

## KARIYER.NET VERİSİYLE KISA VADELİ TARIM DIŞI İŞSİZLİK TAHMİNİ

**Barış Soybilgen\***

### Yönetici Özeti

İşsizlik verisi ekonomistler ve politika yapıcılar için ekonominin genel gidişatını görebilmek açısından önemli bir göstergedir. Bu nedenle işsizlik verisine bu aktörlerin hızlı bir şekilde ulaşması önem arz etmektedir. Gelişmiş ülkelerde işsizlik oranı kısa bir gecikmeyle açıklanmasına karşın, Türkiye'de işsizlik oranı 2 buçuk aylık bir gecikmeyle açıklanmaktadır. Bu nedenle tarım dışı işsizlik oranı için tahmin gücü yüksek bir öncü göstergeye sahip olmak önem kazanmaktadır. Bu ihtiyacı karşılamak için Betam Nisan 2011'den beri Kariyer.net'in ilan başına başvuru sayısını öncü gösterge olarak kullanmaktadır. Ama şu ana kadar detaylı bir biçimde bu öncü göstergenin tahmin performansını incelememiştir. Bu araştırma notunda Kariyer.net verisinin kısa vadeli tahmin performansı incelenmektedir.

Sonuçlar Kariyer.net ilan başına düşen başvuru verisi kullanarak yapılan nokta tahminlerin Sanayi Üretim Endeksi, Tüketici Güven Endeksi, Reel Kesim Güven Endeksi ve RKG'E'nin alt kolu olan gelecek üç aydaki toplam istihdam beklentisi kullanılarak yapılan tahminlere kıyasla mutlak olarak daha iyi sonuç verdiğini göstermektedir. Ayrıca bu aradaki farkların çoğu da istatistiksel olarak anlamlıdır. Yön tahminlerinde bu modelin en az diğer rakip modeller kadar iyi performans gösterdiği ve istatistiksel olarak da yön tahminlerinin anlamlı olduğu görülmektedir. Son olarak da Kariyer.net ilan başına düşen başvuru verisi kullanarak yapılan nokta tahminlerin yansız olduğu görülmektedir. Bunun sonucunda Kariyer.net ilan başına başvuru sayısının tarım dışı işsizlik oranı için iyi bir öncü gösterge olduğu sonucuna varabiliriz.

### 1. Giriş

İşsizlik verisi ekonomistler ve politika yapıcıların ekonominin genel gidişatını görebilmesi açısından önemli bir göstergedir. Bu nedenle işsizlik verisinin istihdam piyasası ile ilgili verileri toplayan kurumlarca kısa bir gecikmeyle açıklanması önem arz etmektedir. Gelişmiş ülkelerde istihdam piyasası ile ilgili veriler çok kısa bir gecikmeyle açıklanmaktadır. Örnek vermek gerekirse Amerika Birleşik Devletleri'nin (ABD) Mayıs 2014 dönemine dair işsizlik verisi 6 Haziran'da açıklanmakta, bunun yanında ekonomistler işsizlik verisi daha açıklanmadan istihdam piyasası ile ilgili bir dizi öncü göstergeye bakarak işsizlik verisi ile ilgili bir tahmin yapabilmektedirler. Buna karşılık Türkiye ekonomisi için işsizlik verisi 2 buçuk ay gecikmeyle açıklanmaktadır. Mayıs 2014 dönemine (üç aylık ortalama Nisan-Mayıs-Haziran) dair işsizlik verisi ancak 15 Ağustos'ta açıklanacaktır. Ayrıca Türkiye ekonomisi için işsizlik verisi tahmini yapılmasını sağlayacak istihdam piyasasına dair öncü gösterge oldukça sınırlı sayıdadır.<sup>1</sup> Tahmin yapmak amacıyla Betam Nisan 2011'den beri Kariyer.net'in ilan başına başvuru sayısını öncü gösterge olarak kullanmaktadır. Bu araştırma notunda bu öncü göstergenin kısa vadeli tahmin performansı incelenmektedir.

\* Barış Soybilgen, Betam, Araştırma Görevlisi, baris.soybilgen@bahcesehir.edu.tr

<sup>1</sup> İstihdam piyasasına dair öncü gösterge olarak İşkur'un açıkladığı bazı veriler vardır. Türkiye'de tarım dışı işsizlik oranının öncü göstergelerine dair ayrıntılı bir çalışma için bkz. Yüncüler ve diğerleri (2013).

