
GAZİ ÜNİVERSİTESİ
Gazi
1926

YAPISAL KIRILMALI ZAMAN SERİLERİNDE ALTERNATİF KESTİRİM YÖNTEMLERİ: TÜRKİYE VE ABD TÜFE ÖRNEĞİ

Mustafa Said DÜZ

YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI
VERİ BİLİMİ YÜKSEK LİSANS PROGRAMI

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TEMMUZ 2024

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Mustafa Said DÜZ
26/07/2024

YAPISAL KIRILMALI ZAMAN SERİLERİNDE ALTERNATİF KESTİRİM YÖNTEMLERİ: TÜRKİYE VE ABD TÜFE ÖRNEĞİ

(Yüksek Lisans Tezi)

Mustafa Said DÜZ

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Temmuz 2024

ÖZET

Ekonomik verilerden oluşturulan zaman serilerinin doğru kestirimi, doğru ekonomik kararlar alabilmek için kritik öneme sahiptir. Tüketici Fiyat Endeksi (TÜFE), enflasyonu ölçmek ve ekonomi politikalarını geliştirmek için hayati bir göstergedir. Geleneksel ARIMA modeli, zaman serisi tahminlerinde yaygın olarak kullanılır, ancak durağanlık gereksinimi ve yapısal kırılmalar karşısındaki sınırlamaları nedeniyle her zaman yeterli olamayabilir. Bu nedenle LSTM, XGBoost ve Prophet gibi daha yeni modeller geliştirilmiştir. Bu çalışma, Türkiye ve ABD TÜFE verilerini kullanarak ARIMA, Prophet, LSTM ve XGBoost modellerinin tahmin performanslarını karşılaştırmayı ve en doğru tahmin modelini belirlemeyi amaçlamaktadır. Farklı yapısal kırılmalara sahip zaman serileri üzerinde yapılan bu karşılaştırma, hangi modelin daha etkili olduğunu ortaya koyarak ekonomik kararların doğruluğunu artırmada önemli katkılar sağlayabilir. Sonuçlar, büyük şirketlerden kamu otoritelerine, çeşitli ekonomik aktörlerin bilinçli kararlar almasına yardımcı olacaktır.

Bilim Kodu	:	20506
Anahtar Kelimeler	:	LSTM, XGBoost, ARIMA, Prophet, makine öğrenmesi, zaman serileri, öngörü, tahmin
Sayfa Adedi	:	53
Danışman	:	Prof. Dr. Meral EBEGİL

ALTERNATIVE FORECASTING METHODS IN TIME SERIES WITH STRUCTURAL BREAK: AN EXAMPLE OF TURKEYAND USA CPI

(M. Sc. Thesis)

Mustafa Said DÜZ

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

July 2024

ABSTRACT

Accurate forecasting of time series created from economic data is crucial for making informed economic decisions. The Consumer Price Index (CPI) is vital for measuring inflation and developing economy policies. The traditional ARIMA model is widely used in time series forecasting but may not always be sufficient due to its requirement for stationarity and limitations in handling structural breaks. Therefore, newer models such as LSTM, XGBoost and Prophet have been developed. This study aims to compare the forecasting performance of ARIMA, Prophet, LSTM, and XGBoost models using CPI data from Turkey and the USA to determine the most accurate forecasting model. By comparing these models on time series with different structural breaks, this study can identify the most effective model, thereby significantly contributing to the accuracy of economic decisions. The results will assist various economic actors, from large corporations to government authorities, in making informed decisions.

Science Code	:	20506
Key Words Number	:	LSTM, XGBoost, ARIMA, Prophet, machine learning, time series, prediction, forecasting
Page Number	:	53
Supervisor	:	Asst. Prof. Dr. Meral EBEGİL

TEŞEKKÜR

Gazi Üniversitesi Fen Fakültesi İstatistik Ana Bilim Dalında yapmakta olan Tez Danışmanım Profesör Doktor Meral EBEGİL'e oldukça zorlandığım tez konusu seçiminden, yayın konusuna kadar özverili desteklerini esirgemediği için; yoğun tez dönemimde yanımda olan ve kendilerine ayırdığım vakti böldüğüm eşim Nilüfer ve kızım Lilya'ya fedakârlıkları için; değerli vakitlerini ayırarak tez jürimde yer alan ve değerli katkılarıyla tezimin kalitesinin artmasına yardımcı olan Prof. Dr. Yaprak Arzu ÖZDEMİR ve Dr. Öğr. Üyesi Şenay ÖZDEMİR'e teşekkürü bir borç bilirim.

İÇİNDEKİLER

Sayfa

ÖZET iv

ABSTRACT v

TEŞEKKÜR vi

İÇİNDEKİLER vii

ÇİZELGELERİN LİSTESİ ix

ŞEKİLLERİN LİSTESİ x

SİMGELER VE KISALTMALAR xi

1. GİRİŞ 1

2. MATERYALLER VE METOTLAR 11

2.1. Yapısal Kırılma Tespit Testleri 11

2.1.1. Chow testi 11

2.1.2. CUSUM testi (birikimli toplam kontrol grafiği) 12

2.2. Zaman Serisi Tahmin Yöntemlerinin İncelenmesi 13

2.2.1. ARIMA yöntemi 13

2.2.2. Prophet yöntemi 14

2.2.3. XGBoost yöntemi 15

2.2.4. LSTM yöntemi 17

2.3. Parametre Optimizasyonu 17

3. UYGULAMA 19

3.1. Veri Toplama 19

3.2. Yapısal Kırılma Testleri 21

3.2.1. Türkiye ve ABD TÜFE zaman serisine Chow testinin uygulanması 21

3.2.2. Kodlama ortamının seçimi ve hazırlanması 21

3.2.3. Türkiye TÜFE zaman serisine Chow testinin uygulanması 21

Sayfa

3.2.4. ABD TÜFE Chow testi uygulanması 23
3.2.5. Türkiye ve ABD TÜFE zaman serisine CUSUM testinin uygulanması ... 24
3.2.6. Kodlama ortamının seçimi ve hazırlanması 24
3.2.7. Türkiye TÜFE zaman serisine CUSUM testinin uygulanması 24
3.2.8. ABD TÜFE zaman serisine CUSUM testinin uygulanması 26
3.3. Kestirim Yöntemleri 27
3.3.1. Türkiye TÜFE tahmin sonuçları 27
3.3.2. ABD TÜFE Tahmin Sonuçları 33
3.4. Yöntemlerin Karşılaştırılması 39
4. SONUÇ VE ÖNERİLER 43
KAYNAKLAR 47
ÖZGEÇMİŞ 53

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

Çizelge 3.1. Türkiye Tüfe tahmin yöntemlerinin sonuçları 40

Çizelge 3.2. ABD Tüfe tahmin yöntemlerinin sonuçları 40

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

Şekil 3.1. Türkiye TÜFE yıllık değişim ilk ve son 5 ay 19

Şekil 3.2. ABD TÜFE yıllık değişim ilk ve son 5 ay 20

Şekil 3.3. Türkiye TÜFE zaman serisi gerçekleşme grafiği 20

Şekil 3.4. ABD TÜFE zaman serisi gerçekleşme grafiği 21

Şekil 3.5. Chow Testi ile Türkiye TÜFE zaman serisinin yapısal kırılma noktası tespiti 22

Şekil 3.6. Türkiye TÜFE Chow testi sonucu 23

Şekil 3.7. ABD TÜFE Chow testi sonucu 23

Şekil 3.8. Chow Testi ile ABD TÜFE zaman serisinin yapısal kırılma noktası tespiti 24

Şekil 3.9. Türkiye TÜFE CUSUM testi grafiği 25

Şekil 3.10. Türkiye TÜFE CUSUM testi sonucu 25

Şekil 3.11. ABD TÜFE CUSUM testi grafiği 26

Şekil 3.12. ABD TÜFE CUSUM testi sonucu 26

Şekil 3.13. Türkiye TÜFE verisi için ARIMA ile tahmin ve gerçekleşme 28

Şekil 3.14. Türkiye TÜFE verisi için Prophet ile tahmin ve gerçekleşme 29

Şekil 3.15. Türkiye TÜFE verisi için XGBoost ile tahmin ve gerçekleşme 31

Şekil 3.16. Türkiye TÜFE verisi için LSTM ile tahmin ve gerçekleşme 32

Şekil 3.17. ABD TÜFE verisi için ARIMA ile tahmin ve gerçekleşme 33

Şekil 3.18 ABD TÜFE verisi için Prophet ile tahmin ve gerçekleşme 35

Şekil 3.19. ABD TÜFE verisi için XGBoost ile tahmin ve gerçekleşme 37

Şekil 3.20. ABD TÜFE verisi için LSTM ile tahmin ve gerçekleşme 39

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	Açıklamalar
$\mathbf{RSS_c}$	Kırılma noktasında bölünmüş iki model için toplam karelerin toplamı
$\mathbf{RSS_u}$	Kırılma noktasız tek model için hesaplanan toplam kareler toplamı
$\mathbf{S_t}$	t zaman değerli zaman serisi
$\mathbf{e_i}$	Regresyon artıkları
$\mathbf{h_i}$	Karar ağaçlarında zayıf öğrenici
$\mathbf{h_t}(\boldsymbol{x})$	İterasyon t'deki zayıf öğrenici
$\boldsymbol{F_{t-1}}(\boldsymbol{x})$	İterasyon t-1'e kadar güçlendirilmiş model
$\boldsymbol{e_t}$	Hata terimi
$\boldsymbol{e_t}$	Hata terimi
$\boldsymbol{y_t}$	Tahmin edilen değer
C(t)	Hücre durumu (Cell state)
d	ARIMA'da durağanlaştırma için gereken fark
eta	Öğrenme oranı
F(x)	Son güçlendirilmiş model
g(t)	Hücre durumu için aday değer hesaplayıcı
g(t)	Prophet fonksiyonundaki trend bileşeni
h(t)	Gizli durum (Hidden state)
h(t)	Prophet fonksiyonundaki tatil ve özel olaylar bileşeni
i	Yineleme sayısı
i(t), f(t), o(t)	Bilgi akışını kontrol eden kapı değerleri
N	Gözlem sayısı
p	Modeldeki parametre sayısı
p	Otoregresiflik derecesi
q	Hareketli ortalama derecesi
$\mathbf{R^2}$	Regresyon modelinin belirleme katsayısı
s(t)	Prophet fonksiyonundaki mevsimsellik bileşeni

Simgeler	Açıklamalar
$\mathbf{y(t)}$	g(t) + s(t) + h(t) + ε(t)
$\boldsymbol{\varepsilon(t)}$	Hata terimi
$\boldsymbol{B}$	ARIMA geri kaydırma operatörü
$\boldsymbol{a}$	CUSUM testi sabit değeri
$\boldsymbol{b}$	CUSUM testi katsayı
$\boldsymbol{\Omega(h_i)}$	Düzenleme terimi
$\boldsymbol{\gamma}$	Güncelleme adım boyutunu kontrol eden öğrenme hızı
$\boldsymbol{\theta(B)}$	Hareketli ortalama polinom terimi
$\boldsymbol{\varphi}$	Otoregresif terimlerin katsayıları
$\boldsymbol{\varphi(B)}$	Otoregresif polinom terimi

Kısaltmalar	Açıklamalar
AR	Otoregresif
ARDL	Gecikmesi dağıtılmış otoregresif sınır testi
ARIMA	Bütünleşik otoregresif hareketli ortalama
CNN	Evrişimli sinir ağları
CPI	Consumer price ındex
CSV	Virgülle ayrılmış değerler
CUSUM	Birikimli toplam kontrol grafikleri
EKG	Elektrokardiyografi
GARCH	Genelleştirilmiş otokorelasyonlu koşullu homoskedastiklik
GRU	Kapı özyinelemeli geçitler (Gated recurrent unit)
LSTM	Uzun kısa süreli bellek
MA	Hareketli ortalama
MAE	Ortalama mutlak hata
MAPE	Ortalama mutlak yüzde hata
MARS	Çok değişkenli uyarlanabilir regresyon eğrileri
ME	Ortalama hata
MLP	Çok katmanlı algılayıcı
MLR	Çok değişkenli lineer regresyon

Kısaltmalar	Açıklamalar
MPE	Ortalama yüzdesel hata
MSE	Ortalama kare hata
RMSE	Kök ortalama kare hata
RNN	Tekrarlayan sinir ağları
SARIMA	Mevsimsel bütünleşik otoregresif hareketli ortalama
sMAPE	Simetrik ortalama mutlak yüzde hata
SVR	Destek vektör regresyonu
TAR	Eşik korelasyon
TÜİK	Türkiye İstatistik Kurumu
XGBoost	Ekstrem gradyan artırma

1. GİRİŞ

Küresel Covid19 salgını sonrası ekonomide yaşanan değişiklikler, fiyat endeksleri ile ilgili tahmin yapmayı zorlaştırmış, mevcut tahmin yöntemlerinin etkinliğini etkilemiştir. Özellikle Covid19 salgını ile küresel çapta başlayan istisnai yüksek enflasyonist ortam sonucunda, gerçekleşmesi uzun sürecek olan projeler ile ilgili öngörülen ve imza altına alınan sözleşme bedellerinin yüklenicilerin enflasyon sonucu artan masraflarını karşılamadığı ve güncellenmesi gerektiği gerekçesiyle, proje sahipleriyle yükleniciler arasında hukuki ihtilaflar doğmaktadır. Dolayısıyla, fiyat endekslerindeki ani artışların öngörülebilir olup olmadığının tespiti, bu ihtilaflar sonucu yapılan hukuki başvuruların sonuçlandırılmasında büyük önem arz etmektedir. Fiyat endekslerini tahmin etmek çok farklı etmenlere bağlıdır. Bu noktada, fiyat endekslerindeki öngörülemez büyük artışların yapısal kırılma olduğu göz önüne alınarak, yapısal kırılmalı zaman serilerinin öngörülebilir olup olmadığının istatistiksel ve makine öğrenmesi yöntemleriyle tespiti, ihtilafların çözülmesinde ciddi rol oynayabilecektir. Bu amaçla da, söz konusu zaman serisindeki yapısal kırılmayı hesaba katarak, en iyi kestirimi yapmayı sağlayan istatistiksel ya da makine öğrenmesi yönteminin tespiti de kilit bir öneme sahiptir.

Bu çalışmanın temel amacı, yapısal kırılmalı zaman serilerinde çeşitli yöntemlerin -ARIMA, Prophet, XGBoost, LSTM- kestirimdeki etkinliklerinin değerlendirilmesi ve kıyaslanmasıdır. Bu çalışma spesifik olarak aşağıdaki soruları cevaplamayı amaçlamaktadır:

- Bu modeller yapısal kırılmaların varlığı durumunda nasıl performans göstermektedir?
- Her bir modelin karmaşık zaman serilerinde güçlü yanları ve kısıtları nelerdir?

Bu çalışmanın, yapısal kırılma durumlarında bu modellerin performanslarını kıyaslayarak literatürdeki önemli bir boşluğu doldurması hedeflenmektedir.

Çalışmanın temelini oluşturan zaman serisi analizi ise; kronolojik sırayla düzenlenen verilerden anlamlı bir özet ve istatistiki bilgi çıkarma çabasıdır (Nielsen, 2019). Zaman serisi analizi, ekonomi, finans, meteoroloji, astronomi ve sağlık bilimleri gibi çeşitli alanlarda, karmaşık veri setlerinden geleceğe yönelik kestirimler yapmak için kullanılır (Brockwell ve Davis, 2002). Bu ve benzeri alanlarda veri noktalarının zaman içindeki değişimini ve trendlerini inceleyerek geleceğe yönelik kestirimler gerçekleştirilir (Chatfield, 2003).

Zaman serisi analizi 1970'te ortaya çıkan Box-Jenkins yöntemi ile sistematikleşmiştir (G. E. P. Box ve Jenkins, 1970). Bununla birlikte, o dönemde şimdiki gibi çok büyük veri setlerini işleyebilecek güçte işlemcili bilgisayarlar yoktu. Bundan dolayı zaman serisi analizi küçük veri kümeleri ile başlamış, bilgisayar teknolojilerinin gelişmesine doğru orantılı olarak zamanla gelişmiş ve şu anda çok daha büyük veri setleri ile eskiye göre çok daha hızlı bir şekilde zaman serisi analizleri gerçekleştirilebilmektedir (Nielsen, 2019).

II. Dünya Savaşı sonrası ekonomide zaman serisi analizi, özellikle veri erişiminin kolaylığı nedeniyle yaygınlaşmıştır. Örneğin, borsada işlem gören bir hissenin fiyatını, kayıtlı olmasından ve gizli olmamasından dolayı öğrenmek zor değildir. Ekonomi alanındaki veriler; tıp alanında mikroskop, EKG cihazı; astronomi alanında gelişmiş teleskoplar; meteoroloji alanında radar, uydu gibi toplanması yoğun teknoloji ve kullanıcı uzmanlığı gerektiren veriler değildir (Nielsen, 2019).

Ekonomik tahminler 19.yy sonları ve 20. yy başlarında ABD ve Avrupa'da dönemsel bankacılık krizlerinin tetiklediği endişeden doğmuştur. ABD'de bu tahminlerin olabildiğince doğru yapılabilmesi ve ekonomik aktörlerin doğru bilgilendirilebilmesi için çeşitli kamu kuruluşları ve özel kuruluşlar oluşturulmuştur. Bu ve benzeri kurumlar, bugün ülkemiz de dahil dünyanın birçok ülkesinde kullanılan ve kamunun anlayabileceği açıklıkta tablolar oluşturmuştur. Ülkemizde de bu fonksiyonu yerine getiren kamu kurum ve kuruluşları mevcuttur. Bu çalışmada üzerinde çalışılacak olan verilerden biri de bu kurumlardan biri olan Türkiye Cumhuriyet Merkez Bankasının internet sitesinden alınmıştır (Türkiye Cumhuriyet Merkez Bankası [TCMB], 2022). Türkiye İstatistik Kurumu (TÜİK), Hazine ve Maliye Bakanlığı, Kalkınma Bakanlığı, Borsa İstanbul gibi çeşitli kurum ve kuruluşlarda ekonomik verilerden oluşan zaman serileri kullanılmaktadır.

