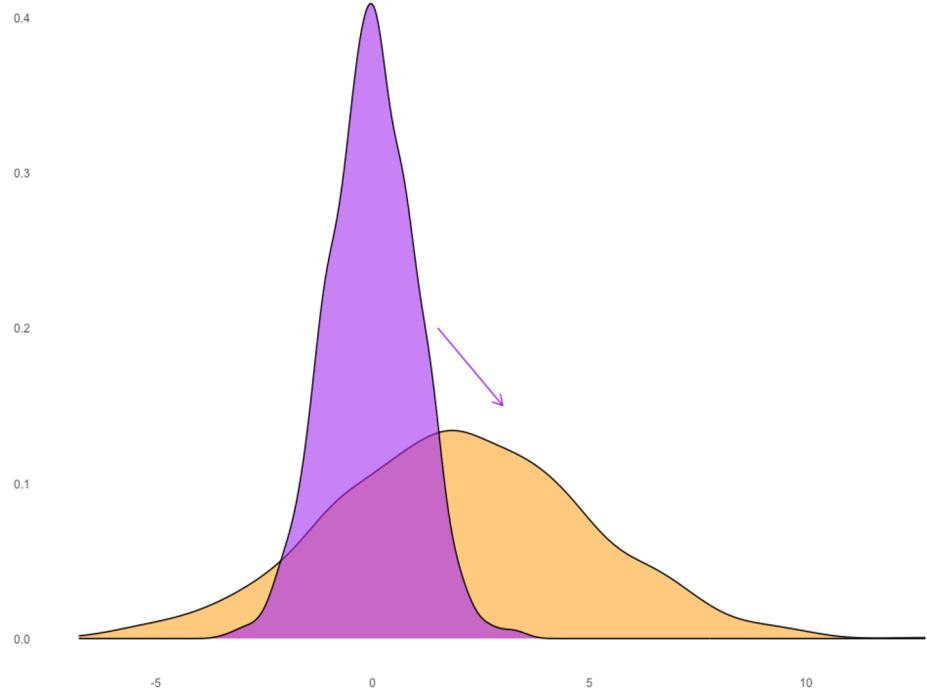


TAHMİN MODELLERİNDE VERİ
KAYMALARININ AÇIKLANABİLİR
YAPAY ZEKÂ ARAÇLARIYLA TESPİT EDİLMESİ

Uğur DAR

Tez Danışmanı: Dr. Öğr. Üyesi Mustafa ÇAVUŞ

Veri Kayması



Veri kayması, eğitim ve yeni gelen veri setleri arasında veri dağılımındaki farklılık nedeniyle model performansının düşmesi durumudur.

