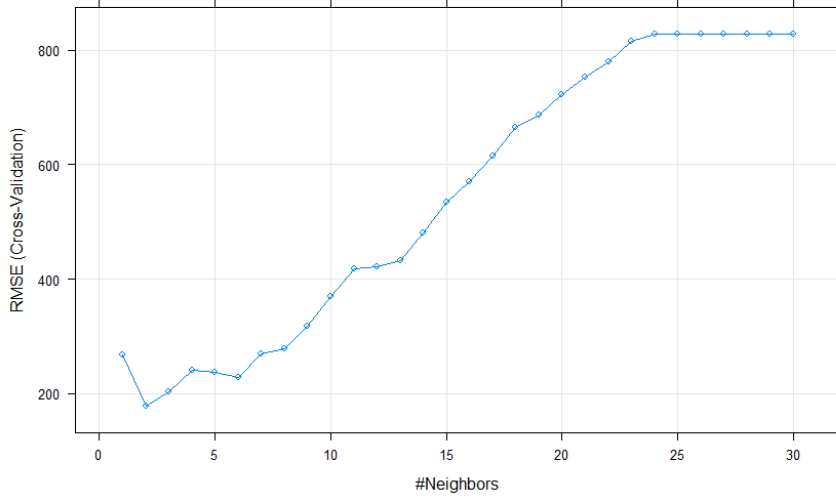
Şekil 9. Modelin Eğitimi Sırasında Hata Kareler Toplamı ve Ceza Grafiği

ElasticNet: Ridge ile Lasso regresyonlarının birleşimidir. Cross-Validation ile L2 düzenlenmesi uygulanmıştır ve maksimum doğruluk düzeyine çıkarılmıştır. Çalışmamızda elde ettiğimiz sonuçlar Şekil 10’da verilmiştir.



Şekil 10. Cross-Validation ile L2 Düzenlenmesi Uygulanması (ElasticNet)

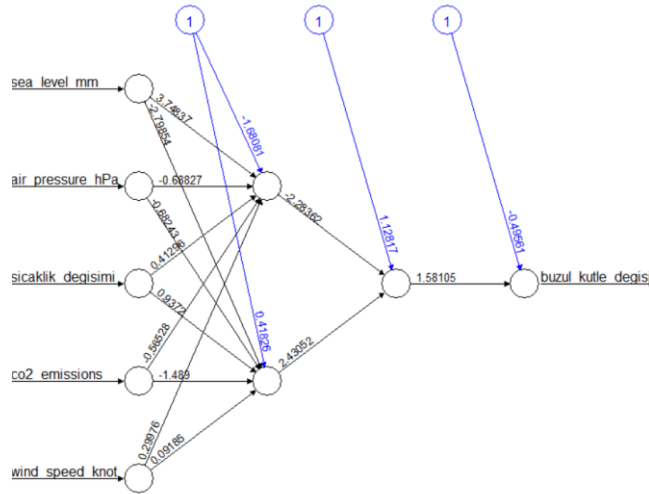
KNN: Sınıflandırılacak olan veri, daha önceki verilerle olan yakınlık ilişkisine göre sınıflandırır. KNN genellikle seçim yapma algoritmalarında kullanılmaktadır. Yakınlık değeri manuel olarak seçilmektedir. Bu araştırma için her ne kadar uygun bir regresyon olmasa da yine de Cross-Validation uygulanarak test edilmiştir. Çalışmamızda elde ettiğimiz sonuçlar Şekil 11’de verilmiştir.



Şekil 11. Cross-Validation ile L2 Düzenleştirilmesi Uygulanması (KNN)

SVR: Diğer modellerin aksine, SVR gerçek ile tahmin edilen değer arasındaki hatayı en minimum değere indirmek yerine en iyi tahmin çizgisini bir eşik değerine sığdırmaya çalışır. Eşik değeri hiper düzlem ile sınır çizgi arasındaki mesafedir. Genellikle az verili veri setleri için kullanılır. Fazla veri içeren veri setlerinde başarısı düşmektedir.

YSA: Yapay sinir ağlarında veri seti için herhangi bir tahmin veya öngörü yapmaya ihtiyaç duymaz. Temel olarak yaptığı iş veri setindeki yapıyı öğrenip tahmin ettirilmesi gereken bağımlı değişkene genelleştirme yapmaktır. Bunu yapabilmek için yapay sinir ağları modelin üzerinde kendini eğiterek ağırlıklar belirlemelidir. Böylece tahminlerinin sapmasını azaltmaktadır. Nöron sayısını belirlemek için Cross-Validation kullanılmıştır. Çalışmamızda elde ettiğimiz sonuçlar Şekil 12’de verilmiştir.