Kariyer.net ilan başına başvuru sayısının tahmin performansını incelemek için bu araştırmada Otoregresif Dağıtılmış Gecikmeler (autoregressive distributed lag [ADL(p,k)]) modeli kullanılmaktadır. Kariyer.net ilan başına başvuru sayısının tahmin performansını karşılaştırabilmek için çeşitli modeller kullanılmıştır. Bu modeller sırasıyla şunlardan oluşmaktadır: 1. tarım dışı işsizliğin gecikmelerini içeren Otoregresif (autoregressive [AR(p)]) model, 2. Sanayi Üretim Endeksi'ni (SÜE) içeren ADL model, 3. Tüketici Güven Endeksi'ni (TGE) içeren ADL model, 4. Reel Kesim Güven Endeksi'ni (RKGE) içeren ADL model ve 5. RKGE'nin alt kolu olan gelecek üç aydaki toplam istihdam beklentisi'ni (RKGE-İst) içeren ADL modeli.

Notun geri kalanı şu şekilde organize edilmektedir: notun 2. kısmında araştırmada kullanılan veriler incelenmektedir; 3. kısmında tahmin modelleri ve metotları tanıtılmaktadır; 4. kısımda modellerin tahmin sonuçları verilmektedir ve tahmin performansları ile istatistiki testler yapılmaktadır; 5. kısımda ise sonuçlara dair kısa bir değerlendirme yer almaktadır.

## 2. Veri

Bu çalışmada asıl olarak Kariyer.net ilan başına başvuru verisinin tarım dışı işsizlik tahmin performansı incelenmektedir. Bu seriyi elde etmek için Betam Kariyer.net'ten ilgili ayın bitiminden hemen sonra mevsimsellikten arındırılmamış toplam ilan sayısı ve ilanlara yapılan toplam başvuru sayısını almaktadır. Bu iki seri TSW ile mevsimsellikten arındırıldıktan sonra toplam ilan sayısının toplam başvuru sayısına bölünmesi ile ilan başına başvuru sayısı bulunmaktadır.

Bu araştırma notunda Kariyer.net verisinin yanında dört ayrı değişkenle de tarım dışı işsizlik oranı tahmin edilmektedir. Bu dört ayrı değişkenden ilki olan SÜE, Türkiye İstatistik Kurumu (TÜİK) tarafından ilgili ayın bitiminden yaklaşık 40 gün sonra mevsimsellikten arındırılmış olarak açıklanmaktadır. Tahmin modellerinde kullanılan diğer bir seri olan TGE, TÜİK tarafından ilgili ay bitmeden önce açıklanmaktadır, ve TÜİK TGE'yi mevsimsellikten arındırmadığı için Betam bu seriyi TSW kullanarak mevsimsellikten arındırmaktadır. Son olarak bu çalışmada kullanılacak iki veri Reel Kesim Güven Endeksi (RKGE) ve RKGE'nin alt kolu olan gelecek üç aydaki toplam istihdam beklentisi (RKGE-İst) ilgili ay bitmeden Türkiye Cumhuriyeti Merkez Bankası (TCMB) tarafından mevsimsellikten arındırılmış olarak açıklanmaktadır.

Bu araştırmada tahmin edilmeye çalışılan seri olan tarım dışı işsizlik oranı TÜİK tarafından ilgili ayın bitmesinden 2 buçuk ay sonra mevsimsellikten arındırılmış bir biçimde açıklanmaktadır. RKGE ve RKGE-İst hariç bütün seriler için 2005.1-2014.2<sup>2</sup> dönemleri arasındaki veri kullanılmaktadır. RKGE ve RKGE-İst serileri için ise 2007.1-2014.2<sup>3</sup> dönemleri arasındaki seri kullanılmaktadır.

## 3. Tahmin Yöntemi

Bu notta Otoregresif Dağıtılmış Gecikmeler (autoregressive distributed lag [ADL(p,k)]) modelini kullanarak örneklem dışı tahmin üretme yaklaşımıyla (out of sample forecast) tahminler oluşturulmuştur.