$$\exists X : P_{t_0}(X, y) \neq P_{t_i}(X, y), i > 0$$

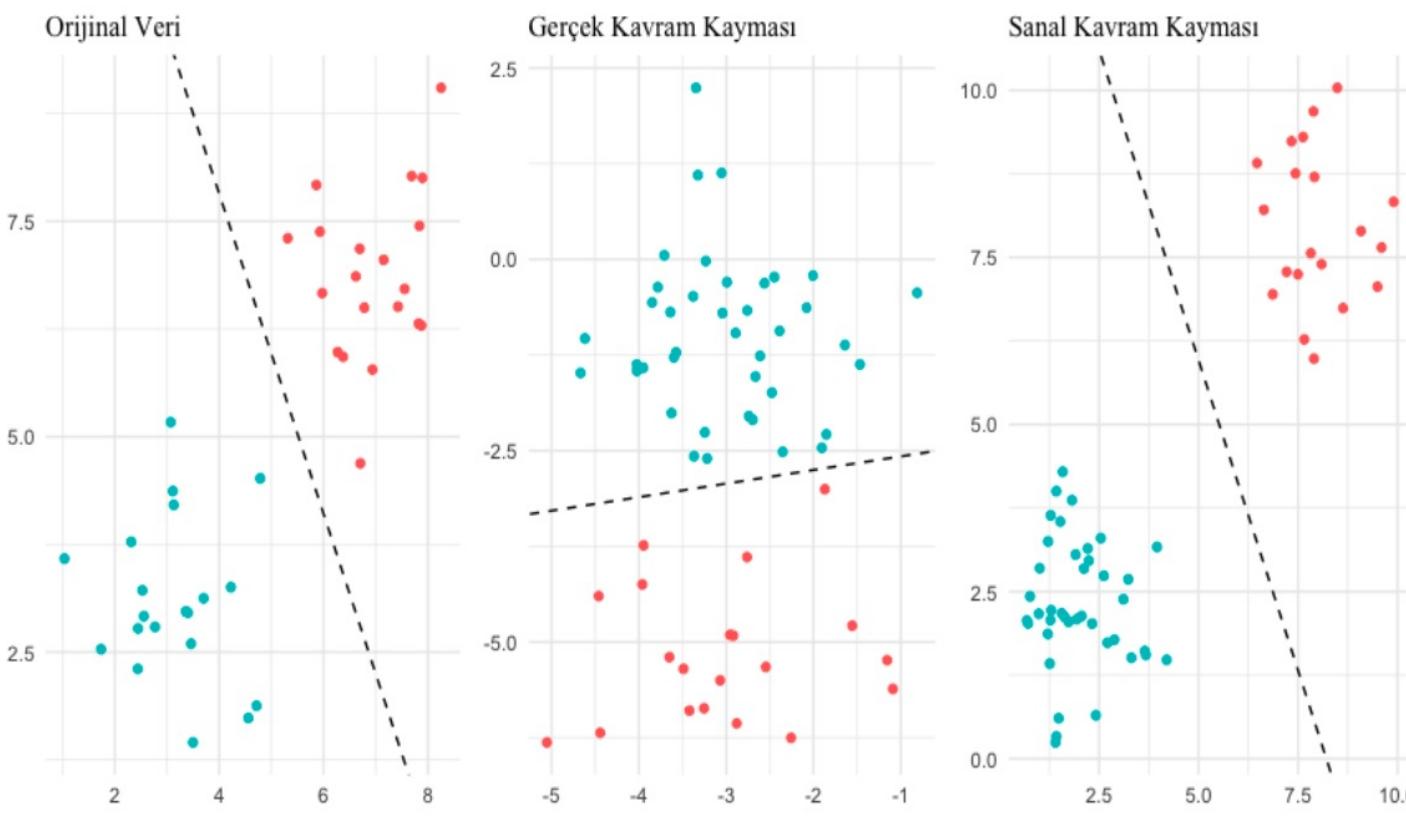
Yığın öğrenme (*batch learning, offline learning*)

Modelin sabit bir veri kümesi veya bir veri yığını üzerinde toplu olarak eğitilmesi. Manuel ya da otomatik bir şekilde yapılabilir. Statik, klasik makine öğrenmesi algoritmaları kullanılabilir.

Çevrimiçi öğrenme (*online learning, incremental learning*)

Modelin veri akışı sırasında sürekli olarak eğitilmesi ve güncellenmesi. Otomatik olarak artırılmış modeller kullanılır.

Kavram Kayması



Celik vd. (2022) kavram kaymasını 2 başlık altında incelemiştir.

- **Gerçek Kavram Kayması**

$$\exists X : P_{t_0}(y | X) \neq P_{t_i}(y | X), i > 0$$

- **Sanal Kavram Kayması**

$$P_{t_0}(y) \neq P_{t_i}(y)$$

$$P_{t_0}(x) \neq P_{t_i}(x)$$

Moreno-Torres vd., (2012) tarafından 3 başlık altında incelenmiştir.

1-Kovaryant kayması

2- Yanıt değişkeninde kayma

3- Kavram kayması

Kavram kayması tespitinde literatürde farklı yöntemler mevcuttur.

Kavram Kayması Tespit Yöntemleri

1- İstatistiksel Testler

Yeni gelen verilerin dağılımı incelenir.