Zaman serilerinde yapısal kırılmalar, zaman içindeki verilerdeki sıra dışı değişikliklerin neden olduğu durağanlığı bozan anormalliklerdir. Bu yapısal kırılmalar, seri içinde belirli bir noktada meydana gelen, serinin ortalamasının değişmesine yol açarak, serinin seyrini etkileyen radikal bir değişime yol açabilmektedir. Bu tür büyük değişimler zaman serisi kestirim modellerinin, trendlerin, mevsimselliğin bozulmasına yol açabilmektedir. Bu da doğru kestirim yapılmasını zorlaştırmaktadır (StataCorp LLC, 2024).

Kırılmaların nedenleri, birçok faktöre bağlı olabilir. Örneğin, ekonomik verilerdeki kırılmalar, ekonomik politika değişiklikleri, savaşlar, pandemiler, doğal afetler, kıtlık ve benzeri sıra dışı olaylar nedeniyle meydana gelebilir (Bai ve Perron, 2003).

Fiyat endeksi zaman serilerinde de kırılmalar zaman zaman görülen bir durumdur. Fiyat endeksi bir ekonomik gösterge olması nedeniyle, bu zaman serilerindeki kırılmaların nedenleri, yukarıdaki paragrafta belirtilen olaylar kapsamında değerlendirilebilir. Bu kırılmalar, enflasyon oranının yapısal olarak değiştiği dönemleri ifade eder ve genellikle ekonomik politikaların veya diğer ekonomik faktörlerin etkisiyle meydana gelir. Son yıllarda Covid-19 pandemisi, Ukrayna-Rusya savaşı ve ülkemizde 2023 yılında meydana gelen Kahramanmaraş merkezli depremler gibi olaylar sonucunda dünya genelinde ve Türkiye'de normal şartlar altında öngörülemeyen büyüklükte bir fiyat enflasyonu yaşanmıştır (Türkiye İstatistik Kurumu [TÜİK], 2024).

Yapısal değişikliğin belirlenmesi, ekonomik zaman serilerinin ve panel verilerin analizinde çok önemli bir adımdır. Zaman aralığı ne kadar uzunsa, 2007-2008 mali krizi, 2020 COVID-19 salgını ve Rusya-Ukrayna savaşı gibi önemli yıkıcı olayların bir sonucu olarak model parametrelerinin değişmiş olma olasılığı o kadar yüksektir. Bu nedenle kırılmaların varlığını tespit etmek ve tarihlendirmek sadece tahmin amacıyla değil, aynı zamanda değişimin itici güçlerini ve bunların ilişkileri üzerindeki etkilerini anlamak için de gereklidir (Ditzen vd., 2021).

Zaman serisi verilerinde yapısal kırılma tespiti, istatistiksel modellerin doğruluğu ve güvenilirliği üzerinde önemli bir etkiye sahiptir. Bu alandaki tespit yöntemlerinin gelişimi, Gregory Chow'un 1960 yılındaki öncü çalışmasıyla başlamıştır. Chow Testi, regresyon modellerinde tek bir yapısal kırılmanın tespit edilmesi için yeni bir yaklaşım sunmuş ve ekonomistlerin model istikrarına yaklaşımını temelden değiştirmiştir (Chow, 1960). Aynı dönemde Quandt (1960), bilinmeyen yapısal değişiklikleri araştırmış ve hata varyansındaki değişiklikleri tartışmıştır (Quandt, 1960). Brown, Durbin ve Evans'ın 1970'lerde tanıttığı CUSUM testi ise yapısal kırılma tespitinin kapsamını genişletmiştir (Brown vd., 1975).

1980'lerde, Dufour'un çoklu rejim Chow Testini tanıtması, karmaşık veri yapılarıyla başa çıkmada önemli bir adım olmuştur (Dufour, 1982). Lo ve Newey ise eş zamanlı denklemler üzerinde çalışarak Chow Testi için ileri geliştirmeler yapmıştır (Lo ve Newey, 1985).

Andrews ve Fair, genel doğrusal olmayan modelleri varyans test analizi kullanarak incelemiş ve daha çok yönlü yöntemlere doğru bir yönelim sağlamıştır (Andrews ve Fair, 1988). Kim ve Siegmund (1989) basit doğrusal regresyon modelleri için Quandt Testini hem kesişim hem de eğim değişikliklerini içerecek şekilde geliştirmiştir (Kim ve Siegmund, 1989). Bu dönemde ayrıca, Ploeger ve diğ. 1988 ve 1989 yılları arasında gecikmeli bağımlı değişkenlerle doğrusal regresyona Quandt Testi önerilmiştir (Kramer vd,. 1988). 1990'ların başlarında, Chihwa ve Ross seri korelasyon hataları olan doğrusal regresyon sorununu ele almıştır (Chihwa ve Ross, 1995). (Cantrell vd., n.d.) ve (Dufour vd., 1994) tarafından genel doğrusal olmayan modellerle tahmin testlerinin uygulandığı önemli ilerlemeler de kaydedilmiştir.

1995 yılında Hansen, yapısal kırılmaların değerlendirilmesinde temel bir araç haline gelen detaylı p-değerleri önermiş ve literatüre önemli bir katkı yapmıştır (Hansen, 1997). Zaman serileri analizinde yapısal kırılma tespit yöntemlerinin evrimi, ekonometrik modellerde yapısal kırılmaları tespit etme ve analiz etme yeteneğini önemli ölçüde artırmıştır. Chow ve Quandt'un temel çalışmalarından 20. yüzyılın sonlarındaki sofistike metodolojilere kadar, bu gelişmeler yapısal kırılma testlerini ekonometrik analizde daha detaylı ve ayrılmaz bir parça haline gelmiştir. Alan gelişmeye devam ederken, bu yöntemlerin ekonometrik analizde daha da detaylı ve ayrılmaz bir parça hale gelmesi muhtemeldir.

Zaman serisi tahminleri, ekonomik karar vericiler için vazgeçilmezdir ve yaklaşmakta olan ekonomik gelişmelere proaktif ve önleyici yanıtlar verilmesini sağlar. Ekonomik veriler içinde, TÜFE özel bir öneme sahiptir çünkü belirli bir pazar veya ülkede nihai tüketicilerin karşılaştığı ortalama fiyat seviyelerini temsil eder. Bir endeks olarak ifade edilen TÜFE, farklı aylar ve yıllar arasında karşılaştırmalı analiz yapılmasını mümkün kılar. Tahmin doğruluğunu en üst düzeye çıkarmak için, tüm yöntemler geçmiş verilerden yararlanır ve tarihin kendisinin, bağlamdan bağımsız olarak, TÜFE üzerinde bağımsız bir etkisi olmadığını kabul eder. Bu otoregresiflik özelliği, zaman serisi tahminlerini doğal olarak geçmiş değerlere bağımlı kılar.

Fiyat endekslerinin tahmini, ekonomik karar vericiler için büyük önem taşımaktadır. Bu konuda literatürde birçok çalışma bulunmaktadır. Bunlardan bazıları şunlardır; Nguyen ve diğ. (2023) ABD TÜFE tahmini için ABD TÜFE verisi kullanarak çok değişkenli lineer regresyon (MLR), destek vektör regresyonu (SVR), gecikmesi dağıtılmış otoregresif sınır

testi (ARDL) ve çok değişkenli uyarlanabilir regresyon eğrileri (MARS) modellerini MAPE, MAE, RMSE ve R^2 ölçütlerine göre kıyaslamış, MARS yönteminin gerçeğe en yakın sonuçları verdiğini tespit etmişlerdir (Nguyen vd., 2023). Sibai vd. (2024) Suudi Arabistan'ın 2013-2020 yılları arası TÜFE verileri ile Ortalama Kare Hata (MSE) ve R^2 ölçütlerine göre Karar Ağacı, k-en yakın komşular (kNN), Lineer Regresyon (LR), Sinir Ağları (NN), Rastgele orman (RF), and Destek Vektör Makinesi (SVM) yöntemleri ile en doğru kestirimi gerçekleştiren yöntemin MLP olduğu tespit edilmiştir (Sibai vd., 2024). Yunus Emre GÜR (2024) Kapı Özyinelemeli Geçitler (GRU), LSTM, Rastgele Orman, Sinir Ağları ve k-En Yakın Komşular modellerinin TÜFE kestirim performanslarını RMSE, MSE, MAPE ve R^2 ölçütlerine göre karşılaştırmış ve GRU yönteminin en gerçeğe yakın kestirimi gerçekleştirdiği sonucuna varmıştır (Gür, 2024).

Çoğunlukla ülke bazında yürütülen bu tahmin çalışmaları, bir ülkeden diğerine fiyat endekslerindeki değişen eğilimler nedeniyle, farklı ülkeler için farklı modellerin daha doğru tahminler sağlamasına yol açabilir.

Son yıllarda, fiyat artışlarının öngörülebilirliği çeşitli sektörlerde tartışma konusu olmuştur. Özellikle Covid-19'dan kaynaklanan yapısal kırılmalar nedeniyle, proje bedellerinin artan maliyetleri karşılamadığı ve yüklenici firmaların zarar ettiği iddialarıyla, kamu kurumlarından sözleşme tutarlarının güncellenmesi talepleri gelmektedir. Bu noktada, yüklenicilerin öngörülemeyen maliyet artışlarını "basiretli tüccar" olarak tahmin edebilmesi gerektiği öne sürülmektedir. Alternatif yöntemlerle fiyat artışlarının öngörülebilirliğinin incelenmesi, literatüre ve ilgili kurumlara yol gösterici olabilir.

Zaman serisi kestirim yöntemleri genel olarak iki kategoride ele alınabilmektedir (Yu, 2024):

- İstatistiksel Yöntemler
- Makine Öğrenmesi Yöntemleri

Zaman serisi tahmini için kullanılan istatistiksel yöntemler, en erken sistematik ve formalize edilmiş yöntemlerdendir. Bu çalışmada kullanılan istatistiksel yöntemler ve kısaca ayırt edici yaklaşımları aşağıda listelenmiştir:

- ARIMA (Otoregresif Entegre Hareketli Ortalamalar): Doğrusal ilişkileri tespit etmekte etkili bir otoregresif zaman serisi tahmin yöntemidir (Data Aspirant, 2024).
- Prophet: Prophet, zaman serisi tahmini için kullanılan bir istatistiksel yöntemdir. Facebook şirketi tarafından geliştirilmiş olup, mevsimsellik, trendler ve tatil etkilerini içeren ek bir regresyon modeline dayanır. Prophet, modelinde gradyan artırma ve L1 düzenlileştirmesi gibi makine öğrenmesi teknikleri kullanırken, altta yatan veri üretim sürecine dair açık varsayımlara dayanan istatistiksel bir modele dayandığı için istatistiksel bir yöntem olarak sınıflandırılması daha uygun olacaktır (Krieger, 2024). Mevsimsellik ve tatil etkisi olan zaman serilerinde etkin bir tahmin yöntemidir.

Makine öğrenmesi yöntemleri, zaman serisi verilerinde tahmin yapmak için veri setini bir kesme tarihi itibarıyla eğitim ve test verileri olarak ayırarak eğitim verilerini modeli eğitmek ve modeli test verilerinde test etmek, tahminin gerçek veriden sapmasını gözlemlemek için kullanılabilir. Bu çalışmada kullanılan makine öğrenmesi yöntemleri aşağıda listelenmiştir:

- XGBoost: XGBoost, dağıtılmış gradyan artırma için çok yönlü ve taşınabilir optimize edilmiş bir kütüphanedir (Uma vd., 2023) (*XGBoost Documentation — Xgboost 1.7.5 Documentation*, n.d.). ARIMA'ya kıyasla, doğrusal olmayan veri setlerinde daha etkin olabilir. ARIMA'ya göre çok daha fazla farklı parametre kombinasyonları deneme imkânı sağlar.
- LSTM: İleri-beslemeli sinir ağlarından farklı olarak, LSTM geri besleme yeteneğine sahip yapay zekâ ve derin öğrenme yöntemidir (Hochreiter ve Schmidhuber, 1997). Uzun ve karmaşık zaman dizilerinde kestirim doğruluğu ön plana çıkmaktadır. XGBoost gibi çok sayıda parametre kombinasyonlarını deneme ve en iyisini seçme imkânı sağlar. Ancak, XGBoost'a göre daha karmaşık olan algoritması (hem uzun hem kısa dönem analiz yaptığı için) XGBoost'a göre çok daha fazla bilgisayar zamanı ve kaynağı kullanmasına neden olur.

Zaman serisi analizinin tarihsel gelişimi incelendiğinde, literatürde çeşitli yöntemler öne çıkmaktadır. İlk olarak otokorelasyon (AR) modelleri, bir zaman serisinin gelecekteki değerlerinin geçmiş değerlerine bağlı olduğunu varsayar ve serinin kendi içindeki önceki değerleri kullanarak gelecekteki değerleri tahmin eder (Yule, 1927). Bunu takiben, hareketli ortalama (MA) modelleri, geçmişte yapılan tahminlerin hatalarını dikkate alarak gelecekteki değerleri tahmin etmeye çalışır ve rastgele şokları modellemek için kullanılır (Slutzky,

1937). AR ve MA modellerinin birleşimi olan ARMA modeli ise serinin kendi iç dinamiklerini ve rastgele şok etkilerini dikkate alır (G. E. P. Box & Jenkins, 1970).

Zamanla, ARIMA (Otokorelasyon Entegre Hareketli Ortalama) modeli, zaman serilerinin durağan olmayan doğasını ele almak için geliştirilmiştir. Bu model, serinin fark alınmış değerleri üzerinden çalışarak durağan olmayan serileri analiz edebilir hale getirir (G. E. P. Box & Jenkins, 1970). Mevsimsel etkileri dikkate alan SARIMA (Mevsimsel ARIMA) modeli ise mevsimsel dalgalanmaları göz önünde bulundurarak daha doğru tahminler yapmayı sağlar (Brownlee, 2018).

ARIMA ve diğer istatistiksel yöntemlerin kısıtlamalarına cevap olarak son yıllarda, makine öğrenmesi ve derin öğrenme yaklaşımları zaman serisi analizinde önemli bir yer tutmaya başlamıştır. Karar ağaçları, destek vektör makineleri ve rastgele ormanlar gibi makine öğrenmesi algoritmaları, doğrusal olmayan ve karmaşık seyirlere sahip serilerde etkili olabilmektedir (Kontopoulou vd., 2023). Derin öğrenme yaklaşımlarından tekrarlayan sinir ağları (RNN), uzun kısa süreli bellek (LSTM) ve evrişimli sinir ağları (CNN), zaman serilerindeki bağımlılıkları ve uzun süreli eğilimleri yakalamada etkilidir (Hochreiter & Schmidhuber, 1997; Lecun vd., 1998).

Makine öğrenmesi ve derin öğrenme tekniklerinin dışında, özellikle finansal zaman serilerinin analizinde kullanılan daha geleneksel istatistiksel yöntemler de bulunmaktadır. Bu yöntemler arasında yer alan eşik otoregresyon (TAR) modelleri, zaman serilerindeki seyir değişikliklerini yakalamada kullanılır (Tong, 1983). Genelleştirilmiş otoregresif koşullu değişen varyans (GARCH) modelleri ise özellikle finansal zaman serilerinde gözlemlenen dalgalanmaları modellemek için geliştirilmiştir. GARCH modelleri, serinin geçmiş dalgalanmalarını dikkate alarak gelecekteki dalgalanmaları tahmin eder (Bollerslev, 1986). Durum uzayı modelleri ve Kalman filtresi ise zaman serisi verilerinin dinamik yapısını analiz etmek için kullanılır ve zamanla değişen sistemleri ve gözlem hatalarını etkili bir şekilde ele alabilir (Kalman, 1960).

Tekil modellerin güçlü yönlerinin beraber kullanılarak, daha doğru tahminler yapılmasını sağlayabileceği düşünülen, Fransızca'da "birlikte" anlamına gelen ve bu bağlamda "birliktelik yöntemleri" olarak da düşünülebilecek Ansambl (ensemble) yöntemleri, birden fazla tahmin modelinin sonuçlarını birleştirerek genel tahmin başarısını artırır. Bu

yöntemler, farklı modellerin güçlü yönlerini bir araya getirerek daha sağlam tahminler yapmayı sağlar (Rokach, 2010).

ARIMA modeli, George Box ve Gwilym Jenkins tarafından 1970'lerde geliştirilmiş olup, durağan olmayan verileri analiz etmek için tasarlanmıştır (G. Box ve Jenkins, 1976). Ekonomik veri setlerindeki durağan olmayan seriler için yaygın olarak kullanılır ve geçmiş verileri kullanarak gelecekteki değerleri tahmin etme kapasitesine sahiptir (Brockwell ve Davis, 2002).

Facebook Prophet, 2017 yılında Facebook tarafından geliştirilen ve kullanımı kolay bir zaman serisi tahmin aracıdır. Özellikle mevsimsellik ve tatiller gibi özel zaman etkilerini dikkate alarak tahminler yapar ve perakende satış tahminleri ve stok yönetimi gibi alanlarda kullanılır (Taylor ve Letham, 2018).

XGBoost, Chen ve Guestrin tarafından 2016'da tanıtılan gelişmiş bir karar ağacı tabanlı öğrenme algoritmasıdır. Hem sınıflandırma hem de regresyon problemlerinde kullanılır ve zaman serisi verilerindeki karmaşık yapıları ve ilişkileri etkili bir şekilde modelleyebilir (Chen ve Guestrin, 2016).