Şekil 12 – Yapay Sinir Ağı Modeli

Ridge modeli kullanılması için yazılımını R dilinde yapılmıştır.

```
df1 <- kutuplar_proje
df1 <- na.omit(df1)
rownames(df1) <- c()
egitim_indeks <- createDataPartition(df1$buzul_kutle_degisimi_gt,
                                     times=1,
                                     p=0.8,
                                     list=FALSE)

egitim <- df1[egitim_indeks,]
test <- df1[-egitim_indeks,]

egitim_x <- eğitim %>% dplyr::select(-buzul_kutle_degisimi_gt)
egitim_y <- eğitim$buzul_kutle_degisimi_gt

test_x <- test %>% dplyr::select(-buzul_kutle_degisimi_gt)
test_y <- test$buzul_kutle_degisimi_gt
```

Şekil 14. Bütün veriler %80 eğitim, %20 test verisi olacaktır.

```
ridge_cv <- cv.glmnet(as.matrix(egitimx),y=egitim_y,
                    alpha=0)

log(ridge_cv$lambda.min)
log(ridge_cv$lambda.1se)

plot(ridge_cv)

defaultSummary(data.frame(obs=test_y,
                          pred= as.vector(predict(ridge_cv,as.matrix(testx))))))
```

Şekil 15. Önce alpha(ceza) manuel olarak denenmiştir.

```
kontrol <- trainControl(method="cv",number=10)
set.seed(33)

ridge_lambda <- data.frame(lambda= seq(0,0.2,length=20))
ridge_model_tuning
ridge_model_tuning <- train(egitimx,egitim_y,
                          method="ridge",
                          trControl = kontrol,
                          tuneGrid = ridge_lambda,
                          preprocess = c("center", "scale"))

ridge_model_tuning$results

defaultSummary(data.frame(obs=test_y,
                          pred= as.vector(predict(ridge_model_tuning,as.matrix(testx)
predict(ridge_model_tuning,as.matrix(egitim_x))
barplot(df1$buzul_kutle_degisimi_gt,col="red"); par(new=T); barplot( predict(ridge_mc
plot(df1$buzul_kutle_degisimi_gt,col="red",pch=21,cex=2,lwd=3); par(new=T); plot( pre
test_y
```

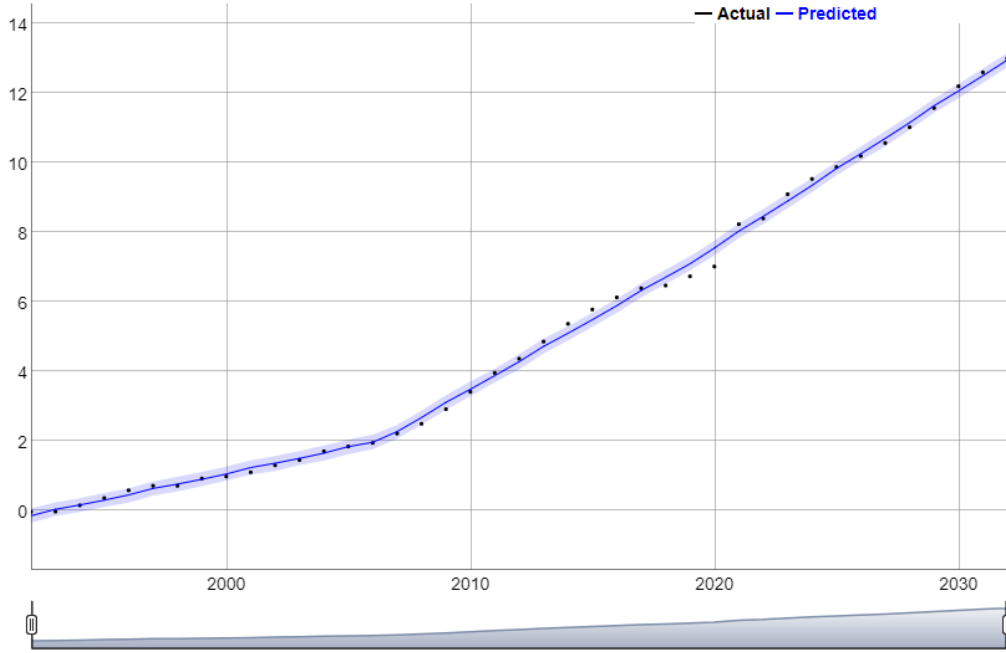
Şekil 16. Hata oranının çok yüksek olmasından dolayı Cross-Validation yapılarak en uygun alpha(ceza) değeri bulunmuştur.

Proje İş-Zaman Çizelgesi

İşin Tanımı	Ocak	Şubat	Mart	Nisan	Mayıs
Literatür Taraması	x	x			
Yazılım Dillerinin Belirlenmesi		x			
Yapay Zeka Metodu Belirlenmesi		x	x		
Veri Seti Araştırması		x	x		
Test Verilerinin Analizi			x	x	x
Proje Raporu Yazımı				x	x

