

FUTBOL MAÇ İSTATİSTİKLERİ FAKTÖR ANALİZİ

Bu projede amaç, futbol maç istatistikleri verileri kullanılarak, oynanan bir futbol müsabakasından hemen sonra ‘‘maçın en iyi oyuncusu’’nu tahmin etmektir.

Futbolda bir çok farklı değişkenin verisi tutulabilmektedir bu sebeple en az değişkenle en iyi tahmini elde etmek için faktör analizi uygulanmıştır.

1) Veri Yapısı

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Gol_Sayisi	Numeric	1	0		None	None	8	Right	Nominal	Input
2	Topla_Oynama	Numeric	2	0		None	None	8	Right	Scale	Input
3	Girisim	Numeric	2	0		None	None	8	Right	Scale	Input
4	Isb.Sut	Numeric	2	0		None	None	8	Right	Nominal	Input
5	Isbz.Sut	Numeric	2	0		None	None	8	Right	Nominal	Input
6	Karsilama	Numeric	2	0		None	None	8	Right	Nominal	Input
7	Korner	Numeric	2	0		None	None	8	Right	Nominal	Input
8	Ofsayt	Numeric	1	0		None	None	8	Right	Nominal	Input
9	Serbest.Vurus	Numeric	2	0		None	None	8	Right	Nominal	Input
10	Kurtaris	Numeric	1	0		None	None	8	Right	Nominal	Input
11	Isb.Pas	Numeric	2	0		None	None	8	Right	Scale	Input
12	Pas	Numeric	4	0		None	None	8	Right	Scale	Input
13	Kat.Mesafe	Numeric	3	0		None	None	8	Right	Scale	Input
14	Faul.Sayisi	Numeric	2	0		None	None	8	Right	Nominal	Input
15	Sari.Kart	Numeric	1	0		None	None	8	Right	Nominal	Input
16	Sari.Kirmizi	Numeric	1	0		None	None	8	Right	Nominal	Input
17	Kirmizi.Kart	Numeric	1	0		None	None	8	Right	Nominal	Input
18	Ilk.Gol	Numeric	2	0		None	None	8	Right	Scale	Input

Veri, 2018 Dünya Kupasında oynanan tüm maçların (64), takım istatistiklerinden oluşmaktadır. Toplamda 18 numerik değişken ve 128 satır bulunmaktadır. Oynanan her maçta her iki takım için de istatistikler tutularak, bir maç için iki adet satır elde edilmiştir. Veri, 64 maç için toplam 128 adet satırdan oluşmuştur.

Her değişken, ismi ilk sırada yazılan (ev sahibi) takım için tutulan istatistikleri ifade etmektedir.

Gol_Sayisi= Atılan gol sayısı.

Topla Oynama= Topla oynama yüzdesi.

Girişim = Toplam gol girişimi.

Isb. Sut= Kaleyi bulan şut sayısı.

Isbz. Sut= Kaleyi bulmayan şut sayısı.

Karsilama= Kaydedilen şutun kaleye ulaşmadan önlenme sayısı.

Korner= Kullanılan toplam korner sayısı.

Ofsayt= Ofsayt pozisyonunda kalma sayısı.

Serbest.Vurus= Kazanılan serbest vuruş sayısı.

Kurtaris= Kalecinin kurtardığı pozisyon sayısı.

Isb.Pas= Toplam isabetli pas sayısı.

Pas= Toplam pas sayısı.

Kat.Mesafe= Futbolcuların oyun esnasında katettiği toplam mesafe (km).

Faul.Sayisi= Yapılan toplam faul.

Sari.Kart= Görülen sarı kart sayısı.

Sari.Kirmizi= İkinci sarı kart sonrası görülen kırmızı kart sayısı.