LSTM ağları, Sepp Hochreiter ve Jürgen Schmidhuber tarafından 1997'de geliştirilmiştir ve RNN'lerin uzun süreli bağımlılıkları modelleme zorluğunu çözmeyi amaçlar (Hochreiter ve Schmidhuber, 1997). Bu modeller özellikle doğal dil işleme ve konuşma tanıma gibi alanlarda kullanılır ve zaman serisi tahmininde uzun süreli geçmiş bilgileri koruma yeteneği ile öne çıkar (Gers vd., 2000).

Sonuç olarak; ARIMA'nın durağanlaştırılabilir zaman serileri için, LSTM'in uzun süreli bağımlılıklar için, XGBoost'un karmaşık veri yapıları için ve Prophet'in mevsimsel etkilere duyarlı seriler için etkili çözümler sunduğu öne sürülebilir.

Literatürde metodoloji olarak doğrudan bu çalışmaya benzer çalışmalar incelendiğinde; Özdemir (2020) makine öğrenimi yöntemleri ile geleneksel yöntemlerin performansını Ortalama Mutlak Yüzdesel Hata (MAPE) ve Simetrik Ortalama Mutlak Yüzdesel Hata (sMAPE) metriklerini kullanarak karşılaştırmış ve genellikle makine öğrenmesi yöntemlerinin geleneksel yöntemlerden daha üstün olduğunu göstermiştir (Özdemir, 2020).

Karatekin ve Başaran (2022) Türkiye gün öncesi elektrik piyasasında, elektrik enerjisi fiyat tahmini için lineer regresyon, polinomiyal regresyon, yapay sinir ağları ve XGBoost yöntemlerini RMSE ölçütüne göre kıyaslamış ve XGBoost modelinin en doğru tahmini yaptığını ortaya koymuştur (Karatekin & Başaran, 2022). Yelgeç (2022) durağan olmayan rüzgar gücü verisi tahmini için MAE, RMSE ve R^2 ölçütleriyle yalnızca XGBoost modelini uygulayarak ayrık dalgacık dönüşümü-Xgboost modelinin daha başarılı sonuçlar verdiği görülmüştür (Yelgeç ve Bingöl, 2022). Oukhouya ve El-Himdi (2023) Fas borsasının tahmini için LSTM, SVR, XGBoost ve MLP makine öğrenmesi yöntemlerini Ortalama Hata (ME), MAE, MSE, Ortalama Yüzdesel Hata (MPE), MAPE, RMSE ve R^2 ölçütlerini kullanarak karşılaştırmış ve SVR'nin en doğru sonuçları verdiğini bulmuşlardır (Oukhouya & El Himdi, 2023). Aqbar ve Supono (2023) ise ekonomik zaman serisi kestirim çalışmalarında LSTM ve XGBoost yöntemlerinin performanslarını RMSE ve MAPE skorlarına göre karşılaştırmış ve LSTM'in daha yüksek kestirim doğruluğu sağladığını belirlemişlerdir (Aqbar ve Supomo, 2023). Aliyev (2024) istatistiksel yöntemlerden ARIMA, SARIMA; makine öğrenmesi yöntemlerinden Rastgele Orman, XGBoost, LSTM ve GRU modellerini R^2, MAPE, MAE ve RMSE ölçütlerine göre karşılaştırmış, en uygun ölçütlerin GRU modeliyle elde edildiğini ortaya koymuştur (Aliyev, 2024).

Çalışmada yer alan iki veri seti birbirinden farklı para birimleri ile fiyat endeksi verilerini göstermektedir. Türk lirası enflasyonu ile dolar bazında enflasyonun farklılık arz etmesi, elbette ki, metotların etkinliğini veri seti bazında farklı yönde etkileyebilecektir. Yine, Dolar bazında enflasyon verisi, Türk Lirası bazında enflasyon verisine göre çok daha fazla geçmiş veri içermekte ve bu da daha fazla yapısal kırılmaya maruz kalmış olma ihtimalini yükseltmektedir. Bütün bunlar göz önüne alındığında, her veri seti türü farklı olarak ele alınmış ve her iki veri setine de birbirinden farklı güçlü yönleri bulunan ARIMA, Prophet, XGBoost ve LSTM metodları ayrı ayrı uygulanmış ve kestirim performansları ayrı ayrı sunulmuştur.

Bu araştırma, zaman serisi verilerinde yapısal kırılmalar bağlamında farklı tahmin modellerinin kapsamlı bir karşılaştırmasını sağlayarak mevcut bilgi birikimine katkıda bulunmayı amaçlamaktadır. Birbirinden oldukça farklı algoritmalara sahip olmalarından dolayı, ARIMA, Prophet, XGBoost ve LSTM yöntemlerinin performanslarını değerlendirerek ve karşılaştırarak, bu çalışma yapısal düzensizliklere sahip çeşitli gerçek dünya senaryolarında tahmin yapmak için en uygun modellerin seçilmesi konusunda

uygulayıcılara ve araştırmacılara değerli kavrayışlar sunmayı hedeflemektedir. Bu doğrultuda, Covid-19 sonrası sıklıkla yaşanan fiyat endekslerinde ani artışın öngörülebilir olup olmadığı konusunda yaşanan ihtilafların çözümlenmesinde de yardımcı olabilecek bir kaynak olması da hedeflenmektedir.

İkinci bölüm olan Materyaller ve Metotlar bölümünde, başlangıçta veri seti tanıtılarak temel bir bağlam oluşturulmuştur. Ardından, kodlama ortamı açıklanarak, hesaplama çerçevesine dair bilgi verilmiştir. Kodlama adımları özetlenmiş ve seçilen yapısal kırılma tespit ve zaman serisi tahmin yöntemleri için teorik açıklamalar sağlanmıştır. Son olarak, her yönteme özgü ince ayar prosedürleri detaylandırılmıştır.

Üçüncü bölüm olan Uygulama bölümünde, uygulanan modellerden elde edilen tahmin doğrulukları gösterilmiş ve karşılaştırılmıştır.

Daha sonra dördüncü bölüm olan Tartışma bölümünde, sonuçlar yorumlanmıştır.

Son olarak, beşinci ve son bölüm olan Sonuç ve Öneriler bölümünde çıkan sonuçlardan yola çıkarak yapısal kırılmaların öngörülebilir olup olmadığı konusunda bazı yorumlarda bulunulmuş ve bu konuda araştırma yapanlara geliştirme aşamasında olan yeni yöntemlerle tahmin performanslarını geliştirme imkanları olduğu konusunda yol gösterilmiştir.

2. MATERYALLER VE METOTLAR

Çalışmada kullanılan yapısal kırılma tespit, zaman serisi tahmin yöntemleri ve parametre ayarlanması aşamaları bu bölümde incelenecektir.

2.1. Yapısal Kırılma Tespit Testleri

Yapısal kırılmalı zaman serileri, ekonomik, finansal veya çevresel verilerde belirli bir döneme ait ani ve beklenmedik değişiklikleri ifade eder. Bu tür kırılmaları tespit etmek ve analiz etmek için Chow Testi, Bai-Perron, Zivot-Andrews ve CUSUM (Birikimli Toplam Kontrol Grafiği) testi gibi yöntemler kullanılır. Birbirinden oldukça farklı yapısal kırılma tespit yöntemleri olduğundan dolayı Chow ve CUSUM testleri bu çalışmada kullanılacak, bu kısımda Chow ve CUSUM testleri açıklanacaktır.

2.1.1. Chow testi

Chow testi, zaman serisi veya panel verilerinde yapısal kırılmaların varlığını test etmek için kullanılır. Gregory Chow tarafından 1960 yılında geliştirilen bu test, özellikle ekonomik modellerde politika değişiklikleri veya dışsal şoklar sonucu oluşabilecek kırılmaları belirlemek için uygundur. Testin temel prensibi, veri setinin iki alt döneme ayrılması ve her iki dönem için ayrı ayrı regresyon modellerinin tahmin edilmesidir. Daha sonra, bu iki model arasındaki farkın istatistiksel olarak anlamlı olup olmadığı değerlendirilir.

Chow testinin formülü Eş. 2.1.'deki gibidir:

$$F = \frac{(RSS_c - RSS_1 + RSS_2)/k}{(RSS_1 + RSS_2)/(N-2k)} \tag{2.1}$$

Burada:

- RSS_c: Tüm veri seti için tek bir regresyon modeli kullanılarak hesaplanan Artık Kareler Toplamı
- RSS_1: Kırılma noktasından önceki veri için hesaplanan Artık Kareler Toplamı.
- RSS_2: Kırılma noktasından sonraki veri için hesaplanan Artık Kareler Toplamı.
- k: Modeldeki bağımsız değişken sayısı.

- N : Gözlem sayısı.

Eğer F istatistiği belirli bir güven aralığındaki kritik değerden büyükse, yapısal bir kırılma olduğu sonucuna varılır (Chow, 1960). F istatistiğinin, F dağılımından elde edilen kritik değerden nispeten en büyük olduğu tarih, yapısal kırılmanın en büyük olduğu tarihtir.

2.1.2. CUSUM testi (birikimli toplam kontrol grafiği)

CUSUM testi, zaman serilerindeki birikimli toplam değişiklikler üzerinden yapısal kırılmaları tespit etmeye yarayan bir yöntemdir. Bu test, özellikle parametrelerin zaman içinde değişip değişmediğini kontrol etmek için kullanılır. Brown, Durbin ve Evans (1975) tarafından geliştirilen CUSUM testi, parametre tahminlerinin birikimli toplamlarının grafik üzerinde gösterilmesiyle yapısal değişiklikleri gözlemlemeyi amaçlar.

CUSUM testi, regresyon artıklarının birikimli toplamını kullanarak yapısal kırılmaları tespit eder. Formül Eş. 2.2'deki gibidir.

$$S_t = \sum_{i=1}^{t} e_i \qquad (2.2)$$

Burada $e_i = Y_i - \hat{Y}_i$, i zamanındaki regresyon artıklarıdır. S_t ise t zamanındaki birikimli toplamdır. S_t serisinin belirgin bir şekilde yükselmesi veya düşmesi, yapısal bir değişiklik olduğuna işaret eder (Brown et al., 1975).

Bu testler, yapısal kırılmaları tespit etmek için kullanıldığında, verilerin zaman serisine uygun şekilde modellenmesi gerektiğini unutmamak önemlidir. Chow testi daha kesin bir kırılma noktası önerirken, CUSUM testi genel eğilimlerdeki değişiklikleri yakalar (Brown et al., 1975).

CUSUM testinde üst ve alt sınırlar, genellikle standart normal dağılım kullanılarak belirlenir. Kritik sınırlar aşağıdaki formüllerle hesaplanır:

- Üst Sınır = $a + bt$
- Alt Sınır = $-a - bt$

Burada,

- a : Sabit bir değer olup genellikle %95 güven seviyesi için yaklaşık 0.948 olarak kullanılır.
- b: Birikimli toplamın eğilimini belirleyen bir katsayıdır ve genellikle gözlem sayısının karekökü ile belirli bir katsayının çarpımı olarak $b = k\sqrt{N}$ şeklinde hesaplanır. Burada,
- N: Toplam gözlem sayısı
- k: Belirli bir güven seviyesine bağlı olarak belirlenen sabit bir katsayıdır (Brown vd., 1975). %95 güven seviyesi için genellikle 0.948'dir.

2.2. Zaman Serisi Tahmin Yöntemlerinin İncelenmesi

Bu kısımda, birbirinden farklı ARIMA, Prophet, XGBoost ve LSTM yöntemleri incelenecektir.

2.2.1. ARIMA yöntemi

ARIMA modeli genellikle ARIMA (p, d, q) fonksiyonuyla temsil edilir. Burada,

- p: Otoregresyon (AR) bileşeninin derecesi, mevcut değeri tahmin etmek için kaç geçmiş değerin dikkate alındığını belirtir.
- d: Zaman serisini durağan hale getirmek için gereken fark derecesi. Bu, Entegre (I) bileşenidir.
- q: Hareketli Ortalama (MA) bileşeninin derecesi, mevcut değeri tahmin etmek için kaç geçmiş tahmin hatasının dikkate alındığını belirtir (Drebee et al., 2023).

ARIMA modeli kelimelerle aşağıdaki gibi ifade edilebilir.

Öngörülen değer = Sabit değer + Y'nin gecikmelerinin p gecikmesine kadar doğrusal kombinasyonu + tahmin hatalarının q gecikmesine doğrusal kombinasyonu

ARIMA, Eş. 2.3.'te gösterilen durağan ARMA modelindeki trendi I (integrated) harfiyle ifade edilen d bileşeniyle fark alarak yataylaştırarak durağanlaştırır:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-3} + \cdots + \varphi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (2.3)$$

Burada:

- y_t: Tahmin edilen değer
- φ: Otoregresif terimlerin katsayıları
- e_t: Hata terimidir.

ARIMA modelinin formülü ise Eş. 2.4'te verilmiştir.

$$\varphi(B)(1 - B^d)Y_t = \theta(B)e_t \quad (2.4)$$

Burada: B geri kaydırma operatörü olup, $B^d Y_t = Y_{t-d}$, d ise zaman serisini durağanlaştırmak için gereken d'inci farktır. e_t sıfır ortalama ve sabit varyanslı, bağımsız ve aynı dağılıma sahip beyaz gürültü terimidir. $\varphi(B)$ ve $\theta(B)$ ise, sırasıyla otoregresif ve hareketli ortalama polinom terimleridir (Özdemir, 2020).

ARIMA modeli, bu üç bileşeni zaman serisi verilerini modellemek için birleştirir. Örneğin, bir ARIMA(2, 1, 3) modeli, 2. dereceden otoregresyon, 1. dereceden fark ve 3. dereceden hareketli ortalama terimlerini içeren bir formüle sahip olacaktır.

2.2.2. Prophet yöntemi

Prophet modeli, Facebook tarafından geliştirilen ve mevsimsel verilerle başa çıkma konusunda esneklik sağlayan sağlam tahminler sunan bir zaman serisi tahmin metodudur. Prophet modeli Eş. 2.5'de verilmektedir.

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \quad (2.5)$$

Burada,

- g(t): Trend bileşenini temsil eder ve uzun vadeli değişiklikleri yakalar.
- s(t): Mevsimsellik bileşenini modelleyerek periyodik desenleri yakalar.
- h(t): Tatilleri ve özel olayları hesaba katar.
- ε(t): Hata terimini temsil eder.

Prophet, kullanım kolaylığı, otomatik tatil modellemesi ve eksik veriler ile aykırı değerlerle etkili bir şekilde başa çıkabilme yeteneği ile bilinir. Çeşitli uygulamalarda zaman serisi verilerini tahmin etmede gösterdiği performans ile popülerlik kazanmıştır (Facebook Core Data Science Team, 2017; Krieger, 2024; Yibin, 2019; Almazrouee et al., 2020).

2.2.3. XGBoost yöntemi

XGBoost, Karar Ağaçlarından başlayarak torbalama, ardından Rastgele Orman, sonra Güçlendirme, ardından Gradyan Güçlendirme ve en sonunda XGBoost (optimize edilmiş gradyan güçlendirme) olmak üzere bu alanda en gelişmiş modeldir (Bhattiprolu, 2021). Güçlendirme, yanlış tahminlerde bulunan örneklerin önemini artırarak geçmiş hatalardan öğrenirken, gradyan güçlendirme kayıp fonksiyonunu minimize etmek için gradyan inişini kullanır (Nanthini vd., 2023). Son olarak, XGBoost, gradyan yönünü belirlemek için ikinci dereceden gradyanları hesaplar, aşırı uyumu önlemek için düzenleme yöntemleri kullanır ve güçlendirme işlemini hızlandırır.

XGBoost genellikle dolandırıcılık tespiti gibi durumlarda sınıflandırma yapabilmek için veya ev fiyatlarını tahmin etmek gibi regresyon analizleri için kullanılır. Ayrıca,, XGBoost, zaman serisi verilerinde tahmin yapmak amacına yönelik olarak uyarlanabilir (Wijaya, 2023).

XGBoost'un başlangıç makalesine (Chen & Guestrin, 2016) dayanarak özetlenmiş ve basitleştirilmiş formüller Eş. 2.6, Eş. 2.7 ve Eş. 2.8'le gösterilmiştir.

- $F(x)$: Son güçlendirilmiş model
- h_i: Zayıf öğrenici
- i: Yineleme sayısı olmak üzere,

$$F(x) = \sum_{i=1}^{n} h_i(x) \tag{2.6}$$

dır. Amaç, belirli bir kayıp fonksiyonu $L(y, F(x))$'yi en aza indirgeyerek ardışık olarak zayıf öğreniciler ekleyerek $F(x)$'i bulmaktır. Kayıp fonksiyonu, modelin tahminlerinin gerçek değerlerden ne kadar uzak olduğunu ölçer. XGBoost gibi gradyan güçlendirme algoritmalarında, modelin performansını optimize etmek için kayıp fonksiyonu minimize

edilir. Kayıp fonksiyonunun doğru tanımlanması ve optimize edilmesi, modelin doğruluğunu ve genelleme yeteneğini artırır.

Bu çalışmada kullanılan kayıp fonksiyonu regresyon problemlerinde kullanılan MSE olup, Eş. 2.7'de gösterilmiştir:

$$L\big(y, F(x)\big) = \frac{1}{n}\sum_{i=1}^{n}\big(\,(y_i - \big(F(x_i)\big)^2 \qquad (2.7)$$

- y_i: $i'inci\ gerçekleşen\ değer$
- $F(x_i)$: $i'inci\ tahmin\ değeri$
- n: $Veri\ noktalarının\ sayısı$

Her iterasyonda, yeni bir zayıf öğrenici, kayıp fonksiyonunu en aza indirerek topluluğa eklenir:

$$F_t(x) = F_{t-1}(x) + \gamma \cdot h_t(x) \qquad (2.8)$$

- $F_{t-1}(x)$: İterasyon $t-1$'e kadar güçlendirilmiş model
- $h_t(x)$: İterasyon t'deki zayıf öğrenici
- γ: Güncelleme adım boyutunu kontrol eden öğrenme hızı (Kawakita vd., 2005)

XGBoost'un diğer ağaç güçlendirme algoritmalarına göre ayırt edici özelliği, zayıf öğrenicilerdeki karmaşıklığı giderene kadar yeniden iterasyon yaparak aşırı uyumu önlemek için düzenleme işlevidir:

$$Obj = \sum_{i=1}^{n} L(y_i, F(x_i)) + \sum_{i=1}^{N} \Omega(h_i) \qquad (2.9)$$

$\Omega(h_i)$: Düzenleme terimi

Optimizasyon, gradyan inişi yoluyla gerçekleştirilir, zayıf öğrenicilerin parametrelerini ayarlayarak amaç fonksiyonunu en aza indirir (Chen ve Guestrin, 2016).