2- Model performansının izlenmesi

Accuracy,

F1-Score,

RMSE,

MAPE,

vb.

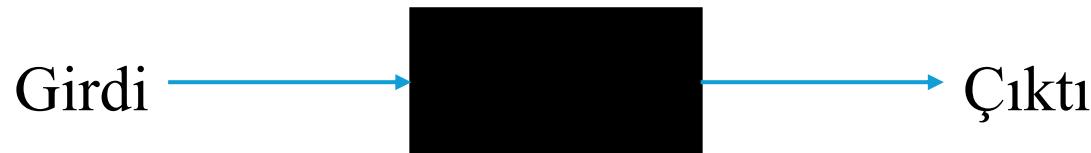
Litratürde istatistiksel testler genel olarak değerlendirme metriğinin dağılımı incellenerekullanılmıştır.

İstatistiksel testler de literatürde kendi içerisinde farklı kategorilerde sınıflandırılmıştır.

İstatistiksel Testler

- Ardışık Analize Dayalı Yöntemler (*Sequential Analysis based Methods*): Kümülatif Toplamlar (Cumulative Sums - CUSUM), Geometrik Hareketli Ortalama, Page Hinkley (*PH*) bu tip yöntemlere örnek olarak verilebilir.
- İstatistiksel Tabanlı Yöntemler (*Statistical-based Methods*): Verinin istatistiksel karakteristiklerini araştırır. Genelde bu karakteristik özellikler ortalama ve standart sapmadır. Drift Tespit Metodu (*Drift Detection Method - DDM*), Erken Drift Tespit Metodu (*Early Drift Detection Method - EDDM*) ve Üstel Ağrılıklı Hareketli Ortalama (*Exponentially Weighted Moving Average - EWMA*) bu tip yöntemlere örnek olarak verilebilebilir.
- Pencere Tabanlı Yöntemler (*Windows-based Methods*): Genellikle sabit bir referans penceresini kullanarak geçmiş dağılımları ve en son dağılımları özetleyen bir kayan pencere kullanılır. Bu pencerelerin dağılımları arasında önemli bir fark olması, veri kaymasının meydana geldiğini öne sürer. Uyarlanabilir Pencereleme (*Adaptive Windowing - ADWIN*), Hoeffding Drift Tespit Yöntemleri (*Hoeffding Drift Detection Methods – HDDM-A testi ve HDDM-W testi*) ve KSWIN (*Kolmogorov-Smirnov Windowing*) dedektörleri bu tip yöntemlere örnek olarak verilebilir.

Açıklanabilir Yapay Zeka (*eXplainable AI*)



- **Güven**
- **Hata tespiti**
- **Yasal ve etik uyumluluk**

(*General Data Protection Regulation -*
GDPR 2018)

- **Model izleme**

Model-Agnostic:

Her makine öğrenimi modeline uygulanabilecek teknikler.