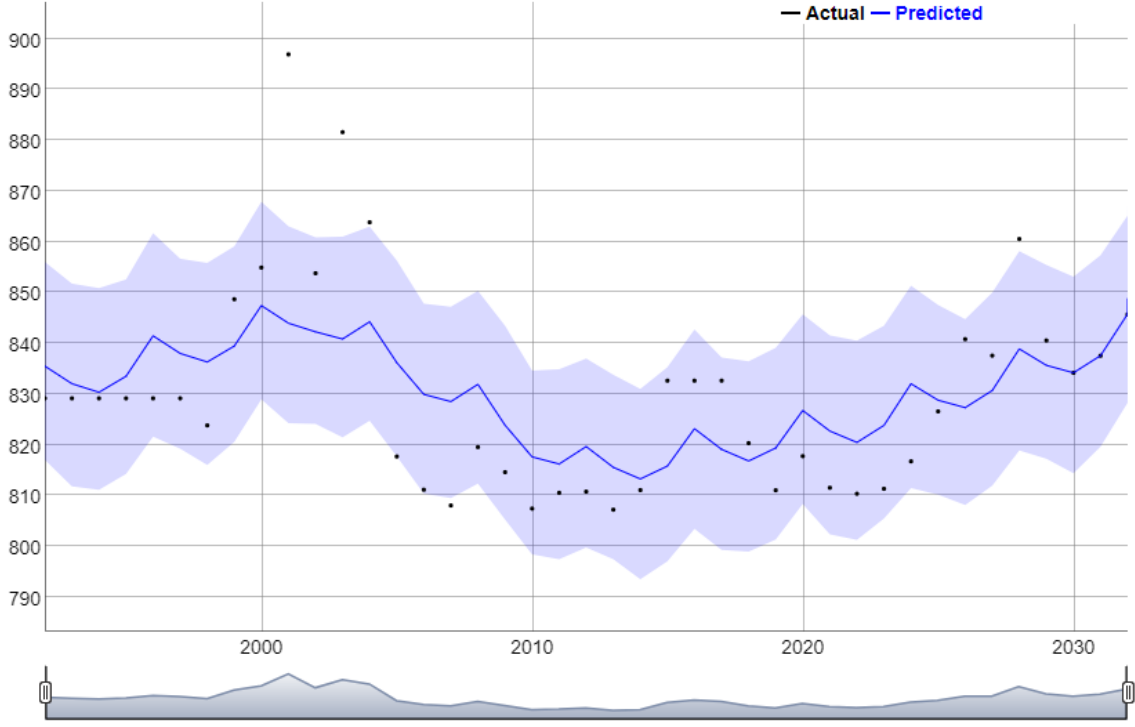
Bulgular

Su seviye değişimini gayet başarılı ve güvenilirliği yüksek bir şekilde eğitilmiş ve 2022 yılından 2032 yılına kadar tahmin etmiştir. Şekil 17’de incelenebilir.

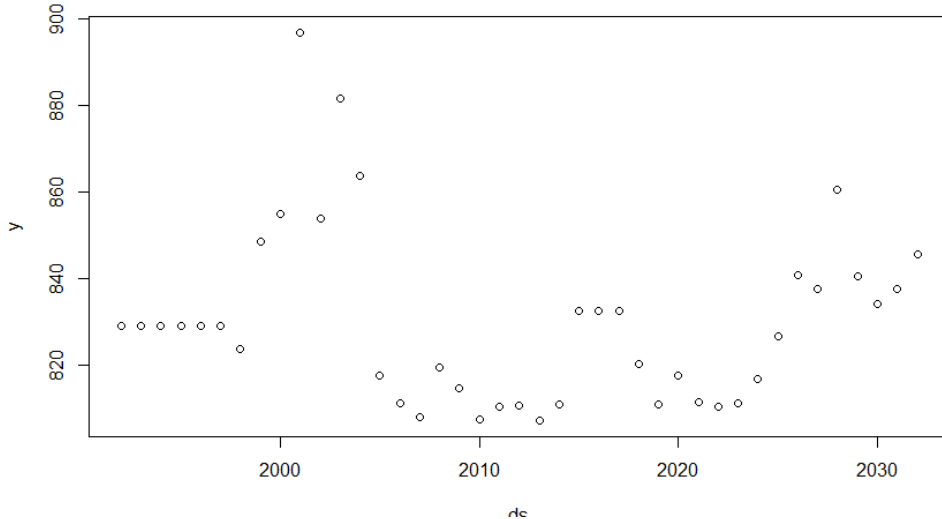


Şekil 17. Su Seviyesi Değişimi ve Tahmini

Modelin hava basıncını kesin tahmin edememesinin sebebi hava basıncının 30 yıl boyunca çok değişken olmasıdır. Tam tahmin tutturamasa da gerçek sonuçlar verdiği maksimum ve minimum değerler arasında kalmaktadır ve beklenenden daha iyi bir sonuç sergilemiştir. Bu sonuçlar Şekil 18 ve 19’da incelenebilir.