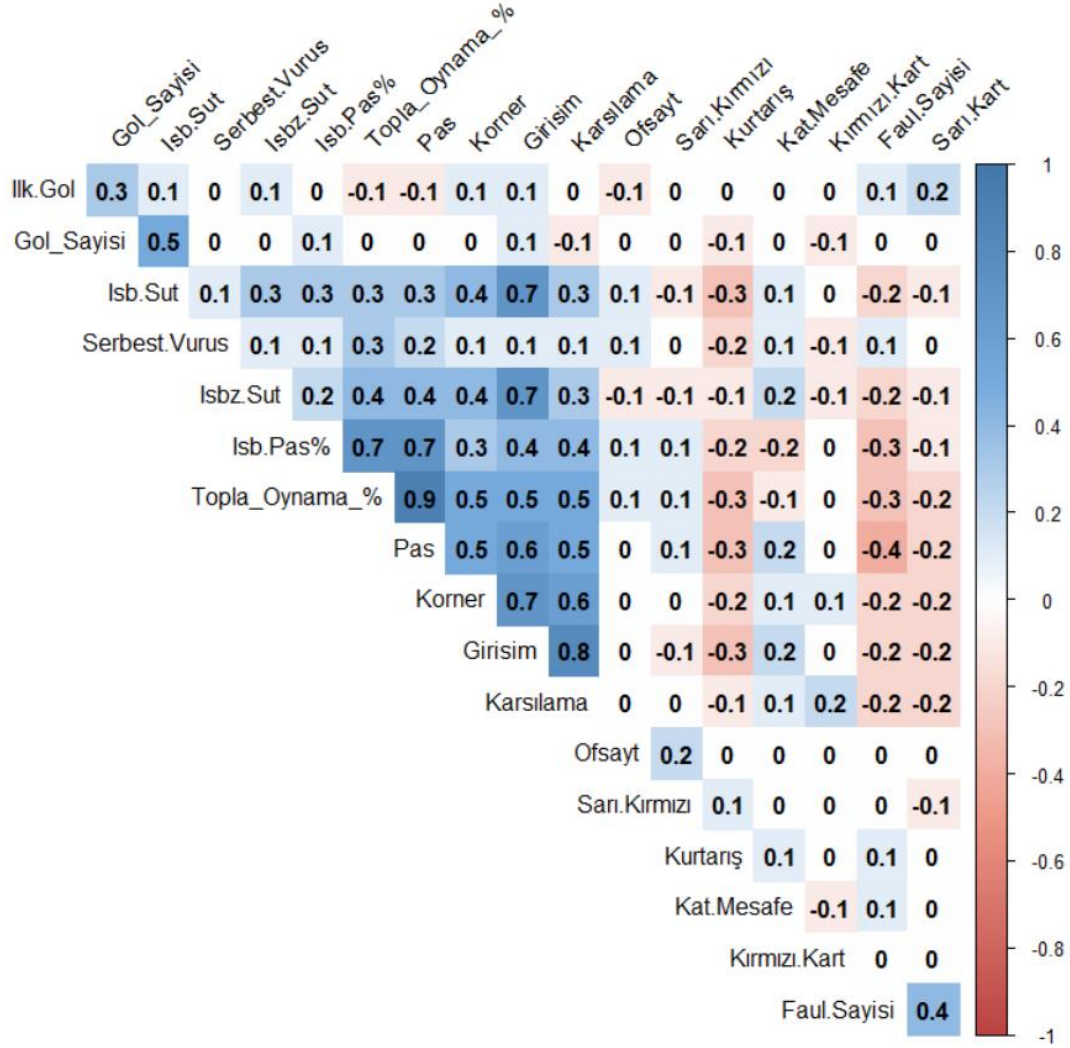
Kirmizi= Görülen kırmızı kart sayısı.

Ilk.Gol= Atılan ilk golün zamanı (dk).

2) Değişkenler Arası İlişkinin Belirlenmesi

Veride değişkenler sınıflama ve sıralama ölçeğinde olmadığından ve örneklem büyüklüğü 30'dan fazla olduğundan parametrik test kullanılmıştır.