2.2.4. LSTM yöntemi

LSTM modeli, benzersiz mimarisi sayesinde verilerdeki ardışık bağımlılıkları yakalama konusunda etkilidir. Bu mimari, üç temel bileşeni içerir. Bunlar,

Giriş Kapısı (i(t)): Bellek hücresinde hangi yeni bilgilerin saklanacağını belirler,

Unutma Kapısı (f(t)): Bellek hücresinden hangi bilgilerin atılacağını düzenler,

Çıkış Kapısı (o(t)): Bellek hücresinin hangi kısmının çıkış olarak kullanılacağını kontrol eder (Baraldi et al., 2017).

LSTM modelleri, ardışık verileri adım adım işler ve her zaman adımında iç bellek hücresini günceller. Eş. 2.9'la gösterilen formül bu süreci özetler:

$$C(t) = f(t) * C(t-1) + i(t) * g(t) \qquad h(t) = o(t) * tanh(C(t)) \qquad (2.10)$$

Burada,

- C(t), zaman içinde bilgiyi tutan hücre durumunu temsil eder,
- h(t), çıkışı sağlayan gizli durumdur,
- i(t), f(t), o(t), bilgi akışını kontrol eden kapı değerleridir,
- g(t), hücre durumu için aday bir değer hesaplar.

Bilgi akışını etkili bir şekilde yöneterek, LSTM modelleri dizi tahmini, dil modelleme ve zaman serisi tahmini gibi görevlerde başarılı olur (Hochreiter ve Schmidhuber, 1997; Kai vd., 2019).

2.3. Parametre Optimizasyonu

Modellerin aşırı uyumla eksik uyum arasında optimum bir denge kurmasını sağlar. Böylece, hem tahminlerin doğruluğunu artırırken, hem de aşırı uyumu (overfitting) engeller. Bu çalışmada her model için farklı parametreler denenmiş ve en optimum parametreler belirlenmiştir.

ARIMA için, en optimum parametreleri otomatik olarak bulmak üzere auto-arima fonksiyonu kullanılmıştır.

Prophet, her uzmanlık düzeyinden kullanıcının kolaylıkla kullanabilmesi amacıyla üretilmiş oldukça otomatik bir araçtır ve varsayılan değerler genellikle en iyi sonuçları verir (Yusof vd., 2021). Dolayısıyla parametre optimizasyonu için çok fazla seçenek yoktur. Bu nedenle Prophet tahmin yöntemi için, değiştirilmesi mümkün olan sayısal olmayan bir parametre manuel olarak değiştirilmiştir. Yapılan değişiklik, Prophet modelini trend içeren veri kümelerine uyum sağlamak için model = Prophet(growth='logistic') kodu ile gerçekleştirilmiştir:

LSTM modeli için başlangıçta veriler ölçeklendirilmiş, daha sonra GridSearch, çapraz doğrulama ve manuel parametre ayarı yapılmıştır. LSTM yönteminde, Farklı geriye dönük "look_back" değerleri teker teker manuel olarak denenmiştir. Nitekim, çok geniş bir geriye dönük look_back değer aralığı ile bir grid araması yapmak bilgisayarın görevi tamamlaması için gereken süreyi çok fazla artırabilmektedir. Ek yoğun katman manuel olarak eklenerek denenmiştir. LSTM katmanlarının sayısı, dönem sayısı, öğrenme oranları gibi farklı kombinasyonlar GridSearch yöntemi ile test edilmiş ve en optimum kombinasyon tespit edilmiştir.

XGBoost modeli için grid araması ve çapraz doğrulama uygulanmıştır. En iyi parametreler elde edilmiş ve ardından model en iyi parametrelerle uygulanmıştır.

3. UYGULAMA

Sırasıyla; veri toplama, yapısal kırılma testleri, tahmin modellerinin uygulanması ve karşılaştırılması aşamaları gerçekleştirilmiştir.

3.1. Veri Toplama

Bu çalışmada, iki farklı veri seti kullanılmıştır. İlki, Türkiye'de 01.2005-10.2022 arası aylık bazda yıllık TÜFE % değişim oranı, ikincisi de ABD'de 01.1913-01-2014 arası aylık TÜFE zaman serisidir. Türkiye TÜFE verisinde her ay için "Tufe_yillik" olarak ifade edilen ve yıllık fiyat endeksini belirten değerler, bir yıl önceki TÜFE değeriyle o ayki TÜFE değeri arasındaki yüzdesel farkı ifade etmektedir. ABD TÜFE veri setinde ise Index adı altında doğrudan endeks değerleri verilmektedir.

Türkiye TÜFE verileri Türkiye Cumhuriyet Merkez Bankası internet sitesindeki Tuketici Fiyatlari bölümünden elde edilmiştir (TCMB, 2022). ABD TÜFE verisi ise ABD İşgücü İstatistikleri Bürosu internet sitesinden elde edilmiştir (US Bureau of Labor Statistics, 2023).

Türkiye TÜFE ve ABD TÜFE ham verileri Python programlama dili ile Anaconda Jupyter Notebook'ta okutulduğunda ilk 5 ve son 5 veri sırasıyla Şekil 3.1 ve Şekil 3.2.'de gösterilmiştir.

	Date	Tufe_yillik	Year
0	2022-10	85.51	2022
1	2022-09	83.45	2022
2	2022-08	80.21	2022
3	2022-07	79.60	2022
4	2022-06	78.62	2022
..	...	...	...
209	2005-05	8.70	2005
210	2005-04	8.18	2005
211	2005-03	7.94	2005
212	2005-02	8.69	2005
213	2005-01	9.24	2005

[214 rows x 3 columns]

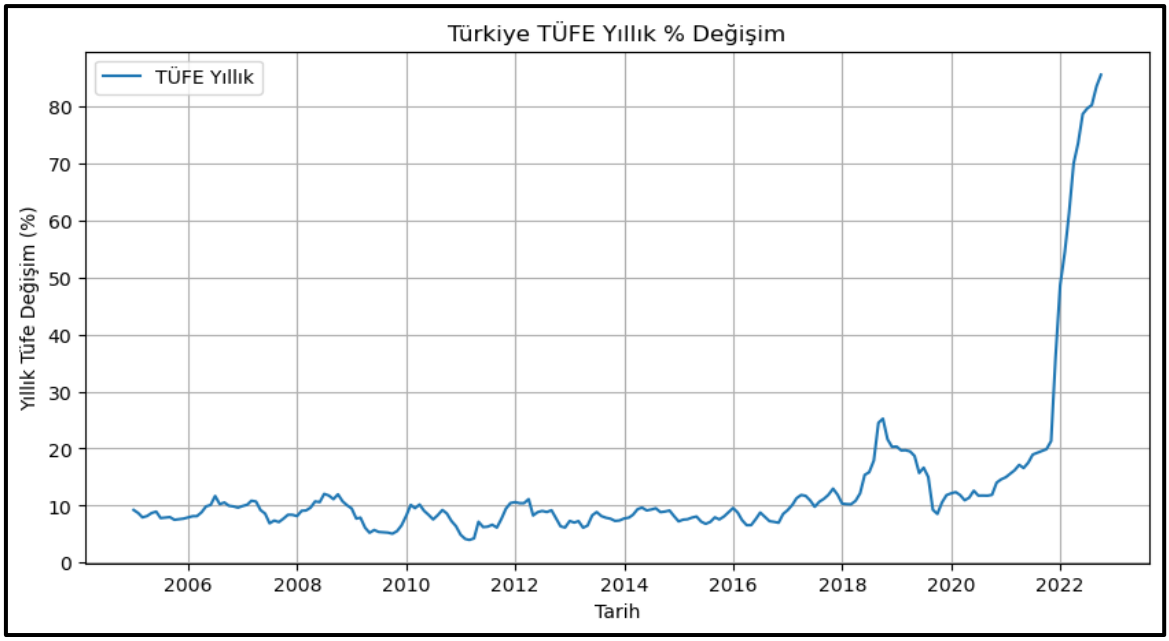
Şekil 3.1. Türkiye TÜFE yıllık değişim ilk ve son 5 ay

	Date	Index	Inflation
0	1913-01-01	9.800	NaN
1	1913-02-01	9.800	0.00
2	1913-03-01	9.800	0.00
3	1913-04-01	9.800	0.00
4	1913-05-01	9.700	-1.02
...	...	...	...
1208	2013-09-01	234.149	0.12
1209	2013-10-01	233.546	-0.26
1210	2013-11-01	233.069	-0.20
1211	2013-12-01	233.049	-0.01
1212	2014-01-01	233.916	0.37

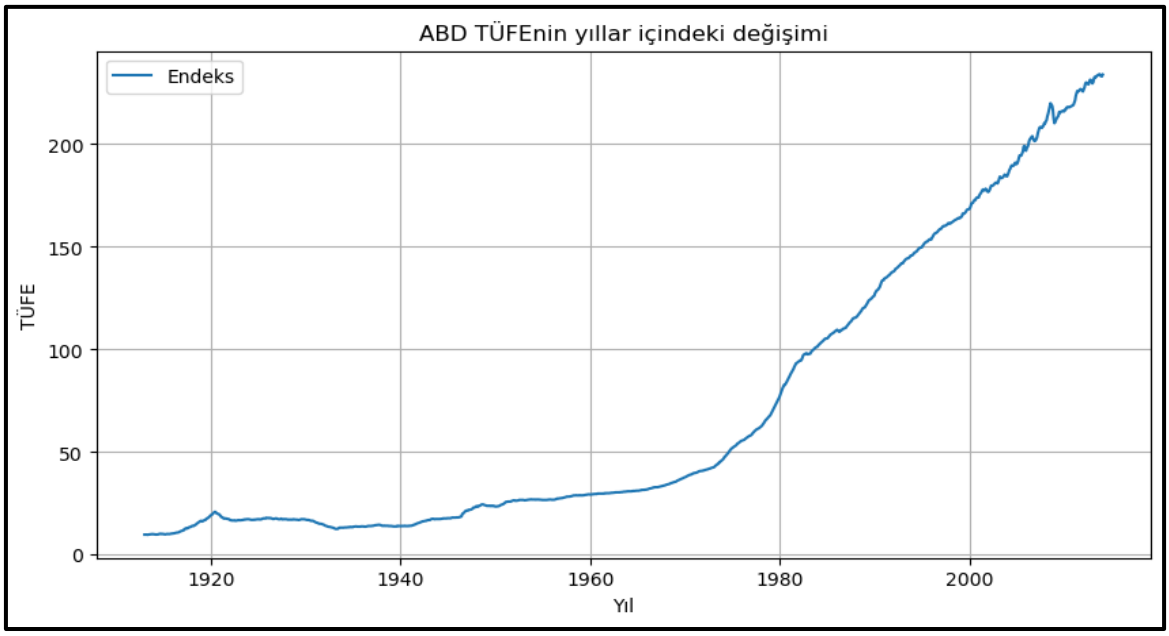
1213 rows × 3 columns

Şekil 3.2. ABD TÜFE yıllık değişim ilk ve son 5 ay

Şekil 3.3 ve Şekil 3.4'te ise iki zaman serisinin grafiksel seyri görselleştirilmiştir.



Şekil 3.3. Türkiye TÜFE zaman serisi gerçekleşme grafiği



Şekil 3.4. ABD TÜFE zaman serisi gerçekleşme grafiği

3.2. Yapısal Kırılma Testleri

3.2.1. Türkiye ve ABD TÜFE zaman serisine Chow testinin uygulanması

Bu kısımda, Türkiye ve ABD TÜFE zaman serisi verilerine ilişkin kırılma noktaları olup olmadığını tespit etmek için Chow testi kullanılacaktır.

3.2.2. Kodlama ortamının seçimi ve hazırlanması

Bu aşamada, Anaconda Jupyter Notebook üzerinde Python programlama dili kullanılmıştır. Öncelikle, gereken kütüphaneler olan; karmaşık matematiksel işlemler için numpy, veri manipülasyonu için pandas, veri görselleştirme için matplotlib, istatistiksel işlemler için statsmodels, bilimsel ve teknik hesaplamalar için scipy kütüphaneleri yüklenmiştir. Daha sonra, gereken kodlarla veri okutulmuştur.

3.2.3. Türkiye TÜFE zaman serisine Chow testinin uygulanması

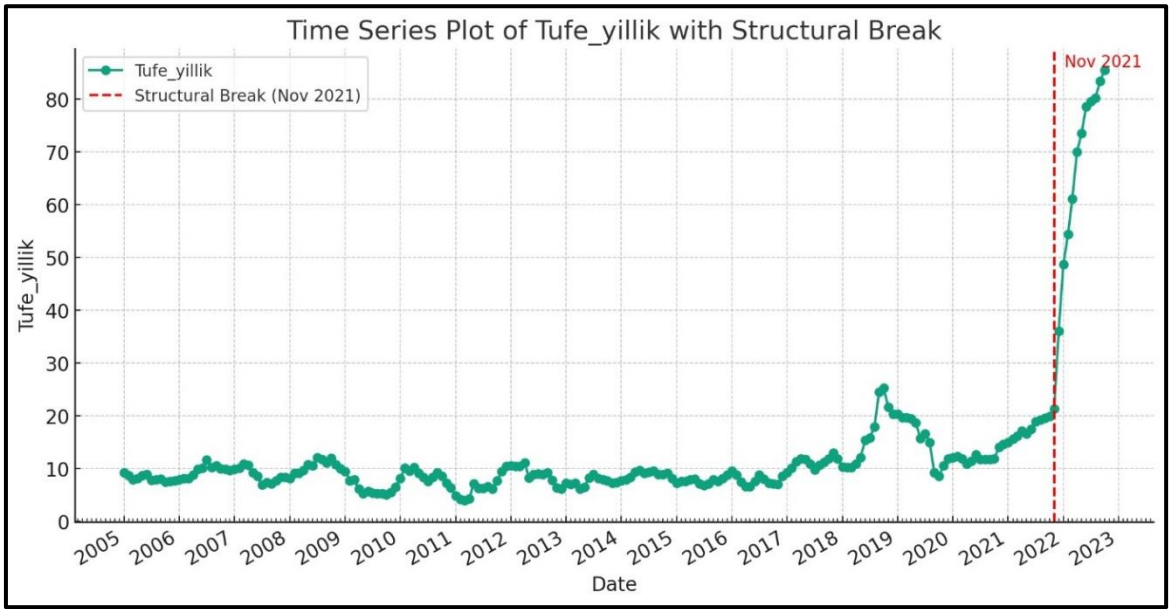
Bu aşamada, Türkiye TÜFE zaman serisine Chow testi uygulanarak yapısal kırılma olup olmadığı tespit edilmiştir.

Öncelikle, veri ön işlemesi Python ile gerçekleştirilmiştir. Pythyon'da CSV dosyasında sütunların Avrupa ülkelerinde virgül noktalık ayıracı olarak kullanıldığından, sütunların yaygın olduğu şekilde noktalı virgülle ayrılmış olmasından dolayı, öncelikle ayıracın ";" olduğu belirtilmiştir. Daha sonra, ondalık ayıracı olan virgül, Python ve pandas paketinde uyumlu bir şekilde çalışabilmesi için noktaya dönüştürülmüş ve TÜFE verileri ondalıklı sayı formatı olan "float"a dönüştürülmüştür. Ardından, Türkiye TÜFE verisindeki 'Date' sütunundaki değerler pandas kütüphanesindeki 'to_datetime' fonksiyonu ile Python'un 'datetime' nesnelerine dönüştürülmüş, tarih formatı da yıl-ay-gün olarak belirlenmiştir. Bu sütundaki yalnızca yıl bilgileri çekilerek 'Year' adıyla oluşturulan yeni ve ayrı bir sütunda sadece yılların gösterilmesi sağlanmıştır.

Daha sonra, Chow testi Python ortamında uygulanmıştır. En büyük kırılmanın ne zaman olduğu Chow testi sonuçlarının maksimum olduğu düzeyin "for" döngüsüyle bulunmasıyla elde edilmiştir. Buna göre elde edilen çıktı aşağıdaki gibidir:

(Timestamp('2021-11-01 00:00:00'), 2609.690788739996)

Çıktı ve Şekil 3.5'te gösterilen grafikten, Kasım 2021'de en büyük kırılmanın gerçekleştiği görülebilmektedir.



Şekil 3.5. Chow Testi ile Türkiye TÜFE zaman serisinin yapısal kırılma noktası tespiti

Ardından, bu tarihteki Chow testi değerleri baz alınarak, Şekil 3.6.'daki çıktı sonuçları elde edilmiştir:

```
Chow İstatistiği: 2609.690789
Kritik Değer: 3.038877
Yapısal kırılma vardır.
```

Şekil 3.6. Türkiye TÜFE Chow testi sonucu

Buna göre, en büyük kırılmanın çok aşikar bir yapısal kırılma olduğu tespit edilmiştir. Çünkü kritik değer olan 3.04, Chow istatistiği değeri olan 2609,7'nin oldukça altında kalmıştır.