Bir adım sonrasını (one-step ahead) tahmin etmek için aşağıdaki gibi bir ADL modeli yazılabilir<sup>4</sup>:

$$y_{t+1} = \sum_{i=1}^p \beta_i y_{t+1-i} + \sum_{j=1}^k \gamma_j x_{t+2-j} + \varepsilon_t$$

<sup>2</sup> Yeni tarım dışı işsizlik oranı serisi 2005.1'den başlıyor, bu araştırma notunu yazarken en son açıklanan seri 2014.2'ydi.

<sup>3</sup> RKGE ve RKGE-İst serileri 2007.1'den başlıyor.

<sup>4</sup> Tahminleri kötüleştirdiği için regresyonun sabit terimi atılmıştır

$y_t$  bağımlı/endojen değişkeni,  $x_t$  ise bağımsız/egzojen değişkeni temsil etmektedir. Modelde kullanılacak bağımsız değişkenler yukarıda bahsedilen Kariyer.net ilan başına başvuru sayısı, SÜE, TGE, RKGE ve RKGE-İst'dir. Bağımlı değişken ise tarım dışı işsizlik oranıdır.<sup>5</sup> ADL modeliyle tahmin yapabilmek için ADL modelinin parameterleri her bir kestirim dönemi (estimation period) için tekrarlanan (recursively) bir biçimde hesaplandıktan sonra, h adım sonrasını tahmin etmek için aşağıdaki model kullanılmaktadır:

$$\hat{y}_{t+h|t} = \sum_{i=1}^p \hat{\beta}_i \hat{y}_{t+h-i|t} + \sum_{j=1}^k \hat{\gamma}_j x_{t+(h+1)-j}$$

Yukarıdaki denklemde dikkat edilmesi gereken nokta t+h dönemdeki tarım dışı işsizlik oranını tahmin edebilmek için t+h dönemindeki egzojen değişkenin mevcut olması gerektiğidir. Bu nedenle ilgili dönemin hemen sonunda veya hemen öncesinde açıklanan Kariyer.net, TGE, RKGE ve RKGE-İst serileriyle iki dönem sonrasına kadar tahmin yapılabilmekte, 40 gün sonra açıklanan SÜE ile ise ancak bir sonraki dönem tahmini yapılabilmektedir.

Otoregresif (autoregressive [AR(p)]) modellerle tahmin yapmak için yukarıdaki denklemlerden egzojen değişkenleri atmak yeterlidir.

İlk kestirim dönemi RKGE ve RKGE-İst haricindeki değişkenler için 2005.1-2010.12, bu değişkenler içinse 2007.2-2010.12'dir. Sonra sabit kayan pencere (fixed rolling window) kestirim yöntemi kullanarak 2011.2-2014.2 dönemi için örneklem dışı tahminler üretilmektedir.

RKGE, RKGE-İst ve SUE kullanılan ADL modellerinde gecikme uzunluğu (lag length) her bir kestirim döneminde Schwarz enformasyon kriteri kullanılarak seçilirken, TGE ve Kariyer.net kullanılan ADL modellerinde ve AR modelinde gecikme uzunluğu her bir kestirim döneminde Akaike enformasyon kriteri kullanılarak seçilmiştir.<sup>6</sup>

#### 4. Tahmin sonuçları ve istatistik testler

Farklı tahmin modellerini karşılaştırmak için Tablo 1'de farklı modellerle yapılan tahminlerin Kök Ortalama Hata Karelerinin (Root Mean Squared Error [RMSE]) sonuçları verilmektedir. Kariyer.net verisi kullanarak yapılan ADL modelinin hem bir sonraki dönem tahmininde hem de iki sonraki dönem tahmininde diğer modellere kıyasla daha küçük RMSE'si olduğu görülmektedir. RMSE'ye göre bir sonraki en iyi model TGE'li ADL modelidir. Bu iki modeli karşılaştırsak TGE'li modelin RMSE'sinin Kariyer.net'li modele kıyasla bir sonraki dönem tahmininde yüzde 19 ve iki sonraki dönem tahmininde yüzde 29 daha yüksek olduğu görülmektedir.