Normallik varsayımı altında Pearson Korelasyon analizi sonucunda değişkenler arası ilişki katsayısı ve ilişkinin yönü Şekil-1'de gösterilmiştir.



Şekil 1 Korelasyon Katsayıları ve Yönü

Şekil 1'e göre mavi ve kırmızı ile renklendirilmiş hücreler, o hücelere karşılık gelen değişkenler arasındaki ilişkinin anlamlı olduğunu göstermektedir. Mavi renkli hücelere karşılık gelen değişkenler arasında aynı yönlü bir ilişkinin bulunduğu, kırmızı renkli hücelere karşılık gelen değişkenler arasında ise ters yönlü bir ilişkinin bulunduğu söylenebilir. Ayrıca renkler koyulaştıkça iki değişken arasındaki ilişkinin daha kuvvetli olduğu yorumu yapılabilir.

3) Temel Bileşenler Analizi

Toplam 18 değişkeni kullanarak, Temel Bileşenler Analizi yapmadan önce her bir değişken için örnekleme yeterliliği ölçümü (measure of sampling adequacy) yaparak, sonucunda düşük ölçüm değerine sahip değişkenler analizden atılarak, analizin KMO örnekleme yeterliliği ölçümü değeri yükseltilebilir.

Her bir değişken için örnekleme yeterliliği ölçümü değerleri Şekil 2’de gösterilmiştir.

Anti-image Matrices																		
Gol_Sayisi	,615 ^a	-0.051	0.083	-0.224	-0.031	-0.013	-0.008	-0.001	0.019	-0.029	-0.140	0.051	-0.067	-0.148	0.168	0.082	0.055	-0.292
Topla_Oynama	-0.051	,736 ^a	-0.084	0.116	0.054	0.070	-0.193	-0.057	-0.296	0.006	-0.112	-0.751	0.444	-0.119	0.012	-0.071	0.005	0.083
Girisim	0.083	-0.084	,523 ^a	-0.958	-0.964	-0.962	-0.165	0.022	0.064	0.036	0.052	0.037	-0.050	-0.004	0.055	0.073	0.109	-0.122
Isb.Sut	-0.224	0.116	-0.958	,375 ^a	0.907	0.907	0.115	-0.047	-0.064	0.030	-0.057	-0.058	0.053	0.038	-0.068	-0.074	-0.103	0.134
Isbz.Sut	-0.031	0.054	-0.964	0.907	,349 ^a	0.932	0.101	0.004	-0.078	-0.053	-0.042	-0.038	0.006	0.023	-0.044	-0.024	-0.063	0.090
Karsilama	-0.013	0.070	-0.962	0.907	0.932	,430 ^a	0.046	-0.016	-0.063	-0.056	-0.082	-0.053	0.036	-0.012	-0.032	-0.064	-0.139	0.103
Korner	-0.008	-0.193	-0.165	0.115	0.101	0.046	,934 ^a	0.021	0.098	0.077	0.092	0.030	-0.034	-0.039	0.052	-0.050	-0.117	-0.096
Ofsayt	-0.001	-0.057	0.022	-0.047	0.004	-0.016	0.021	,486 ^a	-0.066	0.012	-0.145	0.115	-0.141	0.064	-0.009	-0.203	0.021	0.129
Serbest.Vurus	0.019	-0.296	0.064	-0.064	-0.078	-0.063	0.098	-0.066	,526 ^a	0.197	-0.011	0.187	-0.165	-0.117	0.061	-0.003	0.076	-0.068
Kurtaris	-0.029	0.006	0.036	0.030	-0.053	-0.056	0.077	0.012	0.197	,783 ^a	-0.107	0.135	-0.209	0.045	0.040	-0.160	-0.028	0.022
Isb.Pas	-0.140	-0.112	0.052	-0.057	-0.042	-0.082	0.092	-0.145	-0.011	-0.107	,825 ^a	-0.352	0.377	0.082	-0.130	-0.028	0.047	-0.048
Pas	0.051	-0.751	0.037	-0.058	-0.038	-0.053	0.030	0.115	0.187	0.135	-0.352	,700 ^a	-0.606	0.216	0.026	0.003	-0.019	0.031
Kat.Mesafe	-0.067	0.444	-0.050	0.053	0.006	0.036	-0.034	-0.141	-0.165	-0.209	0.377	-0.606	,218 ^a	-0.224	-0.008	0.009	0.074	-0.013
Faul Sayisi	-0.148	-0.119	-0.004	0.038	0.023	-0.012	-0.039	0.064	-0.117	0.045	0.082	0.216	-0.224	,677 ^a	-0.415	-0.120	-0.022	0.066
Sari.Kart	0.168	0.012	0.055	-0.068	-0.044	-0.032	0.052	-0.009	0.061	0.040	-0.130	0.026	-0.008	-0.415	,601 ^a	0.153	-0.029	-0.216
Sari.Kirmizi	0.082	-0.071	0.073	-0.074	-0.024	-0.064	-0.050	-0.203	-0.003	-0.160	-0.028	0.003	0.009	-0.120	0.153	,494 ^a	0.052	-0.083
Kirmizi.Kart	0.055	0.005	0.109	-0.103	-0.063	-0.139	-0.117	0.021	0.076	-0.028	0.047	-0.019	0.074	-0.022	-0.029	0.052	,523 ^a	0.015
Ilk.Gol	-0.292	0.083	-0.122	0.134	0.090	0.103	-0.096	0.129	-0.068	0.022	-0.048	0.031	-0.013	0.066	-0.216	-0.083	0.015	,452 ^a
a. Measures of Sampling Adequacy(MSA)																		