3.2.4. ABD TÜFE Chow testi uygulanması

Bu aşamada, ABD TÜFE zaman serisine Chow testi uygulanarak, yapısal kırılma olup olmadığı tespit edilmiştir.

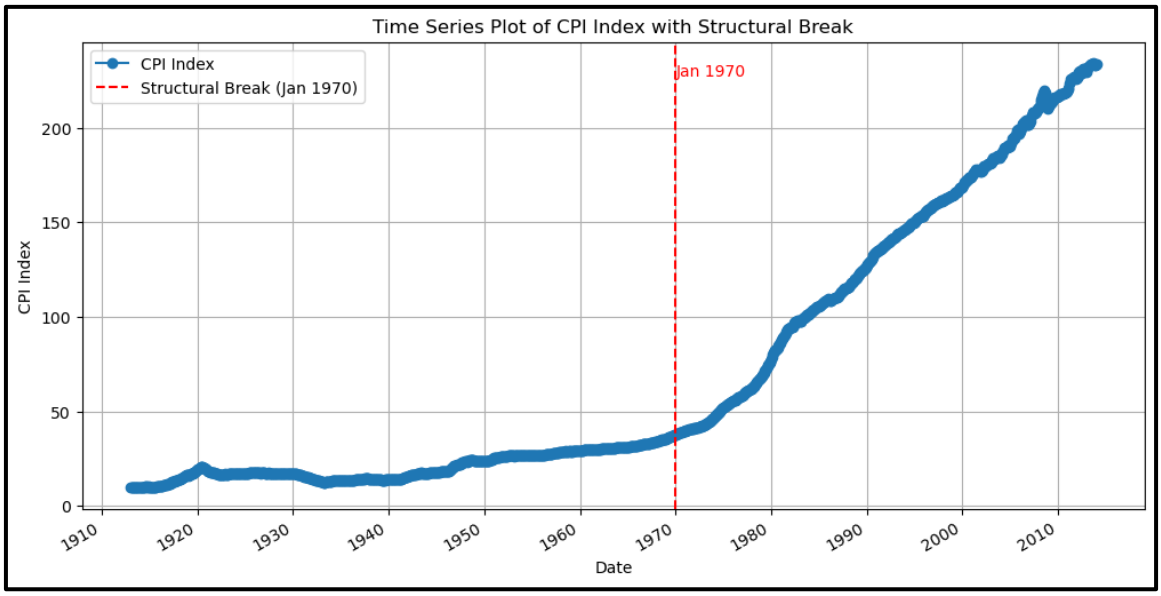
Öncelikle Python ortamında veri ön işlemesi gerçekleştirilmiştir. ABD TÜFE verisindeki 'Date' sütunundaki değerler pandas kütüphanesindeki 'to_datetime' fonksiyonu ile Python'un 'datetime' nesnelerine dönüştürülmüş, tarih formatı da yıl-ay-gün olarak belirlenmiştir.

Daha sonra, Python ortamında en büyük kırılmanın ne zaman olduğu Chow testi sonuçlarının maksimum olduğu düzeyin "for" döngüsüyle bulunmasıyla elde edilmiştir. Buna göre:

```
En büyük yapısal kırılmanın yaşandığı tarih: 1970-01-01 00:00:00,
Bu tarih itibarıyla Chow Testi istatistiği: 50255.54893115546
Kritik Değer: 3.0031675548845884
Yapısal kırılma vardır.
```

Şekil 3.7. ABD TÜFE Chow testi sonucu

Şekil 3.7.'deki çıktı elde edilmiştir. Bu çıktıya göre elde edilen grafik de aşağıdaki şekilde dir:



Şekil 3.8. Chow Testi ile ABD TÜFE zaman serisinin yapısal kırılma noktası tespiti

Şekil 3.7.'deki çıktı ve Şekil 3.8.'deki grafikten de görülebileceği üzere; Chow testi istatisiği F kritik değer ise 3 civarındadır. Buradan da büyük bir yapısal kırılma olduğu açıkça görülebilmektedir.

3.2.5. Türkiye ve ABD TÜFE zaman serisine CUSUM testinin uygulanması

Bu aşamada, Türkiye ve ABD TÜFE zaman serilerine CUSUM testi uygulanarak, yapısal kırılma olup olmadığı tespit edilmiştir.

3.2.6. Kodlama ortamının seçimi ve hazırlanması

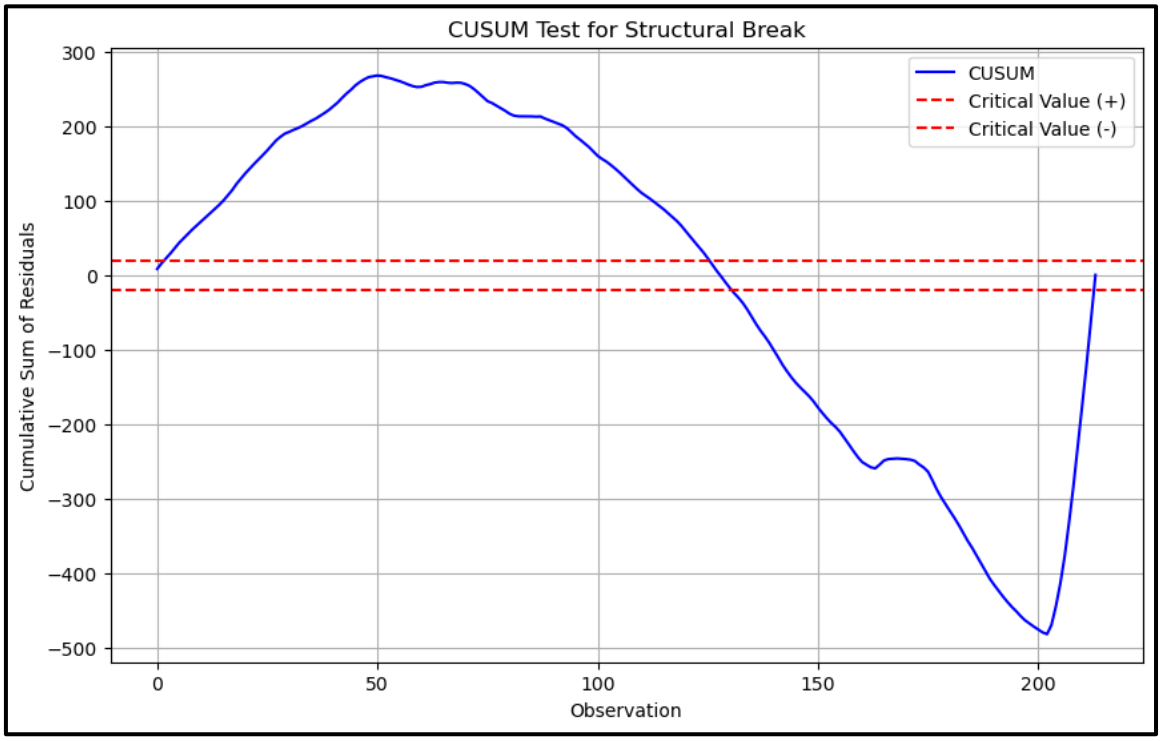
Bu aşamada, Anaconda Jupyter Notebook üzerinde Python programlama dili kullanılmıştır. Öncelikle, gereken kütüphaneler olan numpy, pandas, matplotlib, statsmodels kütüphaneleri yüklenmiştir. Daha sonra, gereken kodlarla veri okutulmuştur.

3.2.7. Türkiye TÜFE zaman serisine CUSUM testinin uygulanması

İlk etapta veri ön işlemesi Python ortamında gerçekleştirilmiştir. Pythyon'da CSV dosyasında sütunların Avrupa ülkelerinde virgül noktalık ayıracı olarak kullanıldığından, sütunların yaygın olduğu şekilde noktalı virgülle ayrılmış olmasından dolayı, öncelikle ayıracın ";" olduğu belirtilmiştir. Daha sonra, ondalık ayıracı olan virgül, Python ve pandas

paketinde uyumlu bir şekilde çalışabilmesi için noktaya dönüştürülmüş ve TÜFE verileri ondalıklı sayı formatı olan “float”a dönüştürülmüştür. Ardından, Türkiye TÜFE verisindeki ‘Date’ sütunundaki değerler pandas kütüphanesindeki ‘to_datetime’ fonksiyonu ile Python'un ‘datetime’ nesnelerine dönüştürülmüş, tarih formatı da yıl-ay-gün olarak belirlenmiştir.

Gereken veri ön işleme işlemlerinden sonra CUSUM testi uygulandığında aşağıdaki grafik ve çıktı elde edilmiştir:



Şekil 3.9. Türkiye TÜFE CUSUM testi grafiği

```
Yapısal Kırılma Tespit Edildi: True
Maksimum Sapma: 482.3741
Kritik Değer: 18.9543
```

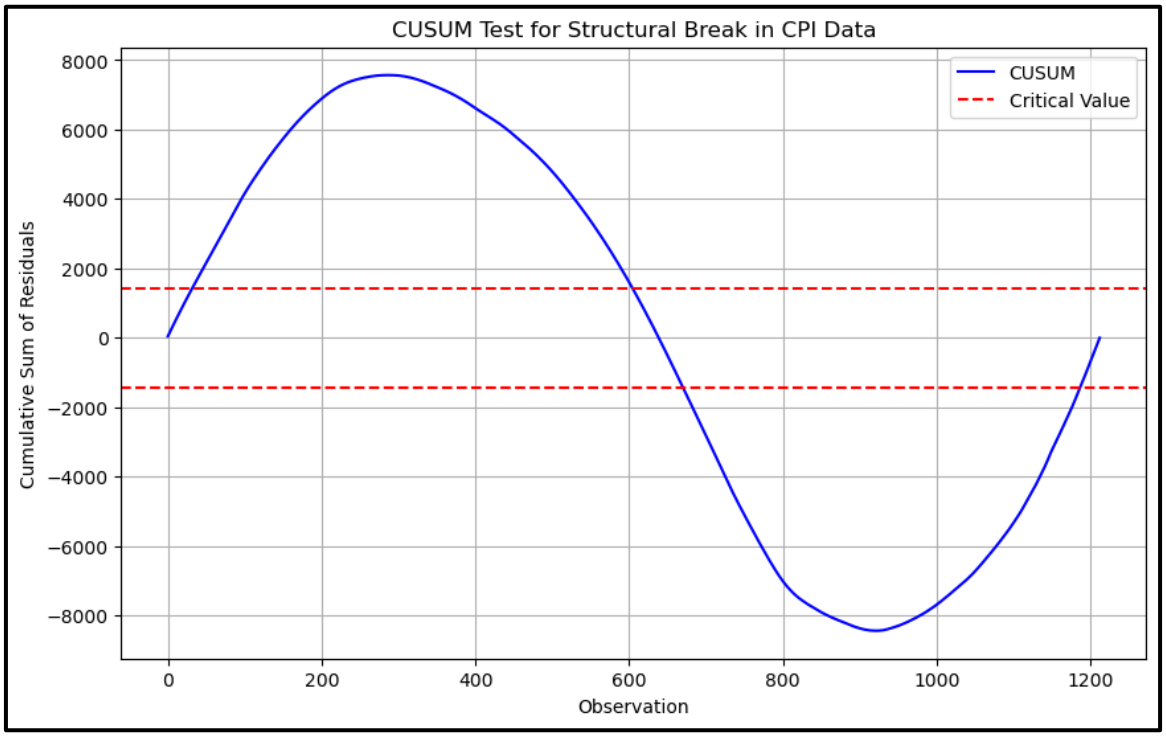
Şekil 3.10. Türkiye TÜFE CUSUM testi sonucu

Grafik ve çıktıdan görülebileceği üzere CUSUM testi maksimum sapmanın kritik değerin oldukça üstünde olduğunu tespit etmiştir. Bu da yapısal kırılma olduğu anlamına gelmektedir.

3.2.8. ABD TÜFE zaman serisine CUSUM testinin uygulanması

İlk olarak veri ön işlemesi gerçekleştirilmiştir. ABD TÜFE verisindeki ‘Date’ sütunundaki değerler pandas kütüphanesindeki ‘to_datetime’ fonksiyonu ile Python'un ‘datetime’ nesnelerine dönüştürülmüş, tarih formatı da yıl-ay-gün olarak belirlenmiştir.

Veriler işlendikten sonra ve teste hazır hale getirildikten sonra uygulanan CUSUM testi uygulanmış olup, sonuçları aşağıdaki grafik ve çıktıdan görülebilmektedir:



Şekil 3.11. ABD TÜFE CUSUM testi grafiği

```
Yapısal Kırılma Tespit Edildi: True
Maksimum Sapma: 8435.627357121377
Kritik Değer: 1425.5668888525947
```

Şekil 3.12. ABD TÜFE CUSUM testi sonucu

Buna göre, maksimum sapma değeri olan 8435,6’nın kritik değer olan 1425,6’dan oldukça yüksek olmasından dolayı yapısal kırılma olduğu sonucuna varılmıştır.

3.3. Kestirim Yöntemleri

Bu bölümde, 2 farklı TÜFE verisi için zaman serisi analizinde yaygın olarak kullanılan yöntemlerle zaman serisi tahmin performans analizi yapılarak sonuçları sunulacaktır. Bu yöntemler daha önce de belirtildiği gibi; ARIMA, Prophet, XGBoost ve LSTM yöntemlerdir. Bu dört yöntem tahmin doğruluğu performansı açısından karşılaştırılacak ve en doğru tahmini sağlayan yöntem belirlenecektir.

3.3.1. Türkiye TÜFE tahmin sonuçları

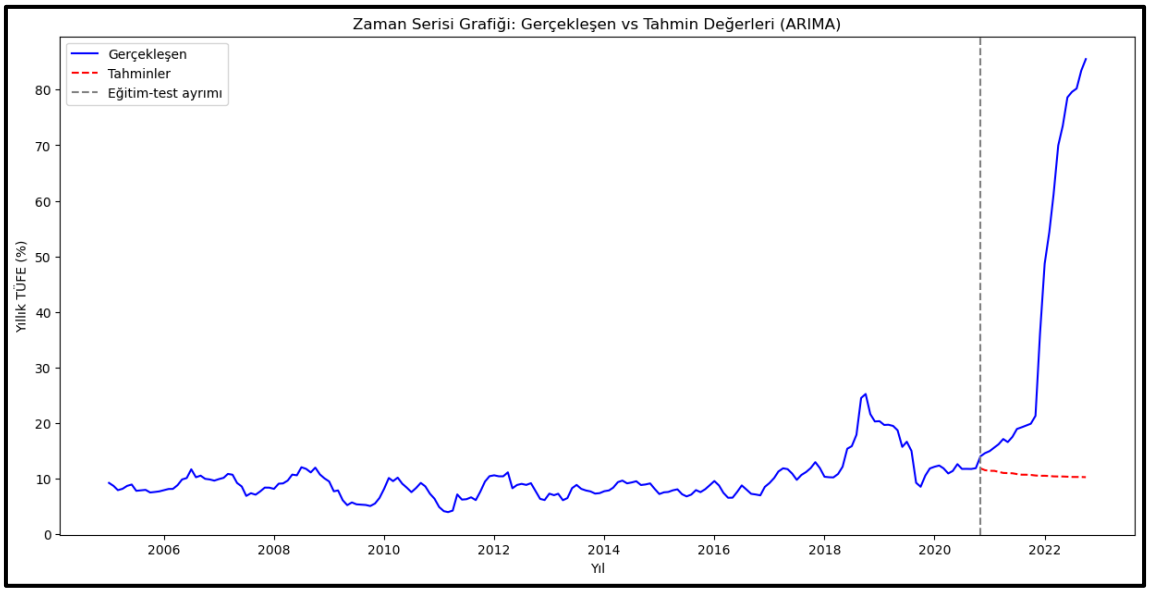
Daha önce yapılan yapısal kırılma tahmin testi sonucuna göre, kırılmanın en büyük olduğu tarih 01.11.2021'dir. Bu tarihi eğitim-test ayrım noktası kabul ederek dört ayrı yöntemin tahmin performansları kıyaslanmıştır.

ARIMA metodu ile Türkiye TÜFE tahmini

Bu bölümde, TÜFE'nin tahmini en bilinen istatistiksel tahmin yöntemlerinden biri olan ARIMA kullanılarak gerçekleştirilmiştir. Bu uygulama için Python programlama dili, Anaconda Jupyter Notebook üzerinden kullanılmıştır. Veriler okutulduktan sonra auto_arima fonksiyonuyla en optimum ARIMA (p, d, q) değerleri sırasıyla (3, 2, 3) olarak bulunmuştur. Bu değerler ve bu değerlerle elde edilen hata değerleri Python ortamında aşağıdaki gibi elde edilmiştir:

- Best ARIMA : (3, 2, 3) AIC=560.193
- MAE : 29.956838832816203
- RMSE : 40.89269794155683
- MAPE : 57.30486917252796
- sMAPE : 89.7401594149148

Bu hataların görsel olarak nasıl bir sapmaya yol açtığını gösteren grafik Şekil 3.13'te yer almaktadır.



Şekil 3.13. Türkiye TÜFE verisi için ARIMA ile tahmin ve gerçekleşme

Sapmadan da görülebileceği üzere, ARIMA'nın durağanlaştırma eğilimi bir miktar yükselmiş olan TÜFE'nin geri düşeceğini, böylece ortalamanın sabit kalacağını öngörmüştür. Türkiye TÜFE veri setinde ARIMA artışı öngörememiştir.

Prophet yöntemi ile Türkiye TÜFE tahmini

Facebook'un Veri Bilimi ekibi tarafından geliştirilen ve başlangıçta fbprophet adı altında indirilebilen Python paketi "prophet", bir tahmin aracıdır. Çoğunlukla bir istatistiksel yöntem olarak kabul edilir ve genellikle bir istatistiksel yöntem olarak sınıflandırılır, ancak gradyan güçlendirme ve çapraz doğrulama gibi makine öğrenmesi yöntemlerini içerdiği için hibrit bir model olarak da kabul edilebilir.