- **Şeffaflık**
- **Güvenilirlik**
- **Adalet**
- **İzlenebilirlik**

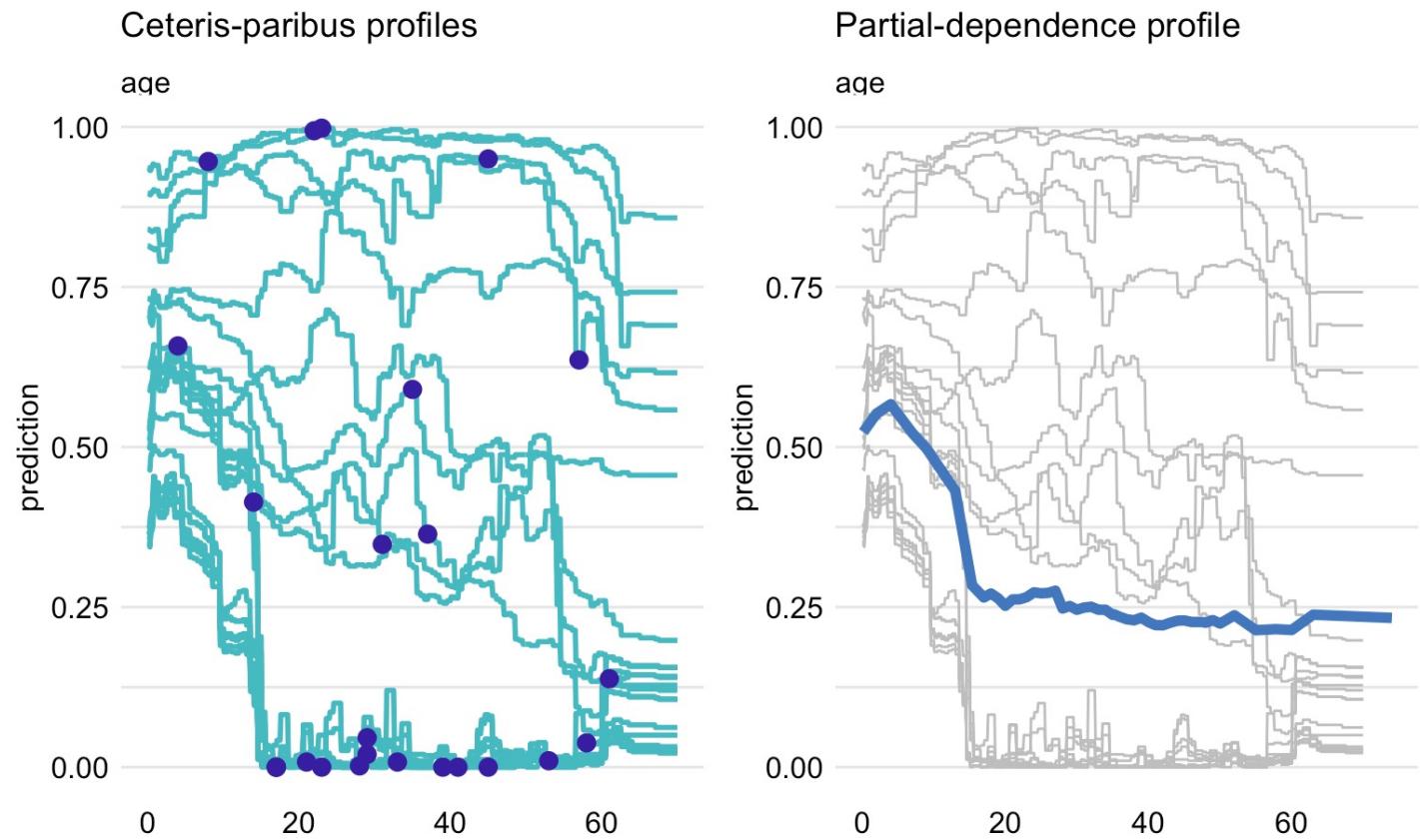
Model-Specific:

Belirli bir makine öğrenimi modeli türü için tasarlanmış teknikler.

Kısmi Bağımlılık Profili (PDP - *Partial-dependence Profile*)

PDP, modelde kullanılan değişkenlerin marjinal dağılımlarının tahmin edilen değerler üzerindeki etkisini hesaplamaktadır (Biecek ve Burzykowski, 2021).

- Ceteris-paribus profillerinin ortalamasıdır.
- Model Agnostic yöntemlerdir, bu sebeple farklı türde modeller ile birlikte kullanılabilirler.



Literatür

- Literatürde değişken önemini ölçmek için XAI araçları SHAP, permüstasyonel değişken önemliliği, artımlı permüstasyonel değişken önemliliği gibi araçlar kullanılmıştır (Muschalik vd. 2022; Fumagalli vd. 2023; Rahmani vd. 2022, . Bu araçların temel problemi genel olarak önemlilik ölçüsü, bağımsız değişken ve yanıt değişkeni arasındaki ilişkinin detaylı olarak incelenmesi zordur.
- Literatürde PDP'ler ile kavram kayması tespitine yönelik artırımlı modellerde kullanılmak üzere Muschalik vd. (2023) tarafından *incremental-PDP*(iPDP) geliştirilmiştir. Ancak yapılan çalışmada kavram kayması tespiti için doğruluk metriği ADWIN ile incelenmiştir ve PDP'ler üzerinden doğrudan kavram kayması tespit edilmemiştir. iPDP artırımlı modellerde yeni gelen veriler ile eski PDP'nin ağırlık ortalamasını alarak yeni PDP oluşturmaya odaklanmıştır.

