

# İşbirlikçi Filtreler İle Kitap Öneri Sistemi

Muharrem Küçükyılmaz

Bilgisayar Mühendisliği

Yıldız Teknik Üniversitesi

İstanbul, Türkiye

Email: muharrem.kucukyilmaz@gmail.com

**Abstract**—Tavsiye sistemleri kullanıcılara kişiselleştirilmiş öneri sistemi sunmaktadır. Bu makalede amaç, tavsiye sistemlerinden birisi olan kullanıcı-tabanlı işbirlikçi filtreler ise pearson katsayısı ile kullanıcılar arasındaki benzerliklere göre kişiye özel önerilerde bulunmaktadır. Kullanıcının puanlamadığı itemlerden en benzer k kişinin o iteme verdiği puanlara bakılarak kullanıcı için en iyisi önerilmektedir.

**Index Terms**—kullanıcı tabanlı , işbirlikçi filtre, pearson coef.

## I. GİRİŞ

Öneri sistemleri makine öğrenmesinin en önemli metodolojilerinden bir tanesidir. Günümüzde Netflix, Amazon, Exxen gibi milyonlarca kullanıcısı olan bir çok firma kişiye özel tavsiyeler sunmak için kullanıcı tabanlı tavsiye sistemlerini kullanmaktadır.

Kullanıcı tabanlı işbirlikçi filtrelerde, kullanıcının diğer kullanıcılar ile benzerliği pearson katsayısı ile hesaplanır. En benzer k kullanıcı k-nearest neighborhood ile bulunur. Kullanıcının puanlamadığı itemler, benzer kişilerin o itemler için verdiği puanlara bakılarak tahmini puanı hesaplanır ve en yüksek puanlı item kullanıcıya tavsiye edilir.

## II. SİSTEM TASARIMI

### A. Veri Kümesi

25 kullanıcının 8 farklı kitap için verdiği puanlar ile elde edilmiş veri kümesi figure II-A'deki gibidir.

### B. Kullanıcı Tabanlı Filtre

Kullanıcı tabanlı filtrede öncelikle bir okuyucunun diğer okuyuculara benzerliği pearson katsayısı kullanılarak hesaplanmış ve benzerlik matrisi oluşturulmuştur. Benzerlik

	TRUE BELIEVER	THE DA VINCI CODE	THE WORLD IS FLAT	MY LIFE SO FAR	THE TAKING	THE KITE RUNNER	RUNNY BABBIT	HARRY POTTER
U1	2	4		3			1	1
U2		5			3	2	1	
U3	3		1	2	2			5
U4		3			4	1		3
U5	2	4	3			2	1	
U6	5	4		3	1		3	1
U7	1	4	5	5	3			4
U8	2	2			4	5	1	
U9			4	2	1			5
U10	3	5	1				4	4
U11	2		2	4		1		2
U12	5	4		2		1	1	3
U13			2		4		4	5
U14		1	2	3	4		5	5
U15		3			5	3		2
U16		3	2	1	1		4	
U17	1	5	1	2		4		4
U18	5		4		2	1	3	5
U19		3		2		4	1	4
U20	2	5	1	1	5	4		4
NU1	4		5	3	2	3		4
NU2		5	2	5	3		2	
NU3	2	1			2	3	4	1
NU4	3	5	1	1		1		3
NU5		2	3	2	1		3	

Fig. 1. Veri Kümesi

formül 1 ile hesaplanmıştır.

$$sim(a, b) = \frac{\sum_{p \in P} (r_{a,p} - \bar{r}_a)(r_{b,p} - \bar{r}_b)}{\sqrt{\sum_{p \in P} (r_{a,p} - \bar{r}_a)^2} \sqrt{\sum_{p \in P} (r_{b,p} - \bar{r}_b)^2}}$$