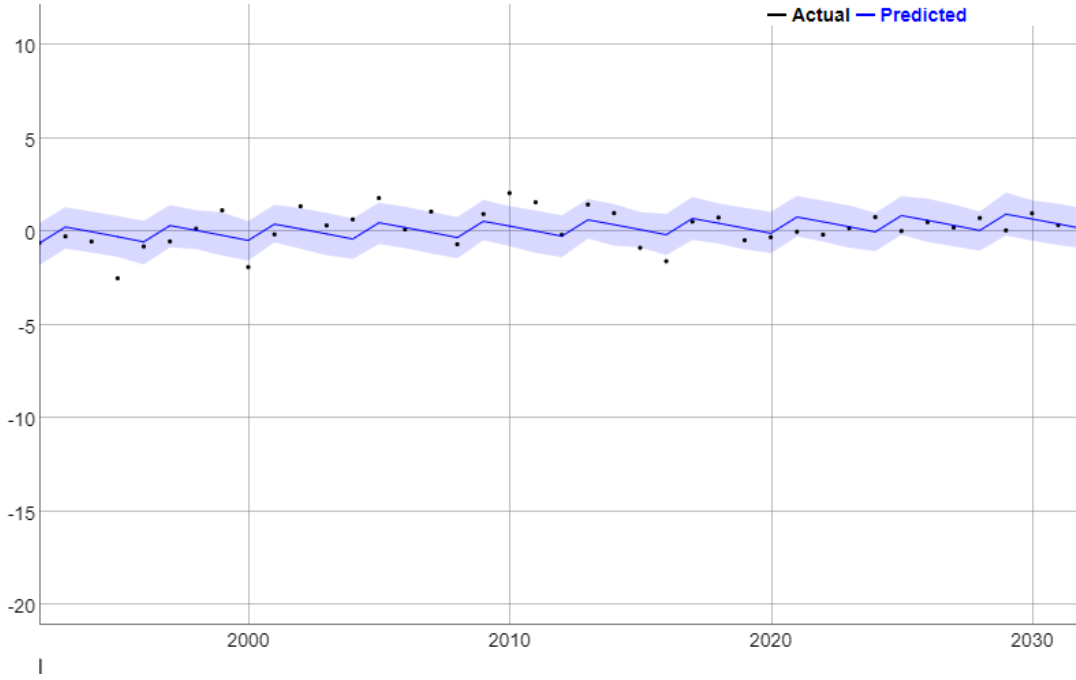
Şekil 18. Hava Basıncı (hPa)



Şekil 19. 30 yıllık hava basıncı dağılımı (hPa)

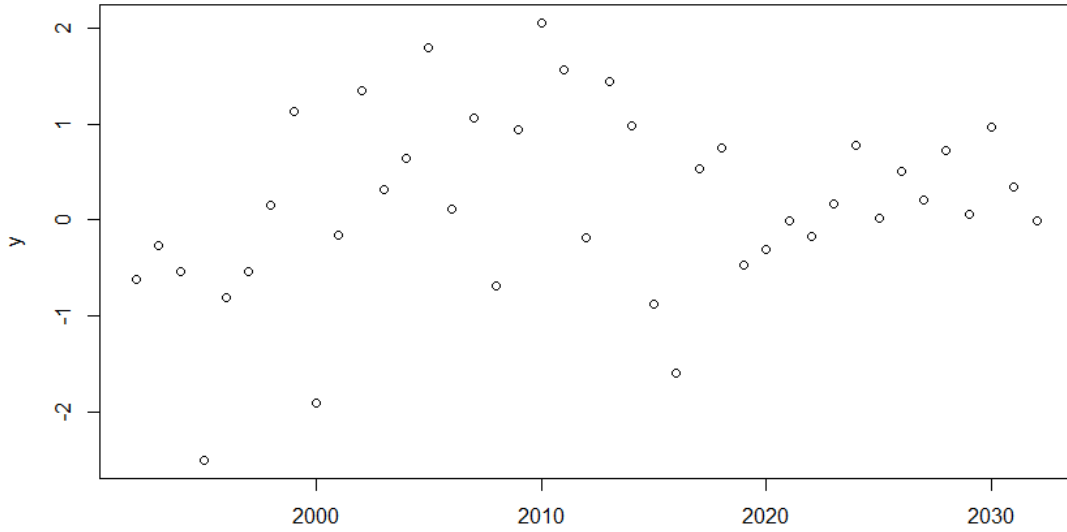
Şekil 19’da görüldüğü üzere 10 yıllık tahmin olabildiğince hata payı yüksek bir değer vermemek için kendini gerçek değerlerin arasında tutmuştur.

Sıcaklık farkı değişimi grafikte görüldüğü gibi dağınık yapıda olmasına model yakın tahminler yapmayı başarmıştır. Aynı hava basıncı gibi sıcaklık değişimi de dağınık bir yapıda olduğundan dolayı modelin su seviyesi değişimi gibi bir öngörü yapması imkansızdır. Bulgular Şekil 20’de incelenebilir.