Profil Farklılıklarının Hesaplanması

- 1- PDP'ler arası Öklidyen uzaklığının hesaplanması (L2).
- 2- PDP'lerdeki değişimin (türev değerleri) arasındaki Öklidyen uzaklığının hesaplanması (L2Der).
- 3- PDP'lerde aynı gözlem noktalarında artış olup olmadığını hesaplanması (*PDI - Partial Disparity Index*).

Bu 3 farklı yöntem kullanılarak PDP'ler arasındaki farklar hesaplanmıştır.

İnterpolasyon

$$\widehat{PDI} \left(\widehat{g_{f_1}^J}, \widehat{g_{f_2}^J} \right) = \frac{1}{m} \sum_{i=1}^m I [sgn \left(der \left(\widehat{g_{f_1}^J} \right) [i] \right) \neq sgn \left(der \left(\widehat{g_{f_2}^J} \right) [i] \right)]$$

: $der \left(\widehat{g_{f_k}^J} \right) [i]$ k. model ile elde edilen PDP'nin türevler vektörünün i. elemanı.

PDI $[0,1]$ aralığında değerler alır. 0 iki PDP'nin şekil olarak benzediğini ifade eder.

İki PDP arasında X değişkenleri aynı değerleri almayıabilir. Bu durumda karşılaştırma yapılamaz. Bu sorunu çözmek için doğrusal interpolasyon uygulanmıştır.

Profil Kayması Tespiti (PDD- *Profile Drift Detection*)

- 1- Model eğitilir.
- 2- Eğitim ve test setleri üzerinden PDP hesaplanır.
- 3- Hesaplanan PDP'ler için karşılaştırma metrikleri, PDI, L2Der ve L2 hesaplanır (Profile1).
- 4- Her yeni gelen yığını için PDP ve metriklerin hesaplanır (Profile2).
- 5- Profile2'de hesaplanan değerler Profile1'de hesaplanan değerlerden büyükse kavram kayması uyarısı ve Profile1'in Profile2 metrikleriyle değiştirilir.
- 6- Her yeni gelen yığın için 4. ve 5. adımlar tekrar edilir.

Deneyler

- 3 sentetik (SEA, Hyperplane, Friedman) ve 3 gerçek hayat (NOA, Ozone, Elec2) veri setleri üzerinde deneyler yapılmıştır.
- Klasik istatistiksel testler HDDM-A, HDDM-W, KSWIN, PH, DDM, EDDM ve PDD kullanılarak farklı yığın sayılarında (3, 5, 15, 20, 25, 30) kavram kayması tespit edilmiştir.
- Kavram kayması tespit edildiğinde model veriye o yığın içerisindeki veriler de eklenecek model tekrar eğitilmiştir. Model olarak her aşamada Rassal Orman modeli ($n\text{tree}=50$) kullanılmıştır.

SEA

- Sentetik veri seti.
- 50.000 gözlem ve 3 bağımsız değişken.
- Bağımsız değişlenler $V1 + V2 < 8$ ise $y=1$ değilse $y=0$ (Street & Kim, 2001).