P :hem a hem b tarafından okunmuş kitaplar

$r_{a,p}$  :a'nın p kitabı için verdiği puan

$r_{b,p}$ :b'nın p kitabı için verdiği puan

$\bar{r}_a$ :a'nın okuduğu kitapların puan ortalaması

$\bar{r}_b$ :b'nın okuduğu kitapların puan ortalaması

Figure 2’de, kullanıcılar arasındaki benzerlik matrisi görülmektedir. Benzerlik değerleri  $[-1, +1]$  arasındadır.  $(-1)$  değeri iki kullanıcı arasında negatif ilişki olduğunu,  $(+1)$  değeri ise iki kullanıcı arasında pozitif ilişki olduğunu göstermektedir. Bu benzerlik matrisinden yararlanarak NU1-5 kullanıcıları için U1-20 kullanıcısı arasından en benzer  $k$  kullanıcı bulunmuştur. NU1-5 kullanıcısının puan vermediği kitaplar için,  $k$  benzer kullanıcıya bakılarak formül 2 ile tahmini puanlar hesaplanmış ve en yüksek puanlı kitap tavsiye edilmiştir.

$$pred(a, p) = \bar{r}_a + \frac{\sum_{b \in N} sim(a, b)(r_{b,p} - \bar{r}_b)}{\sum_{b \in N} sim(a, b)} \quad (2)$$

Figure 3’de akış diyagramı görülmektedir.

### III. DENEYSEL ANALİZ

- NU1-5 kullanıcısı için  $k=1$  ve  $k=3$  iken en benzer kullanıcılara bakılarak önerilen kitaplar tablo 4’deki gibidir. NU1 kullanıcısı için  $k=1$  iken en benzer kullanıcı U9 dur. NU1 kullanıcısının puanlamadığı ‘the da-vinci code’ ve ‘runny babbit’ kitaplarını U9 kullanıcısıda puanlamadığı için kitap önerisi yapılamamıştır.
- NU2 kullanıcısına  $k=1$  için ‘the kite runner’ önerilmişken  $k=3$  için ‘true believer’ kitabı önerilmiştir. Çünkü  $k=1$  iken en benzer kullanıcı U2 dir ve U2 kullanıcısı sadece ‘the kite runner’ kitabını oylamıştır. Ama  $k=3$  için en benzer kullanıcılar U2, U1, U11 dir. Bu üç kullanıcının oyladığı kitaplara bakılarak ‘true believer’ kitabının tahmin değeri daha yüksek çıkmıştır ve bu kitap tavsiye edilmiştir.

### IV. SONUÇ

Kullanıcı tabanlı işbirlikçi filtreler kişiye, benzer kullanıcıyı bulup ona göre öneride bulunduğu için aynı türde kitapları önermek yerine farklı türde kitap da önerir. Kişiye özel kitap önermektedir.

Farklı  $k$  değerleri için farklı kitap önermektedir ve bu optimum  $k$  değerini belirleme problemine yol açmaktadır. Ayrıca hiçbir kitaba puan vermemiş kullanıcıya kitap önerisinde bulunamaz. Bu konu üzerinde geliştirme yapılabilir.