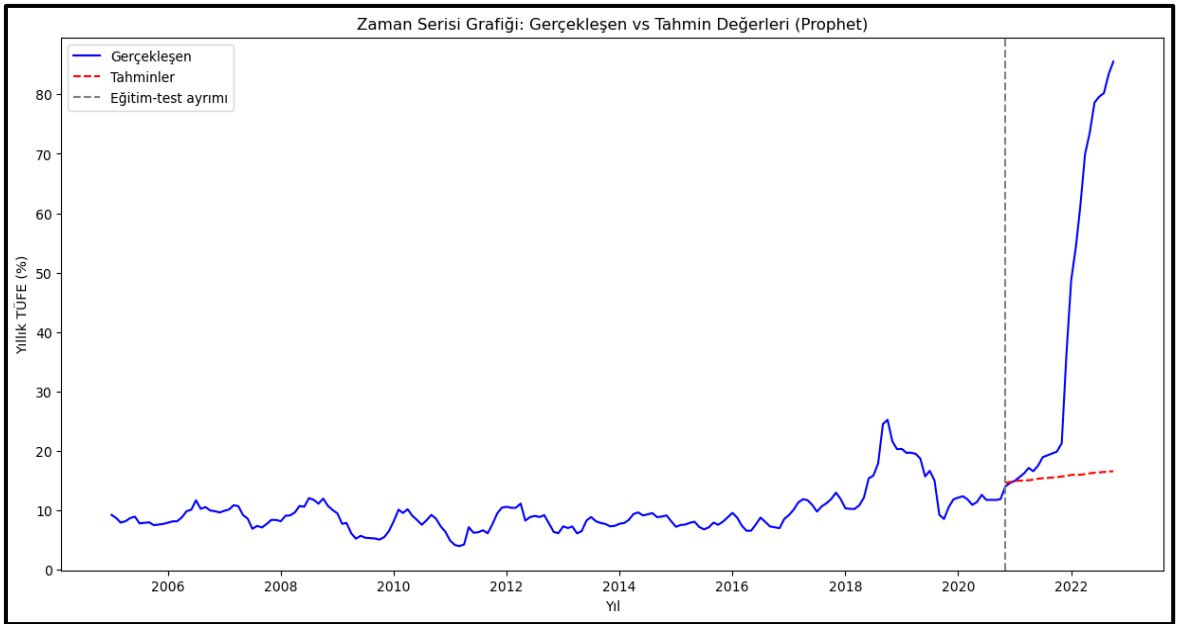
Prophet yöntemi de Türkiye TÜFE verisine Python ile uygulanmıştır. Bu yöntemde parametre değişikliklerinin ciddi bir hata, sapma azaltıcı etkisi genel olarak bulunmamaktadır.

Varsayılan ayarlarla Prophet yöntemi uygulandığında aşağıdaki çıktıyla hata miktarları Python'da elde edilmiştir:

Prophet Model Metrics:

- MAE : 25.098551092086637
- RMSE : 36.769929860234306
- MAPE : 40.78193959079354
- sMAPE : 62.26799834906871

Bu hatalara göre tahmin değerlerinin gerçek değerlerden sapması da Şekil 3.14'te görselleştirilmiştir:



Şekil 3.13. Türkiye TÜFE verisi için Prophet ile tahmin ve gerçekleşme

ARIMA'dan farklı olarak, yukarı yönlü trendi tahmin edebilmiş olmakla birlikte, bu trendin şiddetini tahmin etmede oldukça yetersiz kalmıştır.

XGBOOST Yöntemi ile Türkiye TÜFE Tahmini

Bir makine öğrenmesi yöntemi olan XGBoost'ta öncelikle en iyi parametrelerin tespit edilmesi için gridsearch yapılmıştır. Parametre grid'i aşağıdaki şekilde uygulanmıştır:

```
param_grid = {
    'n_estimators': [100, 200, 300],
    'learning_rate': [0.01, 0.05, 0.1],
    'max_depth': [3, 5, 7],
    'subsample': [0.8, 0.9, 1]
}
```

Bu grid'den en optimum parametre değerleri,

Fitting 3 folds for each of 81 candidates, totalling 243 fits

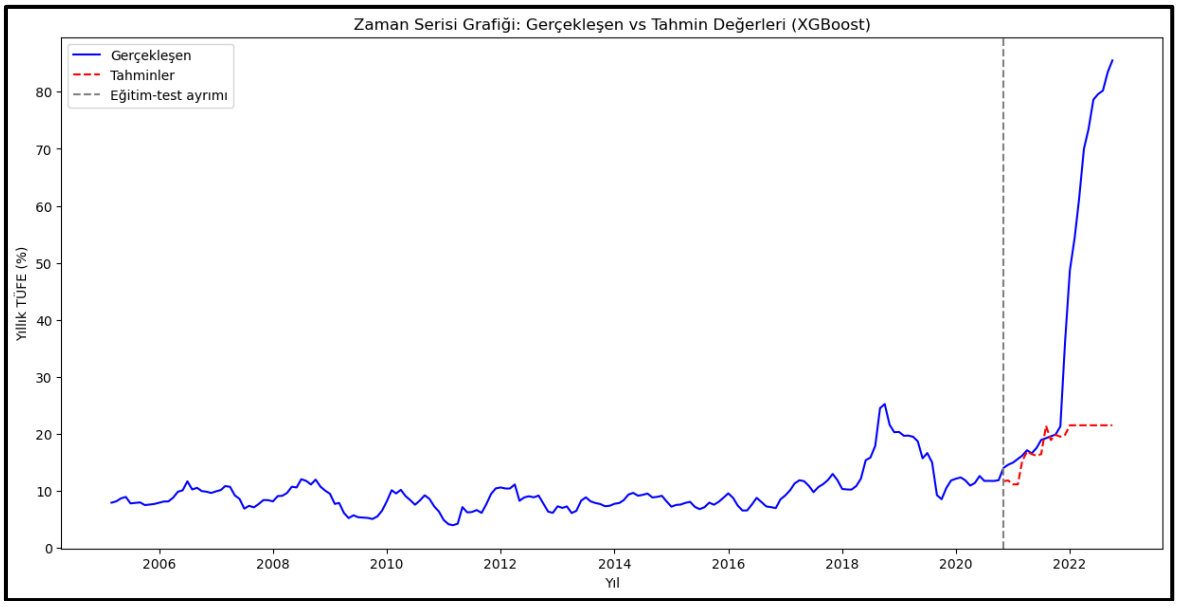
Best parameters: {'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 100, 'subsample': 1.0}

şeklindeki çıktıyla elde edilmiştir.

Bu parametrelere göre hatalar hesaplandığında aşağıdaki sonuçlar elde edilmiştir (Python çıktısı):

- MAE : 22.474854572614035
- RMSE : 33.40545594736417
- MAPE : 36.48245277395946
- sMAPE : 52.9241856658568

XGBoost yöntemi uygulandığında ortaya çıkan tahmin değerleri ile gerçek değerler arasındaki sapmanın görselleştirildiği grafik de Şekil 3.16'da yer almaktadır:



Şekil 3.14. Türkiye TÜFE verisi için XGBoost ile tahmin ve gerçekleşme

Bu makine öğrenmesi yönteminin tahmin doğruluğunun diğerlerinden daha yüksek olduğu hem hata sonuçlarından hem de grafikten gözlemlenebilmektedir. Ancak, artış olacağını doğru tahmin etmesine rağmen, bu artışın devam etmesi yerine duracağını tahmin ettiği için, beklenen sonucu verdiğini söylemek pek mümkün değildir.

LSTM yöntemi ile Türkiye TÜFE tahmini

Öncelikle, LSTM de bir makine öğrenmesi yöntemi olduğundan, aşağıdaki parametreler arasında en optimum olanını tespit etmek için aşağıdaki parametre grid'ine gridsearch uygulanmıştır.

```
param_grid = {
    'units': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100],
    'learning_rate': [0.001, 0.005, 0.01, 0.02, 0.03],
    'batch_size': [1, 5, 10, 20],
    'epochs': [50, 100, 150, 200],
    'look_back': [1, 2, 3, 4, 5]
}
```

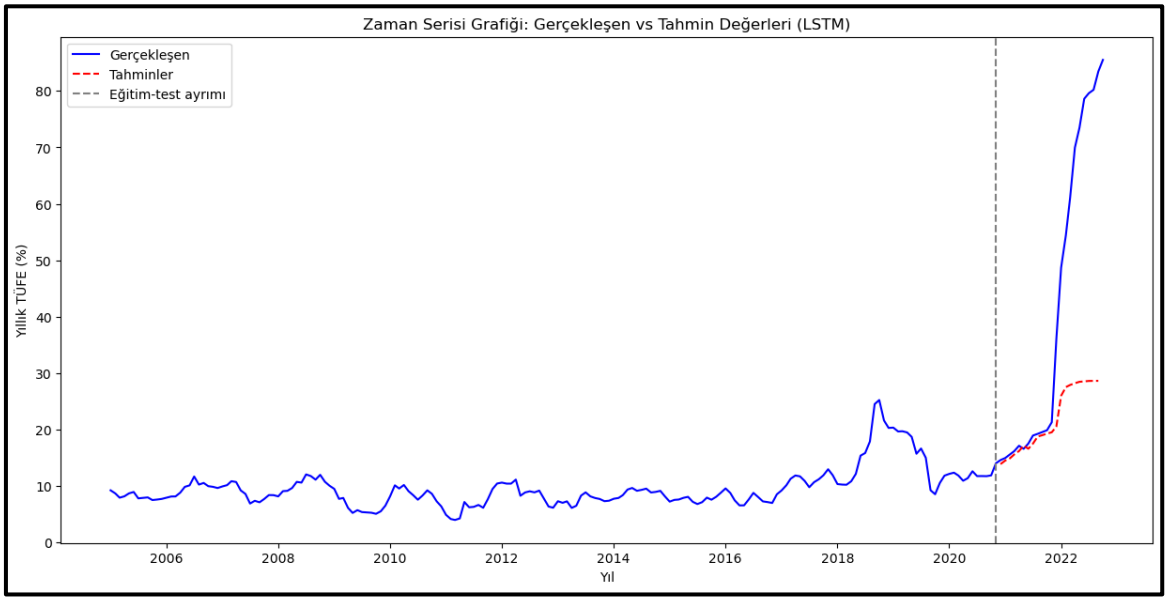
En iyi parametreler aşağıdaki şekilde tespit edilmiştir:

- units : 50,
- learning_rate : 0,01,
- batch_size : 1,
- epochs : 100,
- look_back : 1

En iyi parametrelere göre model tanımlanmış ve uygulanmıştır. Buna göre hatalar aşağıdaki gibi tespit edilmiştir:

- MAE : 18.51950907967307
- RMSE : 27.76235184856819
- MAPE : 30.124912148080817
- sMAPE : 41.020778202372746

LSTM yöntemine göre tahminin gerçekleşme durumundan sapmasını gösteren grafik Şekil 3.17'deki gibidir:



Şekil 3.15. Türkiye TÜFE verisi için LSTM ile tahmin ve gerçekleşme

Sonuçlar ve grafikten de görülebileceği üzere bu yöntem, XGBoost yöntemindeki gibi yukarı yönlü hareketi doğru tahmin etmekle beraber, bu hareketin ne kadar süreceğini doğru tahmin etmekte yetersiz kalmıştır. Esasen bütün istatistiksel ve makine öğrenmesi

yöntemlerinin temel olarak geçmiş verilerden yola çıkarak tahmin yapmasından dolayı, geçmişte hiç gerçekleşmemiş ölçüde bir yükselişi tahmin etme olasılığı yoktur.

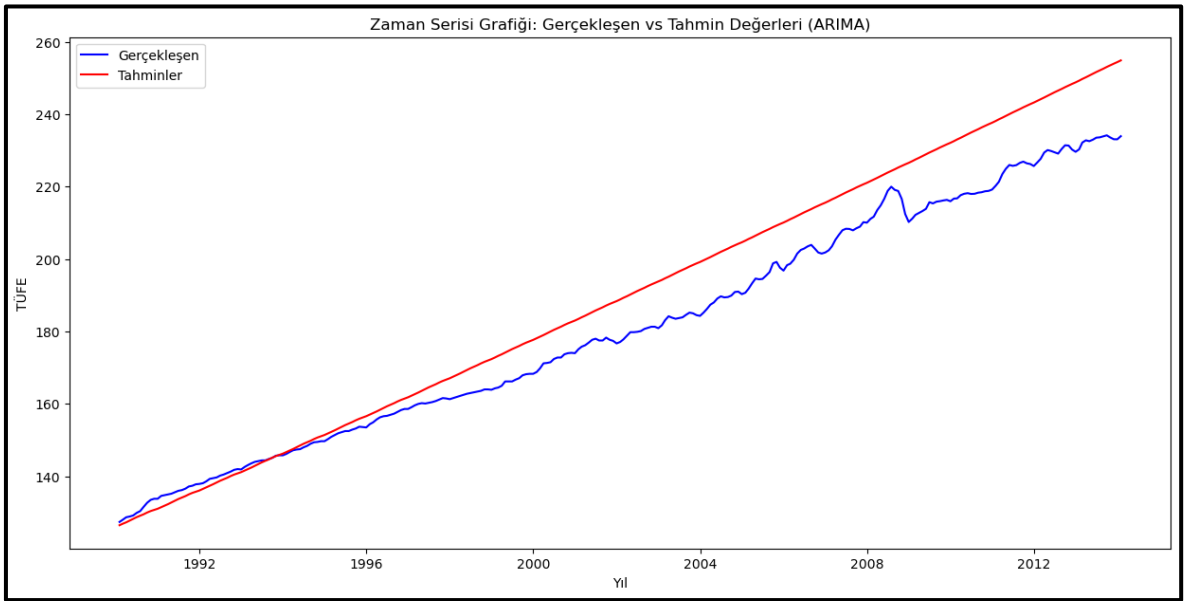
3.3.2. ABD TÜFE Tahmin Sonuçları

Daha önce yapılan yapısal kırılma tahmin testi sonucuna göre, kırılmanın en büyük olduğu tarih olan 01.01.1970 eğitim-test ayrım noktası kabul edilerek dört ayrı yöntemin tahmin performansları kıyaslanmıştır. Veri setinde 01.01.1913'ten 01.01.2014 tarihine kadar aylık bazda ABD TÜFE ve yıllık enflasyon oranları verilmiş olup, bu çalışmada yalnızca TÜFE değerleri kullanılmıştır.

ARIMA metodu ile ABD TÜFE tahmini

Bu bölümde, TÜFE'nin tahmini en bilinen istatistiksel tahmin yöntemlerinden biri olan ARIMA kullanılarak gerçekleştirilmiştir. Bu uygulama için Python programlama dili, Anaconda Jupyter Notebook üzerinden kullanılmıştır.

Şekil 3.18'de görselleştirilen hipotetik tahmin senaryosunda, 01.01.1990 tarihindeymişiz gibi bir tahmin yapılır.



Şekil 3.16. ABD TÜFE verisi için ARIMA ile tahmin ve gerçekleşme

Görüldüğü üzere, başarılı parametre ayarından sonra, ARIMA doğru bir tahmin üretmiştir. ARIMA için temel ön koşul durağanlıktır. Bu nedenle, parametre ayarları yapıldıktan sonra, zaman serisi durağanlaştırılmıştır. Çeşitli kombinasyonlar deney olarak test edilmiştir. Sonuç olarak, p, d ve q parametrelerinin optimal kombinasyonunu belirlemek için auto_arima fonksiyonu kullanılmıştır. Bu işlem sonucunda, ARIMA(0,2,3) en iyi model olarak elde edilmiştir.

Bu modele göre, bu çalışmada diğer yöntemlerle karşılaştırmak için hesaplanan metrikler;

ARIMA Metrikleri:

- MAE : 8,941975819897154
- RMSE : 10,66409670388658
- MAPE : 4,565985025643856
- sMAPE : 4,435629662149783

şeklindedir.

Prophet metodu ile ABD TÜFE tahmini

Prophet, bir "eller serbest" tahmin aracı olarak tasarlanmıştır ve varsayılanları çoğu zaman serisi veri kümesi için uygun olabilir. Prophet'ın basitliği, zaman serisi analizi uzmanlığına sahip olmayan kullanıcılar için kullanıcı dostu olmasını sağlayan güçlü yönlerinden biridir. Prophet'ı eğilim içeren bir veri kümesine uydurmak için yapılacak bir değişiklik, modelin “growth” argümanını “model = Prophet(growth='logistic')” olarak değiştirmekti. Ancak, bu değişiklik istenen sonuçları vermedi.