Yöntem	Yığın Sayısı							
	3		5		10		15	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,8212	0	0,8508	0	0,8264	1	0,8406	1
HDDM-W	0,8212	1	0,8508	1	0,8257	3	0,8392	3
KSWIN	0,8129	3	0,8490	5	0,8268	10	0,8406	15
PH	0,8212	0	0,8508	1	0,8298	1	0,8410	1
DDM	0,8212	0	0,8515	1	0,8267	0	0,8394	0
EDDM	0,8129	3	0,8490	5	0,8268	10	0,8406	15
PDD	0,8129	3	0,8210	2	0,8285	1	0,8311	2
Test	0,8205		0,8502		0,8537		0,8900	
Eğitim	0,9988		0,9990		0,9991		0,9990	
Eğitim-Test	26666	6667	20000	5000	13332	3334	10000	2500

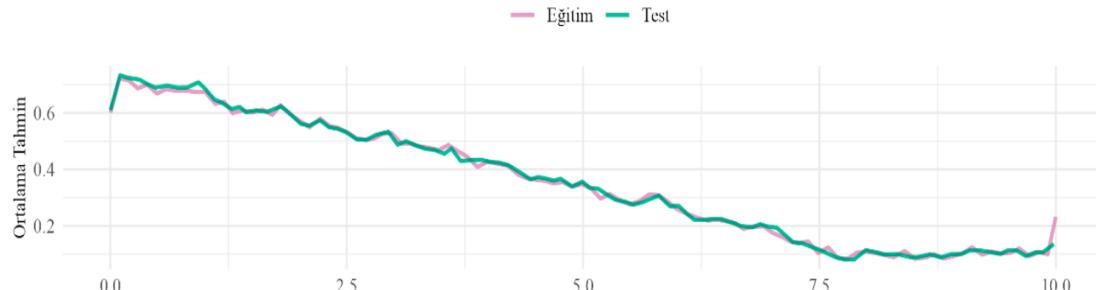
PDD yöntemi diğer istatistik testlere göre model performansı bakımından yakın sonuçlar vermiştir.

PDD, yığın sayısı arttıkça diğer testlere oranla daha az kavram kayması tespit etmiştir. Bu nedenle model yeni gelen yığılarda daha az güncellenmiştir. Bu da model eğitme maliyetini ciddi oranda düşürebilir.

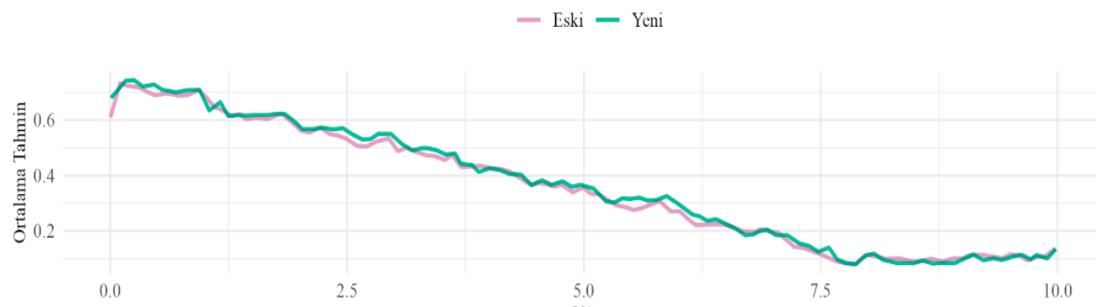
Yöntem	Yığın Sayısı					
	20		25		30	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,8432	2	0,8363	2	0,8384	2
HDDM-W	0,8412	4	0,8343	3	0,8362	3
KSWIN	0,8421	20	0,8372	25	0,8387	29
PH	0,8417	3	0,8374	2	0,8393	3
DDM	0,8430	3	0,8355	1	0,8379	3
EDDM	0,8423	19	0,8372	25	0,8390	13
PDD	0,8338	2	0,8356	2	0,8355	1
Test	0,8896		0,8795		0,9007	
Eğitim	0,9989		0,9988		0,9991	
Eğitim-Test	8000	2000	6666	1667	5713	1429

PDD, PH ve DDM testleri kavram kayması tespitinde ve model performansında birbirine yakın sonuçlar vermiştir