sim	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	0	0.997	-0.943	0.196	0.987	0.473	0.258	0.191	-0.992	0.33	0.648	0.354	-0.832	-0.977	0.381	-0.316	0.299	-0.475	0.073	0.113		-0.872	0.952	-0.467	0.387	-0.756
1	0.997	0	-1	0.424	0.996	0.301	0.346	0	-1	0.524	1	0.898	-0.6	-0.95	0.015	-0.25	0.685	0.385	0.422	0.607	0	0	0.961	-0.942	0.99	-0.516
2	-0.943	-1	0	-0.048	-0.942	-0.291	-0.276	-1	0.526	0.702	-0.242	0.348	0.886	0.873	0.761	0.569	0.624	0.531	0.942	0.448		0.029	0.388	-0.928	0.766	-0.167
3	0.196	0.424	-0.048	0	0.377	-0.666	-0.629	-0.448	-0.555	0.91	0.953	0.864	0.385	0.174	0.474	0.707	0.249	0.414	-0.574	0.238		-0.244	-0.048	-0.556	0.581	-0.999
4	0.987	0.996	-0.942	0.377	0	0.144	0.456	0.101	1	0.066	0.433	0.534	-0.52	-0.983	-0.514	-0.311	0.409	0.235	0.502	0.315		0.658	0.654	-0.916	0.622	-0.378
5	0.473	0.301	-0.291	-0.666	0.144	0	-0.412	-0.689	-0.021	0	0.045	0.561	-0.802	-0.699	-0.196	0.569	-0.381	0.319	-0.474	-0.429		0.484	0.448	0.116	0.324	0.643
6	0.258	0.346	-0.276	-0.629	0.456	-0.412	0	0.234	0.312	-0.156	0.432	-0.779	-0.63	-0.433	-0.943	-0.209	0.202	-0.246	-0.291	-0.209		0.152	0.196	-0.048	-0.376	0.515
7	0.191	0	-1	-0.448	0.101	-0.689	0.234	0	-1	-0.547	-0.871	-0.329	-0.196	-0.049	0.373	-1	0.303	-0.737	0.892	0.431		-0.756	0.152	-0.078	-0.741	-0.81
8	-0.992	-1	0.526	-0.555	1	-0.021	0.312	-1	0	-0.217	-0.538	0.8	0.038	0.14	-0.986	0.812	0.234	0.994	0.992	-0.173		0.849	-0.415	-0.6	0.272	0.953
9	0.33	0.524	0.702	0.91	0.066	0	-0.156	-0.547	-0.217	0	0.507	0.068	0.925	0.136	-0.528	0.487	0.875	-0.23	-0.01	0.962		-0.842	0.677	-0.305	0.817	-0.52
10	0.648	1	-0.242	0.953	0.433	0.045	0.432	-0.871	-0.538	0.507	0	0.041	0.164	-0.169	0.355	-0.962	-0.344	0.473	-0.765	-0.581		-0.203	0.821	-0.412	-0.146	-0.348
11	0.354	0.898	0.348	0.864	0.534	0.561	-0.779	-0.329	0.8	0.068	0.041	0	-0.678	-0.119	-0.22	-0.217	0.8	0.3	0.009	0.824		0.844	0.571	-0.724	0.765	-0.773
12	-0.832	-0.6	0.886	0.385	-0.52	-0.802	-0.63	-0.196	0.038	0.925	0.164	0	0	0.823	-0.41	0.129	0.999	0.101	0.381	0.824		-0.503	0.555	-0.368	0.988	-0.509