Parametre ayarı yapılmadan ve yapıldıktan sonra prophet modellerinin karşılaştırması için hesaplanan metrikler sırasıyla aşağıda verilmiştir:

Ayarlama öncesi Prophet metrikleri:

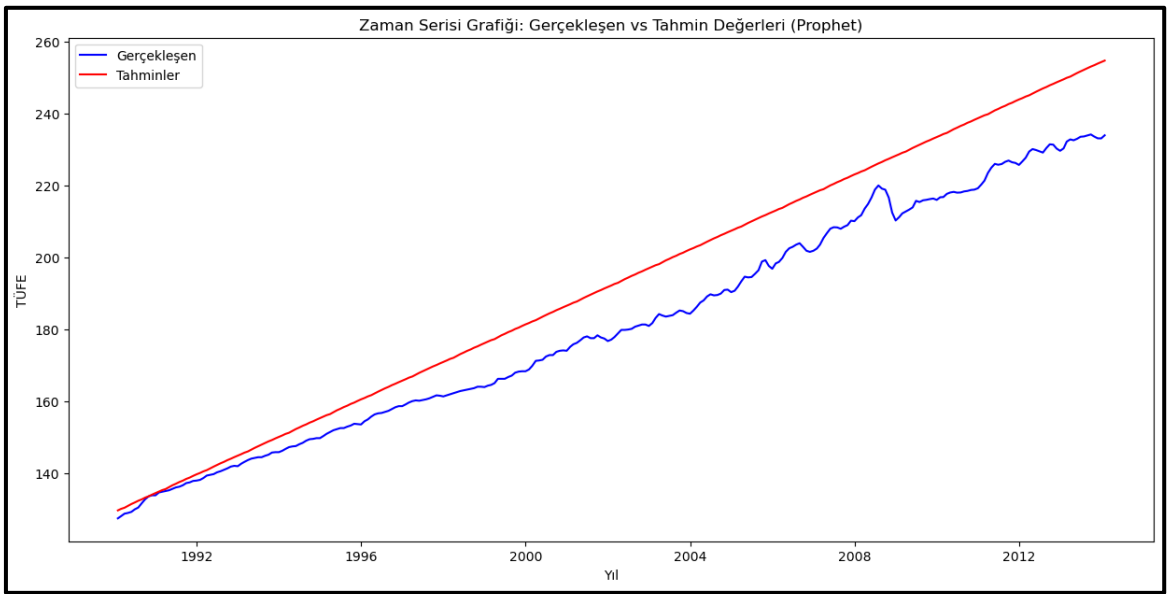
- MAE : 11,125084633707274
- RMSE : 12,488387093086484
- MAPE : 5,819991892307139
- sMAPE : 5,625485176238832

Ayarlama sonrası Prophet Metrikleri:

- MAE : 11,314882510742137
- RMSE : 12,670457753771267
- MAPE : 5,924970616349123
- sMAPE : 5,724601186164466

Görüldüğü gibi, ölçütler çok hafif değişmiş olup, değişiklikler göz ardı edilebilir derecededir. Hatta, hata ölçütlerinde hafif bir artış gözlemlenmektedir. Bu nedenle, parametre ayarlamanın bu veri kümesi için Prophet yöntemi üzerinde istenen etkiye sahip olmadığı varsayılabilir.

Gerçek değerlerle tahmin edilen değerleri karşılaştırmak için oluşturulan grafik Şekil 3.19'da verilmiştir:



Şekil 3.17. ABD TÜFE verisi için Prophet ile tahmin ve gerçekleşme

XGBoost ile ABD TÜFE tahmini

Yarışmalarda en iyi performans gösteren tahmin yöntemlerinden biri olan XGBoost, ABD TÜFE değerlerini tahmin etmek ve gerçek değerlerle karşılaştırmak için kullanıldı.

İlk olarak, xgboost paketi Python'da Jupyter Notebook'ta yüklendi ve zaman serisi tahmini gerçekleştirildi. Ancak Python'da, tahmin değerleri sabit kaldı ve model 1990-01-01'den sonra artan eğilimi yakalayamadı. Daha sonra R, RStudio üzerinden kullanıldı.

Aşağıdaki paketler yüklendi:

- library(xgboost)
- library(caret) #Makine öğrenimi görevleri
- library(lubridate) #Tarih ve zaman serilerini işlemek için paket
- library(ggplot2)
- library(lattice) #Grafik, grafikler

Veri kümesi yüklendi, eğitim ve test değerlerini bölmek için kesim tarihi olarak 1990-01-01 belirlendi. Model eğitildi, en iyi parametreler GridSeach ile bulundu:

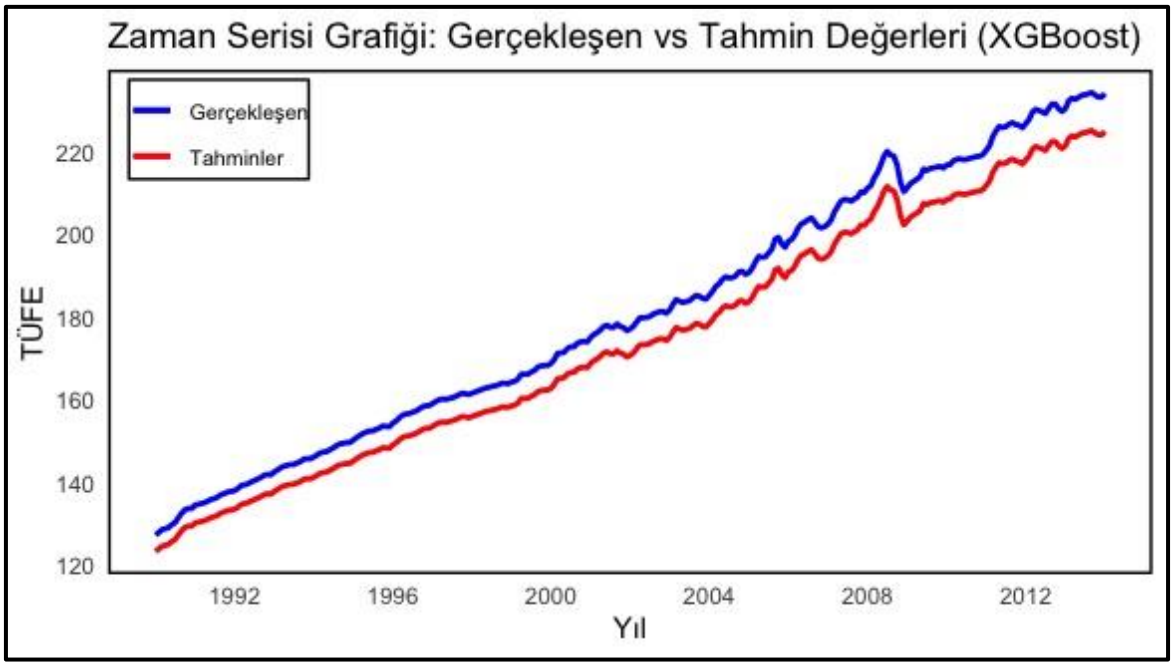
Eta (learning_rate)=0.01, max_depth=7, subsample=0.8, nrounds (n_estimators)=1000

Bu parametrelere göre sonuçlar aşağıdadır:

XGBoost Metrikleri:

- MAE : 6,571802
- RMSE : 6,746269
- MAPE : 3,591228
- sMAPE : 3,657163

Tahmin edilen değerlerin gerçek değerlerden sapmasını gösteren grafik Şekil 3.20’ de yer almaktadır.



Şekil 3.18. ABD TÜFE verisi için XGBoost ile tahmin ve gerçekleşme (R çıktısı)

LSTM ile ABD TÜFE tahmini

LSTM, uzun vadeli ve karmaşık veri ve desenlerle başa çıkma yeteneği nedeniyle akademi ve pratikte en yaygın kullanılan zaman serisi tahmin yöntemlerinden biridir. LSTM'in hem uzun hem de kısa vadeli verileri hesaba katma özelliği onu ayırt edici kılar. Başlangıçta parametre ayarlaması yapılmamış, en yaygın parametreler kullanılmıştır. Daha sonra modelin MAE, RMSE, MAPE ve sMAPE değerleri hesaplanmıştır. Daha sonra grid arama ve çapraz doğrulama uygulandı ve bazı manuel optimizasyonlar gerçekleştirilmiştir. Sonuç olarak önemli ölçüde daha iyi sonuçların elde edildiği görülmüştür. Ayar yapılmadan ve yapıldıktan sonra sonuçlar aşağıdadır:

Ayar yapılmadan önce:

Geriye Dönük: 12 ay (Mevsimsellik nedeniyle standart)

LSTM Metrikleri:

- MAE : 9,422465322776821
- RMSE : 12,54806989386087
- MAPE : 22,819912326387104
- sMAPE : 21,187897498168038

GridSearch ve Manuel Ayarlama Sonrası:

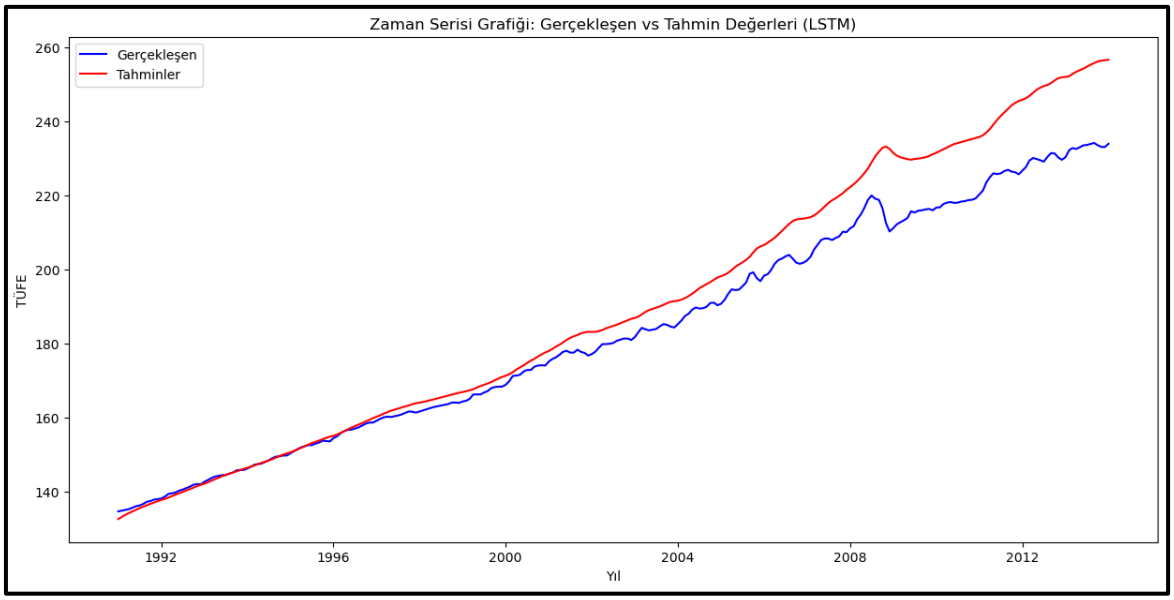
En iyi sonuçları veren parametreler şöyledir:

- Geriye Dönük: 240 ay (manuel olarak belirlendi)
- Yoğun katman sayısı: 1
- LSTM birim sayısı: 50
- LSTM katmanı için aktivasyon fonksiyonu: 'relu' (Düzeltilmiş Doğrusal Birim)
- LSTM katmanı için giriş şekli: (X_train.shape[1], 1)
- Eğitim için kullanılan optimizer: 'adam'
- Kayıp fonksiyonu: 'mse' (Ortalama Kare Hata)
- Eğitim için dönem sayısı: 50
- Eğitim için parti boyutu: 16

Geriye dönük değer hariç diğer parametreler GridSearch ile elde edildi. En iyi geriye dönük değer olan 240 ay manuel olarak belirlendi. Diğer parametrelerin sonuçlar üzerinde önemli bir etkisi olmadığı için, daha iyi doğruluk sonuçları elde etmede geriye dönük değer anahtar rol oynadı.

LSTM Metrikleri:

- MAE : 1,2456506447305475
- RMSE : 1,5764535015543961
- MAPE : 3,067706945487548
- sMAPE : 3,055359659029349



Şekil 3.19. ABD TÜFE verisi için LSTM ile tahmin ve gerçekleşme

LSTM'de en kritik ve ayırt edici parametre değerinin geri dönüş (look_back) parametresi olduğu görülmüştür. Diğer bir deyişle bu parametre değerinin, diğerlerine göre kestirim sonuçları üzerinde etkisinin daha büyük olduğu gözlemlenmiştir. Daha uzun geri dönüş değerleri, zamansal örüntüleri daha iyi yakalamaya yardımcı olur. Ancak, yapılan çalışmalar göstermiştir ki, çok uzun bir geri dönüş değeri, artık etkisi olmayan, yanıltıcı veya gelecekteki değerler üzerinde bir etkisi bulunmayan, verideki alakasız örüntüleri yakalayabilir. Bunun sonucunda, aşırı karmaşık modeller oluşabilir ve bu nedenle aşırı uyum (over-fitting) problemi ortaya çıkabilir. Yaklaşık 100 yıllık veri ile 240 aydan fazla, örneğin 480 ay (40 yıl) geriye dönüş değeri, tüm dönemin ilk 40 yılı için uzun vadeli bir hafıza oluşturulması sonucunda aşırı uyum problemini ortaya çıkarırken, 240 aydan azı ise, LSTM'nin uzun vadeli bellek yeteneklerini sınırlayabilir. Bunun sonucunda optimum geri dönüş süresi 240 ay (20 yıl) olarak belirlenmiştir.

3.4. Yöntemlerin Karşılaştırılması

Türkiye ve ABD TÜFE verileri üzerinde uygulanan tahmin yöntemlerinin bir arada verilip, karşılaştırılmasına imkân tanımak için Çizelge 3.1. ve Çizelge 3.2. oluşturulmuştur. Satırlarda kullanılan tahmin yöntemlerinin isimleri, sütunlarda ise bu yöntemlerin uygulanması sonucu elde edilen performans kriterleri yer almaktadır.

Çizelge 3.1. Türkiye Tüfe tahmin yöntemlerinin sonuçları

Yöntemler / Performans Ölçütleri	MAE	RMSE	MAPE	sMAPE
ARIMA	29,96	40,89	57,3	89,74
Prophet	25,1	36,77	40,78	62,27
LSTM	18,52	27,76	30,12	41,02
XGBoost	22,47	33,41	36,48	52,92

Çizelge 3.2. ABD Tüfe tahmin yöntemlerinin sonuçları

Yöntemler / Performans Ölçütleri	MAE	RMSE	MAPE	sMAPE
ARIMA	8,942	10,664	4,566	4,436
Prophet	11,315	12,670	5,925	5,725
LSTM	1,245	1,576	3,068	3,055
XGBoost	6,572	6,746	3,591	3,657

Yukarıdaki tablolar, bu çalışmada kullanılan dört farklı yöntemin tahmin doğruluk performanslarını, zaman serisi tahminlerini değerlendirmek için yaygın olarak kullanılan dört performans kriteriyle değerlendirildiği şekilde göstermektedir. Bu metriklerdeki daha düşük değerler, gerçek verilere kıyasla değerleri tahmin etmede daha yüksek bir doğruluk derecesi gösterir. Tablolara göre iki zaman serisinde de LSTM yöntemi diğerlerine göre daha başarılı bir tahmin performansı sergilemiş olup, iki seride de makine öğrenmesi yöntemleri istatistiksel yöntemlere göre daha doğru tahminler gerçekleştirmişlerdir. Türkiye TÜFE verisinde Prophet yöntemi ARIMA'ya göre daha gerçeğe yakın tahmin gerçekleştirirken, ABD TÜFE verisinde ARIMA Prophet'e göre daha gerçeğe yakın bir tahmin gerçekleştirmiştir.

Sonuç olarak, ele alınan iki zaman serisinde de LSTM diğer yöntemlere göre tüm metriklerde daha iyi performans göstermiştir. LSTM'deki hatalar en düşüktür. Bu da LSTM'in diğer yöntemlere kıyasla tahminlerin gerçek değerlerden sapmasının en küçük olduğunu gösterir. Bu üstün performans, LSTM'nin karmaşıklığına ve hem uzun hem de kısa vadeli verilerle başa çıkma yeteneğine atfedilebilir.

Prophet, daha ziyade mevsimsel dalgalanmaları ve tatil etkisini ele almak üzere tasarlanmış bir yöntemdir. Ancak, bu çalışmada ele alınan zaman serilerinden Türkiye TÜFE zaman

serisinde yapısal kırılma ve ABD TÜFE verisinde trend ve yapısal kırılma ağır basmaktadır. Mevsimsellikle açıklanabilecek seyir değişiklikleri yok denebilecek kadar azdır.

Öte yandan, ARIMA ortalama ve varyansta büyük değişimlerin yaşanmadığı nispeten durağan zaman serilerini tahmin etmede yeteneklidir. Ancak, ABD TÜFE veri setinde ağırlıklı olarak yukarı yönlü bir eğilim olduğu göz önüne alındığında, ortalama sürekli artmaktadır. Farklar alınarak seriyi durağanlaştırma çabalarına rağmen, ABD TÜFE verisindeki 1970'lerdeki yapısal bir kırılma, durağanlaştırmayı daha zor hale getirmiştir.

ABD TÜFE zaman serisinde makine öğrenmesi yöntemleri olan LSTM ve XGBoost, dört metrikte de ARIMA ve Prophet'e göre daha doğru sonuçlar vermiştir. LSTM ise XGBoost'tan daha doğru bir tahmin performansı sergilemiştir. XGBoost, zaman serisi tahmininde etkili bir performans sergilemesine rağmen, orijinal tasarımı sınıflandırma problemlerinde tahmin yapmak içindir. XGBoost, LSTM'ye kıyasla daha az bilgisayar kaynağı ve zaman kullanırken hesaplama verimliliği sergilese de, veri kümesinin karmaşıklığı arttıkça tahmin üstünlüğü LSTM'e oranla azalır. Bu, LSTM'nin karmaşık modelleme yeteneklerinden kaynaklanmaktadır.

ABD TÜFE verisinde neredeyse 100 yıl boyunca aylık verilerin kapsamlı bir şekilde toplanması ve evrimleşen desenler göz önüne alındığında, LSTM'nin hem kısa hem de uzun vadeli bağımlılıklarla başa çıkma yeteneğinin bu çalışmada kullanılan veri kümesi için diğer modellere göre üstün olduğu varsayılabilir. Bu düşünce, 1970'lerdeki yapısal kırılmanın ve eğilimin baskın olmasını göz önünde bulundurur.

Türkiye TÜFE verisinde de LSTM diğer yöntemlere göre ABD TÜFE verisindeki kadar olmasa da yine tüm metriklerde daha iyi performans göstermiştir. LSTM'nin karmaşıklığı ve hem uzun hem de kısa vadeli verilerle başa çıkma yeteneğinin yaklaşık bir asırlık ABD TÜFE verisinde yarattığı fark, Türkiye TÜFE zaman serisinin daha kısa olmasından dolayı kısıtlı kalmıştır.