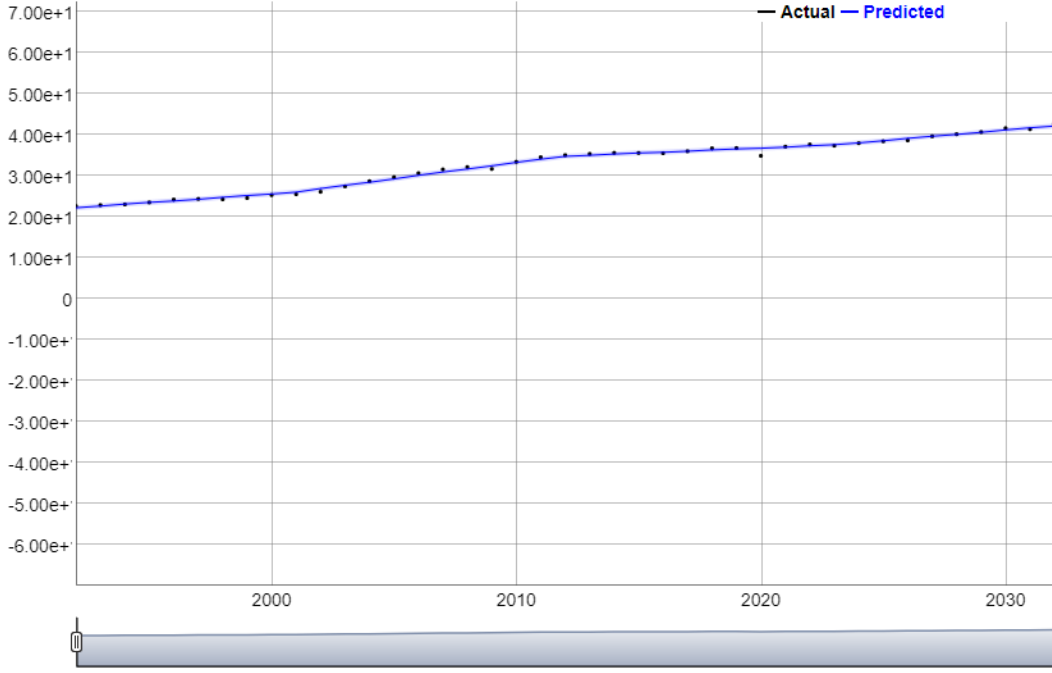
Şekil 20. Sıcaklık Farkı Değişimi

Sıcaklığın farkının 30 yıllık dağılımı yukarıdaki şekilde gibidir. Yaptığı tahminler nokta atışı olmasa da ihtiyacımızı fazlasıyla karşılamaktadır. Şekil 21’de incelenebilir.



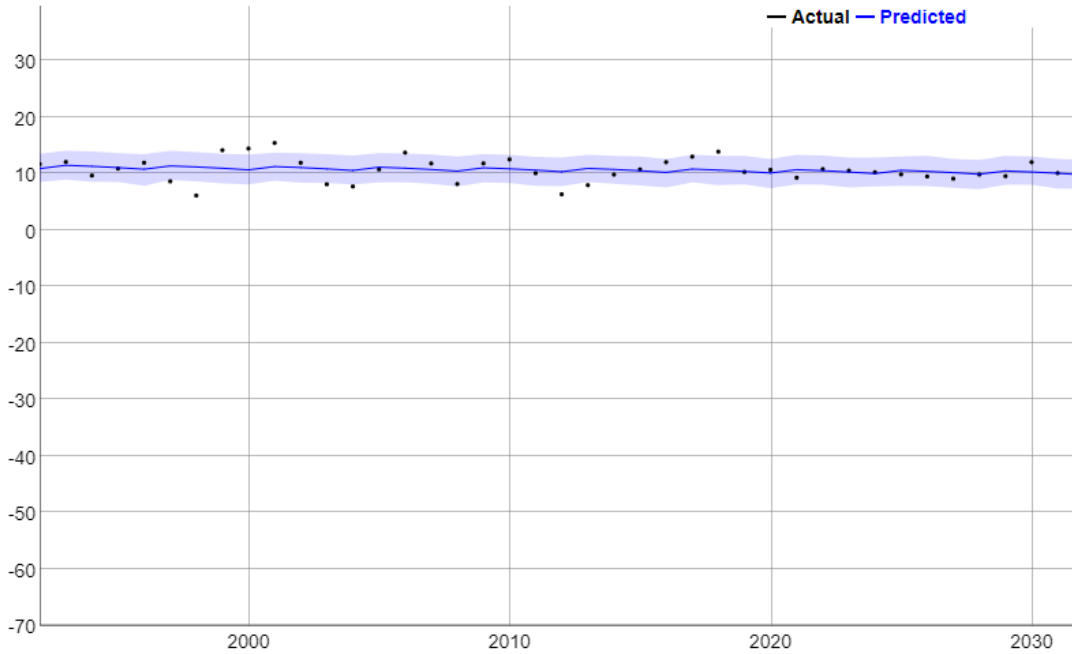
Şekil 21. 30 yıllık sıcaklık farkı dağılımı

Model CO2 emiyosonunu doğruluğu yüksek bir şekilde öğrenip 10 yıllık tahmini yapmıştır. Bu veri setlerinin 10 yıllık tahmini başarıyla yapması sonrasında sıra 10 yıl içindeki buzul erime miktarını bulmaya gelmiştir. Elde ettiğimi bulgular Şekil 22’de verilmiştir.



Şekil 22. CO2 Emisyonu

Model CO2 emiyosonunu doğruluğu yüksek bir şekilde öğrenip 10 yıllık tahmini yapmıştır. Elde edilen sonuçlar Şekil 23'te verilmiştir.