a. Measures of Sampling Adequacy(MSA)

Şekil 2:Anti Image Matrices

Şekil 2’de “anti image” matrisinde köşegen değerlerine bakılarak, KMO değeri 0.5’in altında olan “isabetli şut”, “isabetsiz şut”, “karşılama”, “ofsayt”, “katedilen mesafe”, “sarı.kırmızı” ve “ilk gol” değişkenlerinin analize dahil edilmemesi gerektiği söylenebilir.

Temel bileşenler analizi, belirlenen değişkenlerin çıkarılmasından sonra geriye kalan 11 değişkene uygulanarak elde edilen sonuçlar şu şekildedir:

- KMO ve Bartlett Testi

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,761
Bartlett's Test of Sphericity	Approx. Chi-Square	515,835
	df	55
	Sig.	,000

KMO testi değeri 0.761 ile örneklem büyüklüğünün analiz için iyi derecede (0.7 ile 0.8 arası) olduğu söylenebilir.

- Toplam Varyansın Açıklanma Yüzdesi

Total Variance Explained						
Component	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,810	34,636	34,636	3,810	34,636	34,636
2	1,358	12,344	46,979	1,358	12,344	46,979
3	1,164	10,581	57,561	1,164	10,581	57,561
4	,995	9,042	66,603			
5	,908	8,253	74,856			
6	,828	7,532	82,388			
7	,762	6,930	89,318			
8	,494	4,490	93,808			
9	,293	2,660	96,468			
10	,283	2,572	99,039			
11	,106	,961	100,000			

Extraction Method: Principal Component Analysis.

Initial Eigenvalues (başlangıç özdeğerleri) 1'den büyük 3 faktör vardır. İlk faktör varyansın % 34,636'sını açıklarken ilk iki faktör birlikte varyansın %46,979'unu, üç faktör birlikte varyansın %57,561'ini açıklamaktadır.

- Ortak Varyans

Communalities		
	Initial	Extraction
Gol_Sayisi	1,000	,281
Topla_Oynama	1,000	,812
Girisim	1,000	,592
Korner	1,000	,595
Serbest.Vurus	1,000	,442
Kurtaris	1,000	,360
Isb.Pas	1,000	,565
Pas	1,000	,805
Faul.Sayisi	1,000	,686
Sari.Kart	1,000	,612
Kirmizi.Kart	1,000	,583

Extraction Method: Principal Component Analysis.