Tablo 1. h-dönem sonrası tahminler için Kök Ortalama Hata Kareleri

h	AR	Kariyer.net	SÜE	TGE	RKGE	RKGE-İst
1	0,2277	0,1878	0,2292	0,2233	0,2474	0,2498
2	0,3742	0,2786	--	0,3600	0,3618	0,3669

Bu aradaki farkın istatistiksel olarak anlamlı olup olmadığı anlamak için koşulsuz (unconditional) Giacomonini ve White (2006) testi uygulanmaktadır. Giacomonini ve White testinin sıfır hipotezi iki modelin aynı düzeyde tahmin etme gücü olduğudur. Tablo 2. Kariyer.net'li modelle diğer modeller arasında yapılan Giacomonini ve White testinin sonuçları ile P-Değerlerini göstermektedir. Sonuçlar

<sup>5</sup> Bağımsız ve bağımlı değişkenlerin hepsine birim kök testi uygulanmış ve birim kökün varlığı görülmüştür. Bu nedenle bütün değişkenler denkleme log farkı alınarak koyulmuştur.

<sup>6</sup> RKGE, RKGE-İst ve SUE kullanılan ADL modellerinde Schwarz yöntemi ile gecikme uzunluğu seçmek Akaike yöntemine kıyasla daha iyi tahmin sonuçları vermektedir.

yüzde 10 anlamlılık düzeyinde bir sonraki dönem tahmini için TGE'li model hariç Kariyer.net'li modelin diğer modellerden daha üstün olduğunu göstermektedir. İki sonraki dönem sonrası içinse yüzde 10 anlamlılık düzeyinde RKGE'li model hariç Kariyer.net'li modelin diğer modellerden daha iyi tahmin yaptığı ortaya çıkmaktadır.

Tablo 2. h-dönem sonrası tahminler için Giacomonini ve White testi sonuçları ve P-Değerleri

h	AR	SÜE	TGE	RKGE	RKGE-İst
1	3.13 (0.077)	3.86 (0.050)	2.54 (0.111)	4.51 (0.034)	4.83 (0.028)
2	4.01 (0.045)	-- --	2.77 (0.096)	2.53 (0.112)	3.29 (0.070)

Not: Parantez içindeki değerler P-Değerlerini, parantez içinde olmayan değerler ise test sonuçlarını göstermektedir.

Şekil 1 Kariyer.net'li modelin bir sonraki dönem için yaptığı tahminlerle tarım dışı işsizlik oranlarını göstermektedir. Şekil 2 ise Kariyer.net'li modelin iki sonraki dönem için yaptığı tahminlerle tarım dışı işsizlik oranlarını göstermektedir. Özellikle bir sonraki dönem tahminlerinin tarım dışı işsizlik oranlarını çok yakından takip ettiği Şekil 1'de açıkça görülmektedir. İki sonraki dönem tahminlerinin ise zaman zaman asıl veriyi uzun süre yüksek veya düşük tahmin ettiği görülmektedir. Kariyer.net tahminlerinin yansız (unbiased) olup olmadığı diğer bir deyişle sistematik bir şekilde gerçek değerinin üstünde veya altında tahmin edip etmediğini anlamak için Mincer ve Zarnowitz (1969) testi uygulanmaktadır. Mincer ve Zarnowitz testini uygulamak için aşağıdaki gibi bir denklem oluşturulmaktadır:

$$y_{t+h} = \alpha + \beta \hat{y}_{t+h|t} + \varepsilon_{t+h}$$

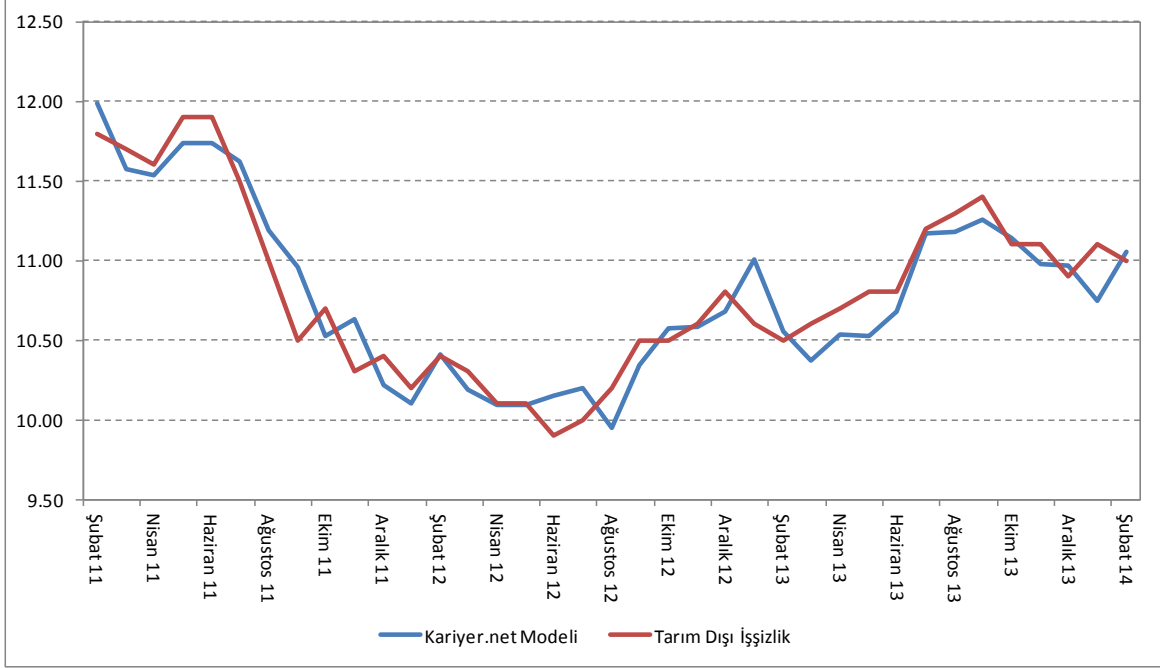
h tahmin ufkunu (forecast horizon),  $y_{t+h}$  t+h zamanındaki gerçekleşen tarım dışı işsizlik oranını,  $\hat{y}_{t+h}$  ise t zamanında t+h zamanı için yapılan tahmini göstermektedir. Eğer tahmin yansız ise  $\alpha = 0$  ve  $\beta = 1$  olduğuna dair ortak hipotezi (joint hypothesis) reddedilmesi gerekmektedir. Genel olarak tahmin hataları heteroskedastis olduğu için regresyonun kovaryans matrisi Newey ve West (1987) prosedürü kullanılarak hesaplanmaktadır. Sonra ortak hipotez Wald testi kullanarak test edilmektedir. Tablo 3'te regresyonun katsayıları ile Wald testinin sonuçları verilmektedir. Wald testinin sonuçlarına göre hem bir sonraki dönem hem de iki sonraki dönem tahmini için ortak hipotezi reddedemiyoruz. Bu da tahminlerin yansız olduğunu göstermektedir.

Tablo 3. h-dönem sonrası tahminler için Mincer ve Zarnowitz testi sonuçları

h	Regresyon Katsayıları		Wald Testi F İstatistiği
	$\alpha$	$\beta$	$\alpha = 0$ ve $\beta = 1$
1	0.51 (0.55)	0.95 (0.05)	0.68 (0.51)
2	1.33 (0.98)	0.88 (0.09)	1.25 (0.30)