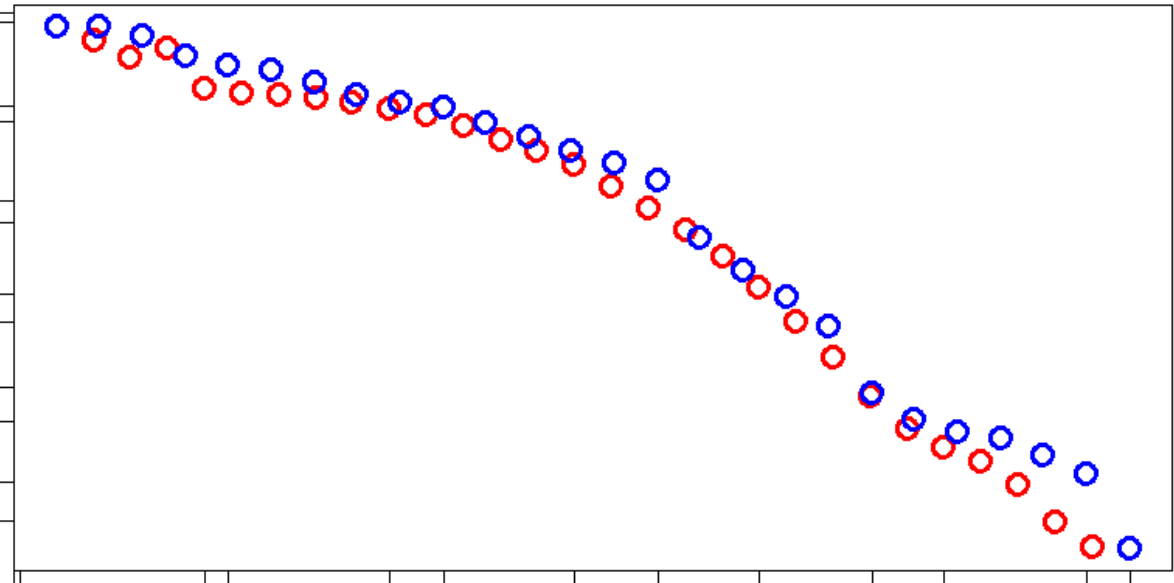
Şekil 23. Rüzgâr Hızı (knot)

Verilerin hepsi teker teker bütün yazılan regresyon modellerine girdi olarak verilmiş ve hangi regresyonun en yüksek güvenilirlik oranına sahip olduğu tespit edilmiştir.

Regresyon	Sonuç		
	RMSE	R ²	MAE
Basit Doğrusal Regresyon	67.7357963	0.9934613	43.2869870
Çoklu Doğrusal Regresyon	84.354790	0.997292	47.905593
PCR	151.0579766	0.9742097	134.0182142
PLS	259.017718	0.902621	193.314292
Lasso	105.2570219	0.9953143	70.8305751
Ridge	45.9772060	0.9952032	32.3026905
ElasticNet	96.5761475	0.9955332	63.4097305
KNN	135.9995720	0.9997713	128.2474312
SVR	95.2853200	0.9916957	78.6001072
YSA	459.0048162	0.8771925	448.8905917

Tablo 2. Sonuçların Karşılaştırılması

Tablodan çıkarılabileceği üzere en iyi sonucu veren Ridge modeli olmuştur. Model eğitilirken verilerin %80'i eğitim, %20 si test için kullanılmıştır. Sonuç olarak verileri gayet yakın bir şekilde tahmin edebilmiştir. Şekil 24'te incelenebilir. Şekil 24'te kırmızı noktalar gerçek, mavi noktalar tahmin edilen değerlerdir. Model over-fitting olmadan gerçek değerlere çok yakın değerler öngörmüştür.



Şekil 24. Tüm yıllarda öngörü noktası uzaklık farkları

```

RMSE      Rsquared      MAE
45.9772060  0.9952032  32.3026905
> predict(ridge_model_tuning,as.matrix(test_x))
[1] -281.2273 -441.5121 -1045.8352 -1866.6988
> test_y
[1] -365.4703 -405.3745 -1044.0487 -1873.7426

```

Şekil 25. Tahmin Çıktıları

Modelin bu başarılı tahminlerinden sonra 2021-2032 yıl arası tahmin edilmesi için girdi olarak regresyona verilmiştir. Tablo 3'te 1992-2032 arasında modelin ve gerçek verilerin birleşimi görüntülenebilir.