Türkiye TÜFE verisinin ARIMA ile tahmin performansı değerlendirildiğinde, tüm durağanlaştırmalara rağmen, 2020 yılındaki ani yapısal kırılma sonucu durağanlık bozulmuş, ortalama ciddi miktarda artmıştır. Alınan fark, yukarı yönlü trendi tahmin etmekte yetersiz kalmıştır.

Prophet’in karşılaştırmalı üstünlüğü olarak ifade edilebilen mevsimsel dalgalanmaları ve tatil etkisini ele alma yeteneği, Türkiye TÜFE verisinde mevsimsellik etkisi bulunmadığından Prophet’in doğru tahmin yapmasına yardımcı olamamıştır. Yine de ARIMA’nın aksine yukarı yönlü hareketi doğru tahmin etmiştir.

Bir makine öğrenmesi yöntemi olarak XGBoost, dört metrikte de ARIMA ve Prophet’e göre daha doğru, LSTM’e göre daha daha düşük performans göstermiştir.

LSTM’in hem kısa, hem uzun dönem verileri tahminde kullanabilme yeteneği, Türkiye TÜFE verisinde LSTM’in diğer yöntemlere göre daha doğru tahminler yapmasını sağlamıştır.

4. SONUÇ VE ÖNERİLER

Küresel Covid19 salgını sonrası fiyat endekslerinde yaşanan büyük artışların ne ölçüde öngörülebilir olup olmadığı, hangi para birimi üzerinden ve hangi ülkede tahmin ve gerçekleşmenin kıyaslandığına göre büyük değişiklikler göstermektedir.

Çalışmamızda yaklaşık 18 yıllık Türkiye TÜFE verisi üzerinden bu çalışma yapıldığında, özellikle 2020'ye kadar TÜFE'de zaman zaman dalgalanmalar olsa da, genel olarak durağan bir seyir izlediği, ancak 2020'de hem Covid19, hem de Türk Lirasında devalüasyondan kaynaklı aşırı bir artış olduğu gözlenebilmektedir. Bu da tahmin yöntemlerinin doğruluk göstergelerine bakıldığında, ekonomik aktörlerin öngöremeyeceği bir durum gibi görülebilmektedir. Ekonomide yaşanan ani değişiklikler, fiyat endeksleri ile ilgili tahmin yapmayı zorlaştırmış, mevcut tahmin yöntemlerinin etkinliğini olumsuz yönde etkilemiştir.

Özellikle Covid19 salgını ile küresel çapta başlayan istisnai yüksek enflasyonist ortam sonucunda, uzun süreli sözleşme akdeden ekonomik aktörler arasında hukuki ihtilaflar doğmaktadır. Bu ihtilafların giderilebilmesi için istatistik ve benzeri alanlarda akademisyenlerden bilirkişi desteği alınmaktadır. Dolayısıyla, fiyat endekslerindeki ani artışların öngörülebilir olup olmadığının tespiti, bu ihtilaflar sonucu yapılan hukuki başvuruların sonuçlandırılmasında büyük önem arz etmektedir.

Bu noktada, fiyat endekslerindeki öngörülemez büyük artışların yapısal kırılma olduğu göz önüne alınarak, yapısal kırılmalı fiyat endeksi zaman serilerinde geleceğin öngörülebilir olup olmadığının istatistiksel ve makine öğrenmesi yöntemleriyle tespiti, ihtilafların çözülmesinde ciddi rol oynayabilecektir. Bu amaçla da, söz konusu zaman serisindeki yapısal kırılmayı hesaba katarak, en iyi kestirimi yapmayı sağlayan istatistiksel ya da makine öğrenmesi yönteminin tespiti de kilit bir öneme sahiptir. Kamu - özel sektör - akademi işbirliği bu aşamada büyük önem arz etmektedir.

Yapısal kırılmalar, kısa vadede değil, uzun vadede olması beklenen olaylardan kaynaklanmaktadır. Dolayısıyla, bu çalışmada yer alan iki zaman serisinden, yaklaşık 18 yıllık bir zaman serisi olan Türkiye TÜFE verisinin içinde geçmişte olan bir yapısal kırılmayı tespit edebilme ihtimali, yaklaşık 100 yıllık ABD TÜFE zaman serisindekine göre çok daha azdır. Yani, örneğin eğer bir proje yüklenicisi, 2020 yılında ya da önce bir proje

için ihale teklifi verirken Şekil.1'den görülebileceği üzere, fiyat seviyelerinin artmayacağını öngörebilir, uzun süreli bir sözleşme ise sözleşme bitimine kadar mevcut fiyat seviyelerini gözeterek fiyat teklifi verir. Bir yapısal kırılma ve buna bağlı sıra dışı bir fiyat artışı yaşandığında da, bunun öngörülemez olduğunu ileri sürerek, proje sahibi kurumdan ilave finansman talep eder. Ancak, zaman serilerinde tahmin yaparken, eğer olabildiğince geriye giderek analiz gerçekleştirilirse, bu analizde, geçmişte meydana gelmiş bir yapısal kırılmanın da tespit edilip, analize dahil edilmesi daha mümkün olacaktır. Türk Ticaret Kanununda "basiretli tacir" olarak ifade edilen şartı sağlamak isteyen bir ekonomik aktör, yalnızca yakın geçmişi değil, olabildiğince uzak geçmişi de ve olası yapısal kırılmaları da tahminlerine dahil etmelidir.

Her ne kadar bu ihtilaflarda kimi ülkelerde Covid19 sonrası kamu otoriteleri fiyat artışlarından kaynaklı maliyet artışı yükünün çoğunu karşılasa ve hatta bazı ülkelerde arabulucu uluslararası kurumlar bu yükün tamamen proje sahibi tarafından karşılanması gerektiği yönünde hükümler tesis etseler de; dolar gibi rezerv bir para biriminde dahi yeterince geçmişe gidildiğinde yapısal kırılmaların varlığı, zaman içerisinde değer kaybettiği görülebilecektir. Dolayısıyla, basiretli bir tacirin öngörebileceği risklerin maliyeti vatandaşlara yüklenmemelidir.

Bu bağlamda, Covid19 sonrası yaşanan yapısal kırılma sonrası yaşanan ihtilaflardan alınan dersle, halkın vergileriyle finanse edilen projelerde, vatandaşların üzerine ilave mali yük binmemesi açısından, özellikle proje şartnamelerine riskin tamamen kamu otoritesi tarafından üstlenilmeyeceği, fiyat teklifi verirken, projenin süresi ve olabildiğince eskiye dönük bütün yapısal kırılmalar dikkate alınarak ihalelerde fiyat teklifi verilmesi, ileride yaşanabilecek ihtilafların çözümünde faydalı olacaktır.

Her ne kadar farklı tahmin yöntemleriyle zaman serisi tahminleri gerçekleştirilmeli ise de, bu çalışmada yer verilen her iki veri setinde de en gerçeğe yakın tahminleri gerçekleştiren LSTM yöntemi uygulanmalıdır. Tek dezavantajı diğer yöntemlere göre daha yoğun bilgisayar kaynağı ve süresi kullanması olan bu yöntemin hem uzun, hem kısa dönem geçmiş verileri eş zamanlı ele alabilen karmaşık algoritması bu yönteme ciddi bir rekabet avantajı sağlamaktadır. LSTM'e de alternatif olacak ve geliştirilmekte olan Kapalı Tekrarlayan Birim (Gated Recurrent Unit), Evrişimli Sinir Ağları (CNN), Dönüştürücü Modeller (Transformer Models), Zamansal Evrişimli Ağ (Temporal Convolutional Network),

DeepAR, Pencereleme ile Beslemeli İleri Sinir Ağları gibi derin öğrenme yöntemlerinin de daha fazla geliştirilmesi zaman serisi tahminlerinin daha da doğru yapılabilmesine olanak sağlayabilecektir.

KAYNAKLAR

Aliyev, P. (2024). İşsizlik oranı öngörülerinde makine öğrenimi yaklaşımları: türkiye üzerine bir uygulama. *Igdir University Journal of Faculty of Economics and Administrative Sciences*, 11, 1–14.

Almazrouee, A. I., Almeshal, A. M., Almutairi, A. S., Alenezi, M. R. and Alhajeri, S. N. (2020). Long-term forecasting of electrical loads in kuwait using prophet and holt–winters models. *Applied Sciences*, 10(16), 5627.

Andrews, D. W. K. and Fair, R. C. (1988). Inference in nonlinear econometric models with structural change. *The Review of Economic Studies*, 55(4), 615-640

Aqbar, K. and Supomo, R. A. (2023). Performance analysis of LSTM and XGBoost models optimization in forecasting crude palm oil (CPO) production at palm oil mill (POM). *International Journal of Computer Applications*, 185(17), 37–44.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22.

Baraldi, L., Grana, C. and Cucchiara, R. (2017). Hierarchical boundary-aware neural encoder for video captioning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1657-1666.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

Box, G. and Jenkins, G. (1976). *Time series analysis: Forecasting.* New York: Holden-Day, 41-56.

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. New York: John Wiley & Sons, 69-92.

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to time series and forecasting* New York: Springer, 145-171.

Brown, R. L., Durbin, J. and Evans, J. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 37(2), 149-163.

Cantrell, R. S., Burrows, P. M. and Vuong, Q. H. (1991). Interpretation and use of generalized Chow tests. *International Economic Review*, 4, 725-741.

Chatfield, C. (2003). *The analysis of time series*. Boca Raton: Chapman and Hall/CRC, 75-91.

Chen, T., & Guestrin, C. (2016, August 13). *XGBoost.* Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/2939672.2939785

Chihwa, K. and Ross, S. L. (1995). A Cusum test in the linear regression model with serially correlated disturbances. *Econometric Reviews*, 14(3), 331-346.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591.

Ditzen, J., Karavias, Y. and Westerlund, J. (2021). *Testing and Estimating Structural Breaks in Time Series and Panel Data in Stata*.

Drebee, H. A., Razak, N. A. A. and Mohsen, A. A. (2023). Maize production forecasting in Iraq: A box-jenkins approach for the period of 2022-2026. *IOP Conference Series: Earth and Environmental Science*, 1259(1), 012128.

Dufour, J.-M. (1982). An application of the chow test for parameter constancy to structural change models. *The Review of Economics and Statistics*, 12(3), 45-61.

Dufour, J.-M., Ghysels, E. and Hall, A. (1994). Generalized predictive tests and structural change analysis in econometrics. *International Economic Review*, 35(1). 199-229.

Gers, F. A., Schmidhuber, J. and Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471.

Gür, Y. E. (2024). Consumer price ındex forecasting in Turkey: A comparison of deep learning and machine learning approaches. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, 36, 312–344.

Hansen, B. E. (1997). Approximate asymptotic p values for structuras-change tests. *Journal of Business & Economic Statistics*, 15(1), 60-67.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.

İnternet: Bhattiprolu, S. (2021). 193 - What is XGBoost and is it Really Better than Random Forest and Deep Learning? In YouTube video. URL: https://www.youtube.com/watch?v= XXHhrlL-FWc&t=19s, Son Erişim Tarihi, 14.03.2024.

İnternet: Brownlee, J. (2018). A Gentle Introduction to SARIMA for Time Series Forecasting in Python. URL: https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/, Son Erişim Tarihi, 14.03.2024.

İnternet: Data Aspirant. (2024). Ultimate Guide for ARIMA Time Series Forecasting. URL: https://dataaspirant.com/arima-time-series-forecasting/, Son Erişim Tarihi, 14.03.2024.

İnternet: Facebook Core Data Science Team. (2017). Facebook/prophet: Tool for Producing High Quality Forecasts for Time Series Data that has Multiple Seasonality with Linear or Non-Linear Growth. URL: https://github.com/facebook/prophet, Son Erişim Tarihi, 14.03.2024.

İnternet: Krieger, M. (2024). Time Series Analysis with Facebook Prophet: How it Works and How to use it. URL: https://towardsdatascience.com/time-series-analysis-with-facebook-prophet-how-it-works-and-how-to-use-it-f15ecf2c0e3a, Son Erişim Tarihi, 14.03.2024.

İnternet: StataCorp LLC. (2024). Structural Breaks - Features Overview. URL: https://www.stata.com/features/overview/structural-breaks/, Son Erişim Tarihi, 14.03.2024.

İnternet: US Bureau of Labor Statistics. (2023). Consumer Price Index. URL: https://www.bls.gov/cpi/, Son Erişim Tarihi: 04.05.2024.

İnternet: Wijaya, C. Y. (2023). Leveraging XGBoost for Time-Series Forecasting. URL: https://www.kdnuggets.com/2023/08/ leveraging-xgboost-timeseries-forecasting.html, Son Erişim Tarihi, 14.03.2024.

İnternet: XGBoost Documentation — xgboost 1.7.5 documentation. (n.d.). URL: https://xgboost.readthedocs.io/en/stable/, Son Erişim Tarihi, 09.06.2024.

İnternet: Yu, F. (2024). A Thorough Guide to Time Series Analysis. URL: https://towardsdatascience.com/a-thorough-guide-to-time-series-analysis-5439c63bc9c5, Son Erişim Tarihi, 14.03.2024.

Kai, Y., Zhiqiang, J., Shitao, C., Yuedong, Y., & Nanning, Z. (2019). An Integrated Localization System with Fault Detection, Isolation and Recovery for Autonomous Vehicles. *Knowledge-Based Recurrent Attentive Neural Network for Small Object Detection*. https://doi.org/10.1109/ITSC48978.2021.9565076

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45.

Karatekin, C. ve Başaran, T. (2022). Gün öncesi piyasasında elektrik enerjisi fiyatının veri analizi ile tahmin edilmesi. *Journal of the Institute of Science and Technology*, 12(4), 2075–2084.

Kawakita, M., Minami, M., Eguchi, S. and Lennert-Cody, C. E. (2005). An introduction to the predictive technique AdaBoost with a comparison to generalized additive models. *Fisheries Research*, 76(3), 328–343.

Kim, H. J. and Siegmund, D. (1989). The likelihood ratio test for a change-point in simple linear regression. *Biometrika*, 76(3), 409-423.

Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I. and Matsopoulos, G. K. (2023). A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. *Future Internet*, 15(8), 255.

Krämer, W., Ploberger, W. and Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica: Journal of the Econometric Society*, 4, 1355-1369.

Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the Institute of Electrical and Electronics Engineers*, 86(11), 2278–2324.

Lo, A. W. and Newey, W. K. (1985). A large-sample Chow test for the linear simultaneous equation. *Economics Letters*, 18(4), 351-353.

Nanthini, K., Pyingkodi, M., Sivabalaselvamani, D., Kumari, S. and Kumar, T. (2022). Performance analysis of machine learning algorithms in heart diseases prediction. In *IoT Based Control Networks and Intelligent Systems: Proceedings of 3rd ICICNIS 2022* (pp. 407-423). Singapore: Springer Nature Singapore.

Nguyen, T.-T., Nguyen, H.-G., Lee, J.-Y., Wang, Y.-L. and Tsai, C.-S. (2023). The consumer price index prediction using machine learning approaches: Evidence from the United States. *Heliyon*, 9(10), e20730.

Nielsen, A. (2019). Practical time series: Prediction with statistics & machine learning. *O'Reilly*, 21(2), 219–221.

Oukhouya, H. and El Himdi, K. (2023, April). Comparing machine learning methods—svr, xgboost, lstm, and mlp—for forecasting the moroccan stock market. In *Computer Sciences & Mathematics Forum* (Vol. 7, No. 1, p. 39). MDPI.

Özdemir, O. (2020). *Performance Comparison of Machine Learning Methods and Traditional Time Series Methods for Forecasting*. Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi Doğal ve Uygulamalı Bilimler Enstitüsü, Ankara, 7-24.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American statistical Association*, 55(290), 324-330.

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), 1–39.

Sibai, F., El-Moursy, A. and Sibai, A. (2024). Forecasting the consumer price ındex: A comparative study of machine learning methods. *International Journal of Computing and Digital Systems*, 15, 2142–2210.

Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica: Journal of the Econometric Society*, 4(1), 105-146.

Taylor, S. J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45.

TCMB. (2022, November). Türkiye Cumhuriyet Merkez Bankası [TCMB] Tüketici Fiyatları.URL: https://www.tcmb.gov.tr/wps/wcm/connect/TR/TCMB+TR/main+menu/istatistikler/enflasyon+verileri/tuketici+fiyatlari, Son Erişim Tarihi: 14.03.2024.

Tong, H. (1983). *Threshold models in non-linear time series analysis*. Berlin: Springer Science & Business Media, 21.

Türkiye İstatistik Kurumu (TÜİK). (2024). Tüketici Fiyat Endeksi, Nisan 2024. URL: https://data.tuik.gov.tr/bulten/index?P=tuketici-fiyat-endeksi-nisan-2024-53614, Son Erişim Tarihi: 07.05.2024.

Uma, N., R. Saktheeswari, N.R., I., Sugacini, M. and Kavishree, S. (2023). Smart accident detection and alert system. *International Journal of Engineering Research & Technology*, 4(8), 114-118.

Yelgeç, M. A., & Bingöl, O. (2022). Ayrık dalgacık dönüşümü ve Xgboost ile rüzgâr gücü tahmini. *Uluslararası Teknolojik Bilimler Dergisi*, 14(2), 58–65.

Yibin, N. (2019). Machine learning techniques applied to stock price prediction. *Towards Data Science*, 6(3), 41-67.

Yule, G. U. (1927). *On a method of ınvestigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers.* London: Philosophical Transactions of the Royal Society of London, 389-420.

Yusof, U. K., Khalid, M. N. A., Hussain, A. and Shamsudin, H. (2021). Financial Time Series Forecasting Using Prophet. In F. Saeed, F. Mohammed and A. Al-Nahari (Eds.), *Innovative systems for ıntelligent health ınformatics.* Berlin: Springer International Publishing, 485-495.



Gazili olmak ayrıcalıktır