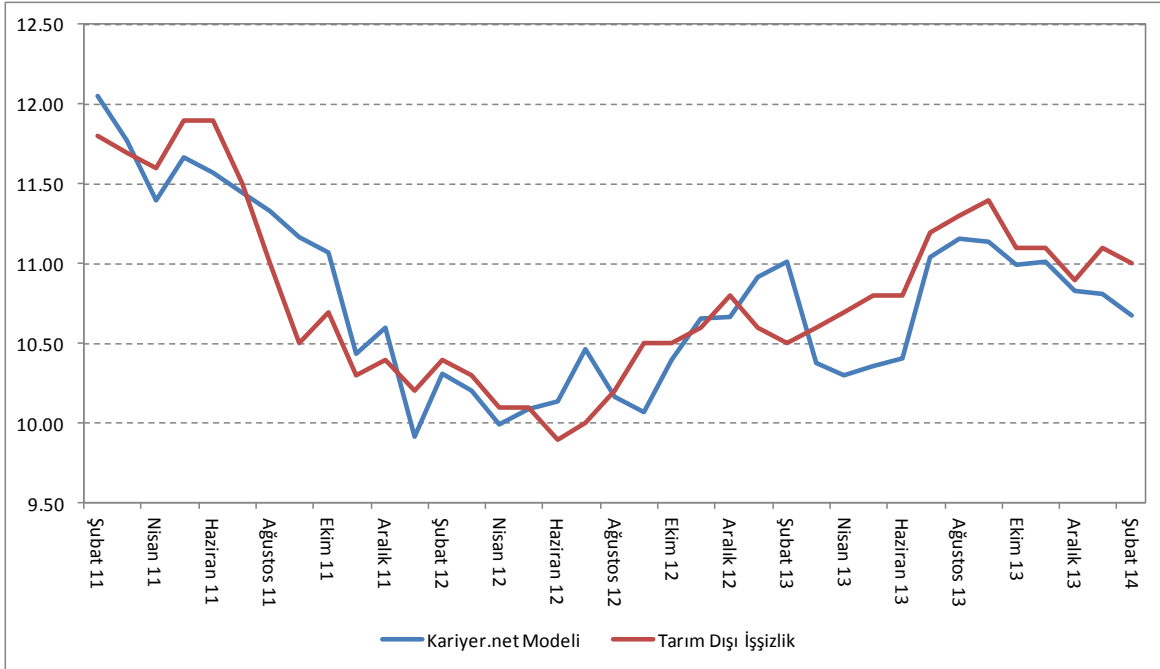
Not: Regresyon katsayılarının altındaki parantez içindeki değerler standard sapmayı, parantez içinde olmayan değerler regresyon katsayılarını göstermektedir. Wald Testi F İstatistiği altındaki parantez içindeki değerler P-Değerlerini, parantez içinde olmayan değerler F-istatistiğini göstermektedir.

Şekil 1. Tarım Dışı İşsizlik Oranları ve Kariyer.net modelinin bir sonraki dönem tahminleri



Kaynak: Tüik ve yazarın kendi hesaplamaları.

Şekil 2. Tarım Dışı İşsizlik Oranları ve Kariyer.net modelinin iki sonraki dönem tahminleri



Kaynak: Tüik ve yazarın kendi hesaplamaları.

Nokta tahminlerinin yanında yön tahminleri de işsizliğin mevcut durumu hakkında ekonomistlere ve politika yapıcılara önemli bilgiler sağlar. Modellerin yön tahmini performansları 2x2 kontenjans tablosuna (contingency table) dayalı Fisher'in Tam Olasılık / Kesin Testi (Fisher's exact test) kullanarak test edilecektir.<sup>7</sup> 2x2 kontenjans tablosu aşağıdaki gibi oluşturulmaktadır:

<sup>7</sup> Metodun daha ayrıntılı açıklaması için bkz. Henriksson ve Merton (1981) ve Joutz ve Stekler (2000).

Tablo 4. Kontenjans Tablosu

	A>0	A≤0	
F>0	n <sub>00</sub>	n <sub>10</sub>	n <sub>00</sub> + n <sub>10</sub>
F≤0	n <sub>01</sub>	n <sub>01</sub>	n <sub>01</sub> + n <sub>01</sub>
	n <sub>00</sub> + n <sub>01</sub>	n <sub>10</sub> + n <sub>01</sub>	

Tablo 4'te  $A = y_{t+h} - y_t$ , t+h dönem sonra gerçekleşen değerin t döneminde gerçekleşen değere göre farkını,  $F = \hat{y}_{t+h|t} - y_t$ , t döneminde yapılan t+h dönem sonrası tahminin t döneminde gerçekleşen değere göre farkını göstermektedir. Her bir hücrede o hücreye tekabül eden şartları sağlayan kaç tane gözlem olduğu görülmektedir. Tablo 4 kullanılarak tek yönlü Fisher kesin test istatistiği hesaplanmaktadır. Fisher's kesin testinin sıfır hipotezi gerçekleşen ve tahmin edilen değerler arasında hiç bir ilişki olmadığıdır.