yillar	sea_level(mm)	air_pressure(hPa)	sicaklik_degisimi	co2_emissions	wind_speed(knot)	buzul_kutle_degisimi(gt)
1992	-0.04	829.148	-0.616	22569624280	11.65	-19.265175
1993	-0.03	829.148	-0.263	22803495349	12.0325	-89.98625
1994	0.15	829.148	-0.54	22964583359	9.6325	-175.0938167
1995	0.36	829.148	-2.508	23453944453	10.85083333	-128.55
1996	0.58	829.148	-0.804	24154961206	11.90083333	-330.5953333
1997	0.71	829.148	-0.535	24300550431	8.6	-354.973075
1998	0.71	823.8145455	0.147	24206271270	6.085	-365.4703417
1999	0.92	848.6818182	1.134	24518962288	14.11363636	-382.72445
2000	0.98	854.9125	-1.911	25234207250	14.415	-405.3745417
2001	1.1	896.9316667	-0.158	25451054498	15.39583333	-433.50075
2002	1.3	853.7981818	1.351	26039191477	11.89	-464.1977583
2003	1.45	881.6153846	0.32	27366723291	8.085	-516.9868167
2004	1.7	863.8466667	0.645	28627533033	7.699166667	-588.326725
2005	1.84	817.68125	1.791	29602777196	10.70833333	-648.7669667
2006	1.94	811.141	0.109	30583530081	13.69166667	-718.8013167
2007	2.21	808.0366667	1.063	31493741589	11.7725	-821.701525
2008	2.49	819.5414286	-0.686	32069444008	8.134166667	-934.4690417
2009	2.91	814.5983333	0.934	31607042410	11.78083333	-1044.048708
2010	3.41	807.42	2.056	33343300453	12.4825	-1171.273575
2011	3.95	810.524	1.564	34468451046	10.04583333	-1332.866158
2012	4.36	810.7636364	-0.182	34974075146	6.314166667	-1498.677183
2013	4.85	807.2	1.441	35283025443	7.926666667	-1679.160633
2014	5.36	811.067	0.982	35534442952	9.7975	-1873.742567
2015	5.77	832.63	-0.876	35496406339	10.70833333	-2036.857833
2016	6.12	832.63	-1.597	35452459298	12.00083333	-2130.8983
2017	6.38	832.63	0.529	35925737687	12.97583333	-2200.492225
2018	6.459346	820.3175	0.744	36646139785	13.8575	-2322.097158
2019	6.723758	811.0358333	-0.47	36702502903	10.26916667	-2503.671825
2020	7.004578	817.7483333	-0.308	34807259099	10.64	-2630.81115
2021	8.22	811.5218182	-0.01636464	37038866581	9.243636364	-2640.12
2022	8.382292	810.3443	-0.173335	37573400530	10.765378	-2870.297
2023	9.079806	811.3213	0.172698	37288176894	10.4970322	-3101.389
2024	9.517244	816.7365	0.7711729	37891878909	10.2036758	-3245.253
2025	9.858278	826.574	0.0201585	38356303784	9.830818	-3360.805
2026	10.1754	840.8097	0.5002112	38564895352	9.4684944	-3464.971
2027	10.54861	837.5626	0.2092211	39562388949	9.079856184	-3589.946
2028	11.00185	860.5476	0.72185	40078614618	9.813857	-3739.306
2029	11.55239	840.5476	0.05668164	40603201852	9.523645	-3924.32
2030	12.17782	834.1713	0.9717272	41549126548	11.991562	-4129.844
2031	12.57782	837.5404	0.34374336	41297183554	10.058979	-4264.671
2032	12.97782	845.6709	-0.0140657	42623604611	9.7752983	-4398.745

Tablo 3. 1992-2032 Yılları Arası Buzul Erime Verileri ve Tahminleri

Sonuç ve Tartışma

Yapay zekâ modelimizin 10 yıllık öngörülerine göre buzulların erime miktarı katlanarak artacaktır. 1992 yılında 19.26 gt (gigaton) eriyen buzullar, küresel ısınmanın etkisiyle beraber yıldan yıla erime miktarını katlamış ve yapay zekâ modelimizin öngörüsüne göre 2032 yılında yaklaşık olarak 43.98 gt buzulun eriyeceği tespit edilmiştir. Bunun yanı sıra buzulların erimesi nedeniyle dünya çapında su seviyeleri yükselmektedir. 1992 ile 2021 yılları arasında 8.98 cm artan deniz seviyesi, modelimizin tespitlerine göre 10 yıl içinde 11.7 cm daha yükselecektir. Bu seviyedeki artış devam ederse 2100 yılına kadar su seviyesi yaklaşık olarak 100-120 cm arasında artacak ve önemli yaşam alanları sular altında kalacaktır. Su altında kalacak yerler; Türkiye'nin Marmara kıyıları, Çukurova, Karadeniz kıyıları, Sinop kıyıları gibi birçok önemli tarım arazisini içermektedir. Karbondioksit emisyonu ise 1992 yılından itibaren 2032 yılında neredeyse 1.8 katına çıkacak ve bu da küresel ısınmanın etkisinin artmasına sebep olacaktır (Schädel ve ark., 2016). Özellikle Grönland ve Antarktika'daki buzullar tatlı sudan oluşmaktadır ve eridiklerinde ortaya çıkan tatlı suyun %69'u okyanuslara gitmektedir (BusinessInsider, 2021). Tatlı suyun okyanuslara karışması, akıntıların yavaşlaması hatta durmasına sebep olabilmektedir. Lapland Üniversitesi'nin yaptığı bir araştırmaya göre buzulların erimesini engellemek için buzul denizlerine "deniz duvarı" yerleştirilerek sıcak suyun geçişini engelleyebilir ve buz parçalarının eriyip düşmemesi için fiziksel bir destek olabileceği kanısına varılmıştır. Doğal yolla etkisini azaltmanın bir yolu ise karbondioksit emisyonumuzu azaltmamızdır. Modelimizin tahminine göre 2032 yılında 42,623,604,611 ton değerlerine kadar çıkabilecek olan karbondioksit emisyonunu azaltmanın bir yolu da topluma bilinç kazandırmaktır. Toplu taşıma kullanmak, geri dönüşüm kutularını kullanmak, enerji tasarrufu yapmak gibi karbon ayak izimizi azaltacak eylemlere başvurulmalıdır.