Tüm varyansın ortak olduğu varsayımı ile işlem yapan temel bileşenler analizinde, “gol sayısı” ile ilgili varyansın 0.281’i ortak iken “topla oynama” ile ilgili varyansın 0.812’si ortaktır.

- Bileşenler Matrisi

Component Matrix^a

	Component		
	1	2	3
Topla_Oynama	,896	,071	,065
Pas	,895	-,015	,057
Girisim	,758	,073	,109
Isb.Pas	,751	-,001	-,016
Korner	,705	,047	,309
Faul.Sayisi	-,457	,642	,256
Sari.Kart	-,336	,584	,397
Serbest.Vurus	,262	,551	-,263
Kurtaris	-,400	-,436	,096
Kirmizi.Kart	,024	-,157	,746
Gol_Sayisi	,132	,271	-,436

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Bileşenler matrisi, 3 temel bileşen sonucuna ulaştı (başlangıç özdeğeri 1’den büyük olarak seçim yaptığımız için).

1. Bileşende “topla oynama”, “pas”, “girişim”, “isabetli pas”, “korner”, “faul sayısı” ve “kurtarış” değişkenlerinin etkisi diğerlerine göre fazladır.

$$Y_1 = 0.896X_1 + 0.895X_2 + 0.758X_3 + 0.751X_4 + 0.705X_5 - 0.457X_6 - 0.4X_9$$

2. Bileşende “faul sayısı”, “sarı kart”, “serbest vuruş” ve “kurtarış” değişkenlerinin etkisi diğerlerine göre daha fazladır.

$$Y_2 = 0.642X_6 + 0.584X_7 + 0.551X_8 - 0.436X_9$$

3. Bileşende “kırmızı kart” ve “gol sayısı” değişkenlerinin etkisi diğerlerine göre daha fazladır.

$$Y_3 = 0.746X_{10} - 0.436X_{11}$$

- Döndürülmüş Bileşenler Analizi

Rotated Component Matrix^a

	Component		
	1	2	3
Topla_Oynama	,871	-,184	,138
Pas	,854	-,259	,097
Korner	,763	-,028	-,107
Girisim	,756	-,119	,079
Isb.Pas	,699	-,239	,140
Kurtaris	-,419	-,188	-,386
Faul.Sayisi	-,241	,790	,065
Sari.Kart	-,095	,774	-,061
Kirmizi.Kart	,225	,226	-,693
Serbest.Vurus	,258	,243	,562
Gol_Sayisi	,035	-,030	,528

Extraction Method: Principal Component Analysis.

Rotation Method: Equamax with Kaiser

Normalization.

a. Rotation converged in 5 iterations.

Equamax, faktörleri basitleştiren varimax yönteminin ve değişkenleri basitleştiren quartimax yönteminin birleşimi bir dönüşüm yöntemidir. Bu yöntemde, yüksek oranda bir faktöre yüklenen değişkenlerin sayısı ve bir değişkeni açıklamak için gereken faktörlerin sayısı en aza indirilir.

Dönüştürme sonucu,

1. Faktörde, dönüştürme işlemi yapılmadığı durumdaki faktör değişkenlerinden “faul sayısı” değişkeni bulunmamaktadır. Diğer değişkenlerin katsayılarında fazla değişiklik gözlemlenmemiştir.
2. Faktörde, dönüştürme işlemi yapılmadığı durumdaki faktör değişkenlerinden “serbest vuruş” ve “kurtarış” değişkenleri bulunmamaktadır. Diğer değişkenlerin katsayılarında artış gözlemlenmemiştir.
3. Faktörde, dönüştürme işlemi yapılmadığı durumdaki faktör değişkenlerine ek olarak “serbest vuruş” değişkeni bulunmaktadır. Diğer değişkenlerin katsayılarında azalış gözlemlenmiştir.