Tablo 5. Bir sonraki dönem için Kontenjans Tablosu değerleri ve Fisher Kesin Test Sonuçları

	A>0 F>0	A>0 F≤0	A≤0 F>0	A≤0 F≤0	Doğru Tahminler (%)	P-Değeri
AR	11	5	6	15	70.3	0.0175
Kariyer.net	8	8	3	18	70.3	0.0230
SÜE	9	7	7	14	62.2	0.1449
TGE	10	6	6	15	67.6	0.0416
RKGE	8	8	9	12	54.1	0.4602
RKGE-İst	8	8	8	13	56.8	0.3483

Tablo 5 bir sonraki dönem tahminleri için oluşturulan kontenjans tablosu değerlerini, modelin doğru tahmin performansını ve Fisher kesin testinin P-Değerlerini göstermektedir. Bir sonraki dönem tahmininde Kariyer.net'li modelin 2011.2-2014.2 dönemi arasında yüzde 70 oranında doğru yön tahmini yaptığı görülmektedir, ve yüzde 5 anlamlılık düzeyinde, gerçekleşen ve tahmin dönemi arasında hiç bir ilişki olmadığına dair sıfır hipotezi reddedilmektedir. Tablo 5'te AR modelinin de 2011.2-2014.2 döneminde benzer bir başarımlık sağladığı görülmektedir. Tablo 6 iki sonraki dönem tahminleri için oluşturulan kontenjans tablosu değerlerini, modelin doğru tahmin performansını ve Fisher kesin testinin P-Değerlerini göstermektedir. İki sonraki dönem için Kariyer.net'li modelin yüzde 67.6 oranında doğru yön tahmini yaptığı görülmektedir, ve sıfır hipotezini yüzde 10 anlamlılık düzeyinde reddediği görülmektedir. TGE'li modelin de iki sonraki dönem yön tahminlerinde Kariyer.net'e benzer bir başarımlık gösterdiği, hatta sıfır hipotezini yüzde 5 güven aralığında reddediği görülmektedir.

Tablo 6. İki sonraki dönem için Kontenjans Tablosu değerleri ve Fisher Kesin Test Sonuçları

	A>0 F>0	A>0 F≤0	A≤0 F>0	A≤0 F≤0	Doğru Tahminler (%)	P-Değeri
AR	8	7	8	14	59.5	0.2466
Kariyer.net	8	7	5	17	67.6	0.0592
TGE	10	5	7	15	67.6	0.0394
RKGE	8	7	9	13	56.8	0.3413
RKGE-İst	6	9	10	12	48.6	0.5048

## 5. Sonuç

Kariyer.net ilan başına düşen başvuru verisi kullanarak yapılan nokta tahminler diğer bütün rakip modellerle yapılan nokta tahminlerden mutlak olarak daha iyi sonuç vermektedir. Giacomoni ve White testine göre istatistiksel olarak da yüzde 10 anlamlılık düzeyinde çoğu rakip modelden daha iyi nokta tahmin yaptığı görülmektedir. Nokta tahminlerinin, Mincer ve Zarnowitz testinin gösterdiği

üzere, yansız olduđu anlaşılmaktadır. Bu sonuçlara dayanarak Kariyer.net verisi kullanılan modelin yön tahminlerinde en az diđer rakip modeller kadar iyi performans gösterdiđi ve istatistiksel olarak da yön tahminlerinin anlamlı olduđu görülmektedir.

## Referanslar

Giacomini, R. ve White, H. (2006) Tests of conditional predictive ability, *Econometrica*, **74**, 1545-1578.

Henriksson, R. D. ve Merton, R. C. (1981) On market timing and investment performance, two statistical procedures for evaluating forecasting skills, *Journal of Business*, **54**, 513-533.

Joutz, F. ve Stekler, H. O. (2000) An evaluation of the predictions of the federal reserve, *International Journal of Forecasting*, **16**, 17-38.

Mincer, J. ve Zarnowitz, V. (1969) The evaluation of economic forecasts, in *Economic Forecasts and Expectations*. Zarnowitz, J. (Ed.), National Bureau of Economic Research, New York.

Newey, W. K. ve West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, **55**, 703-708.

Yüncüler, B. G., Şengül, G. ve Yavuz, A. (2013) A quest for leading indicators of the Turkish Unemployment Rate, *Central Bank of the Republic of Turkey*, WP: 13/41.