13	0.977	-0.95	0.873	0.174	-0.983	-0.699	-0.433	-0.049	0.14	0.136	-0.169	-0.678	0.823	0	-0.052	0.118	-0.038	0.07	-0.155	0.143		-0.396	-0.507	0.445	-0.273	0
14	0.381	0.015	-0.761	0.474	-0.514	-0.196	-0.943	0.373	-0.986	-0.528	0.355	-0.119	-0.41	-0.052	0	-0.902	-0.648	-0.56	-0.833	0.238		-0.87	-0.377	0.309	-0.294	-0.953
15	-0.316	-0.25	0.569	-0.707	-0.311	0.569	-0.209	-1	0.812	0.487	-0.962	-0.22	0.129	0.118	-0.902	0	0.72	0.261	-0.464	0.298		0.564	-0.303	0.511	0.841	0.642
16	0.299	0.685	0.624	-0.249	0.409	-0.381	0.202	0.303	0.234	0.875	-0.344	-0.217	0.999	-0.038	-0.648	0.72	0	-0.486	0.726	0.915		-0.566	0.597	-0.429	0.521	-0.691
17	-0.475	0.385	0.531	0.414	0.235	0.319	-0.246	-0.737	0.994	-0.23	0.473	0.8	0.101	0.07	-0.56	0.261	-0.486	0	-0.028	-0.534		0.7	0.022	-0.542	0.623	0.741
18	0.073	0.422	0.942	-0.574	0.502	-0.474	-0.291	0.892	0.992	-0.01	-0.765	0.3	0.381	-0.155	-0.833	-0.464	0.726	-0.028	0	0.712		0.246	0.296	-0.609	0.127	-0.786
19	0.113	0.607	0.448	0.238	0.315	-0.429	-0.209	0.431	-0.173	0.962	-0.581	0.009	0.824	0.143	0.238	0.298	0.915	-0.534	0.712	0		-0.597	0.167	-0.444	0.651	-0.659
20																				0		0	-0.488	-0.331	0	0.977
21	-0.872	0	0.029	-0.244	0.658	0.484	0.152	-0.756	0.849	-0.842	-0.203	0.844	-0.503	-0.396	-0.87	0.564	-0.566	0.7	0.246	-0.597		-0.488	0	-0.926	0.46	-0.473
22	0.952	0.961	0.388	-0.048	0.654	0.448	0.196	0.152	-0.415	0.677	0.821	0.571	0.555	-0.507	-0.377	0.303	0.597	0.022	0.296	0.167		-0.488	0	-0.926	0	-0.881
23	-0.467	-0.942	-0.928	-0.556	-0.916	0.116	-0.048	-0.078	-0.6	-0.305	-0.412	-0.724	-0.368	0.445	0.309	0.511	-0.542	-0.542	-0.609	-0.444		-0.331	-0.926	0	-0.881	0.599
24	0.387	0.99	0.766	0.581	0.622	0.324	-0.376	-0.741	0.272	0.817	-0.146	0.765	0.988	-0.273	-0.294	0.841	0.521	0.623	0.127	0.651		0	0.46	-0.881	0	-0.481
25	-0.756	-0.516	-0.167	-0.999	-0.378	0.643	0.515	-0.81	0.953	-0.52	-0.348	-0.773	-0.509	0	-0.953	0.642	-0.691	0.741	-0.786	-0.659		0.977	-0.473	0.599	-0.481	0