Çeşitli hava durumu verilerinin bulunduğu veri seti kullanılarak eğitilen yapay zekâ modeli, Antarktika'nın buz kütle değişimini yüksek doğruluk ve kullanılabilirlik oranında hesaplamıştır. Çalışmamız, bu sayede Birleşmiş Milletler Sürdürülebilir Kalkınma Hedefleri'nden İklim Eylemi'ni desteklemektedir, bilimsel çalışmalara ve toplumsal bilinci artırmaya katkı sağlamaktadır.

Öneriler

Yaptığımız çalışma Antarktika'nın tamamında meydana gelen buzul kütle değişimini göstermektedir. Antarktika'nın küçük bölgelere ayrılmış bir şekilde incelenmesi daha yararlı sonuçlar doğuracaktır. Ancak böyle bir yöntem izlenmesi için veri setinin de her bölgeye göre özelleştirilmesi gerekmektedir.

Araştırmalarımızda, herkese açık şekilde Antarktika'nın yağış verisine ulaşamadık. Eğer bu veri seti temin edilebilirse hata oranının daha da düşük olmasını bekliyoruz.

Araştırmamızın hata oranını azaltabilmek için Antarktika'dan paylaşılan verilerin daha geçmişe dayalı ve eksiksiz bir şekilde temin edilebilirse hata oranında ciddi bir düşüş yaşanacaktır.

Kaynaklar

Abram, N. J., Mulvaney, R., Wolff, E. W., Triest, J., Kipfstuhl, S., Trusel, L. D., ... & Arrowsmith, C. (2013). Acceleration of snow melt in an Antarctic Peninsula ice core during the twentieth century. *Nature Geoscience*, 6(5), 404-411.

Akın, G. (2009). *Ekoloji-çevrebilim ve çevre sorunları*. Tiydem Yayıncılık, Ankara.

Akın, G. (2013). Yüzyılımızın Temel Sorunlarından Biri; Buzulların Erimesi. *Antropoloji*, (25), 9-27.

Bulut, E., & Alma, Ö. G. (2011). Kısmi En Küçük Kareler Regresyonu Yardımıyla Optimum Bileşen Sayısını Seçmede Model Seçme Kriterlerinin Performans Karşılaştırması. *Istanbul University Econometrics and Statistics e-Journal*, 15, 38-52.

Hall, M. H., & Fagre, D. B. (2003). Modeled climate-induced glacier change in Glacier National Park, 1850–2100. *BioScience*, 53(2), 131-140.

Hopkinson, C., & Young, G. J. (1998). The effect of glacier wastage on the flow of the Bow River at Banff, Alberta, 1951–1993. *Hydrological Processes*, 12(10-11), 1745-1762.

Kadioğlu, M. (2007). *Küresel iklim değişimi ve Türkiye: bildiğiniz havaların sonu*. Güncel Yayıncılık.

Lee, J. R., Raymond, B., Bracegirdle, T. J., Chades, I., Fuller, R. A., Shaw, J. D., & Terauds, A. (2017). Climate change drives expansion of Antarctic ice-free habitat. *Nature*, 547(7661), 49-54.

Létréguilly, A. (1988). Relation between the mass balance of western Canadian mountain glaciers and meteorological data. *Journal of Glaciology*, 34(116), 11-18.

Ludescher, J., Bunde, A., Franzke, C. L., & Schellnhuber, H. J. (2016). Long-term persistence enhances uncertainty about anthropogenic warming of Antarctica. *Climate dynamics*, 46(1), 263-271.

Moore, R. D., & Demuth, M. N. (2001). Mass balance and streamflow variability at Place Glacier, Canada, in relation to recent climate fluctuations. *Hydrological Processes*, 15(18), 3473-3486.

Radić, V., & Hock, R. (2006). Modeling future glacier mass balance and volume changes using ERA-40 reanalysis and climate models: A sensitivity study at Storglaciären, Sweden. *Journal of Geophysical Research: Earth Surface*, 111(F3).

Schädel, C., Bader, M. K. F., Schuur, E. A., Biasi, C., Bracho, R., Čapek, P., ... & Wickland, K. P. (2016). Potential carbon emissions dominated by carbon dioxide from thawed permafrost soils. *Nature climate change*, 6(10), 950-953.

Shea, J. M., & Marshall, S. J. (2007). Atmospheric flow indices, regional climate, and glacier mass balance in the Canadian Rocky Mountains. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 27(2), 233-247.

Shea, J. M., Moore, R. D., & Stahl, K. (2009). Derivation of melt factors from glacier mass-balance records in western Canada. *Journal of Glaciology*, 55(189), 123-130.

Son erişim tarihi 10 Mayıs 2022: <https://www.businessinsider.com/what-if-all-ice-earth-melted-overnight-2019-10#:~:text=There%20would%20be%20mass%20flooding,would%20leak%20into%20the%20atmosphere.>

Son erişim tarihi 10 Mayıs 2022: <https://medium.com/analytics-vidhya/forecasting-using-facebooks-prophet-library-ce628e76586b>

Son erişim tarihi 10 Mayıs 2022: <https://www.tableau.com/learn/articles/time-series-forecasting>

Ekler

Projenin tüm kodları, veri setleri ve yapay zekâ modeli çıktıları ek belge şeklinde yüklenmiştir.