Fig. 2. Benzerlik matrisi

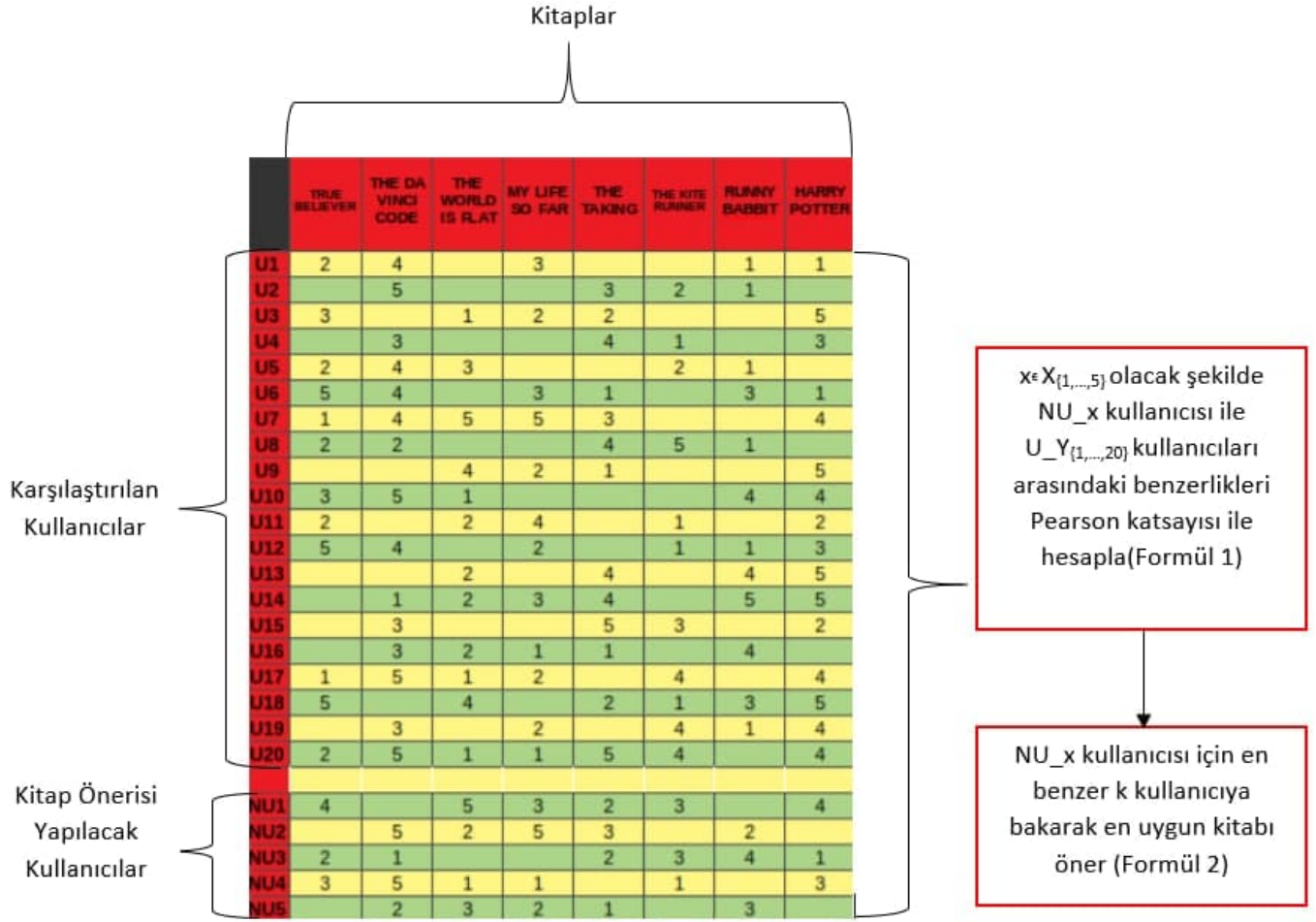


Fig. 3. Sistem Akış diyagramı

	K	En Benzer 1. Kişi	En Benzer 2. Kişi	En Benzer 3. Kişi	Puanlanmamış Kitap 1	Puanlanmamış Kitap 2	Puanlanmamış Kitap 3	ÖNERİLEN KİTAP
NU1	1	U9			THE DA VİNÇİ CODE	RUNNY BABBİT		
NU1	3	U9	U12	U18	THE DA VİNÇİ CODE	RUNNY BABBİT		THE DA VİNÇİ CODE
NU2	1	U2			TRUE BELIEVER	THE KITE RUNNER	HARRY POTTER	THE KITE RUNNER
NU2	3	U2	U1	U11	TRUE BELIEVER	THE KITE RUNNER	HARRY POTTER	TRUE BELİVER
NU3	1	U16			THE WORLD IS FLAT	MY LIFE SO FAR		THE WORLD IS FLAT
NU3	3	U16	U14	U15	THE WORLD IS FLAT	MY LIFE SO FAR		THE WORLD IS FLAT
NU4	1	U2			THE TAKING	RUNNY BABBİT		THE TAKING
NU4	3	U2	U13	U16	THE TAKING	RUNNY BABBİT		RUNNY BABBİT
NU5	1	U9			TRUE BELIEVER	THE KITE RUNNER	HARRY POTTER	HARRY POTTER
NU5	3	U9	U18	U6	TRUE BELIEVER	THE KITE RUNNER	HARRY POTTER	TRUE BELİVER

Fig. 4. NU1-5 kullanıcıları için önerilen kitaplar