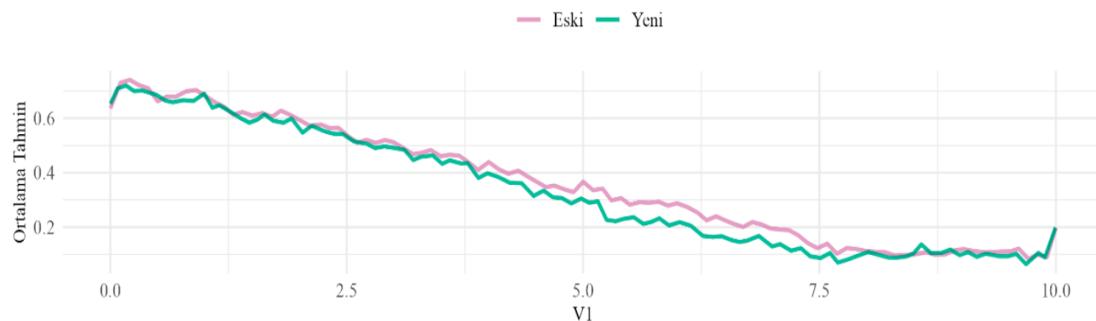
Test - Doğruluk: 0.89 - PDI: 0.14 - L2Der: 0.13 - L2: 0.03



İndeks: 1 - Doğruluk: 0.85 - PDI: 0.19 - L2Der: 0.16 - L2: 0.06

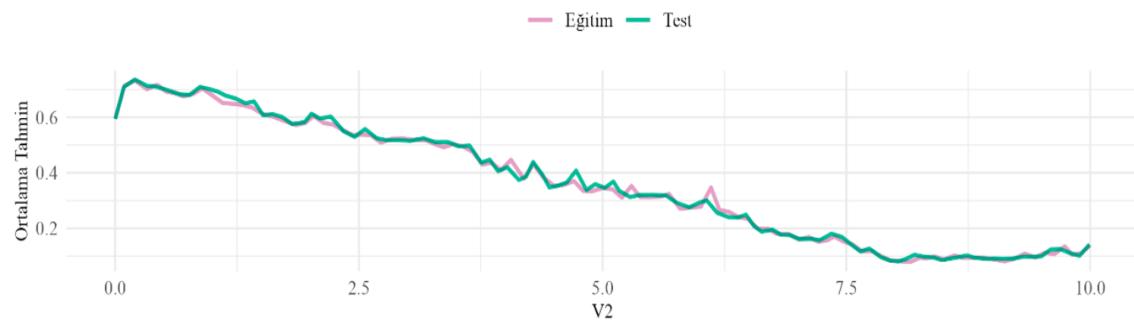


İndeks: 2 - Doğruluk: 0.85 - PDI: 0.22 - L2Der: 0.17 - L2: 0.12

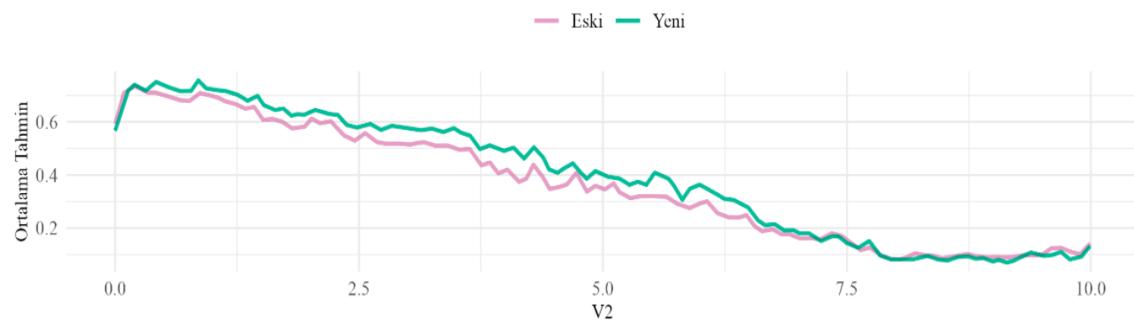


Y1ğın-15

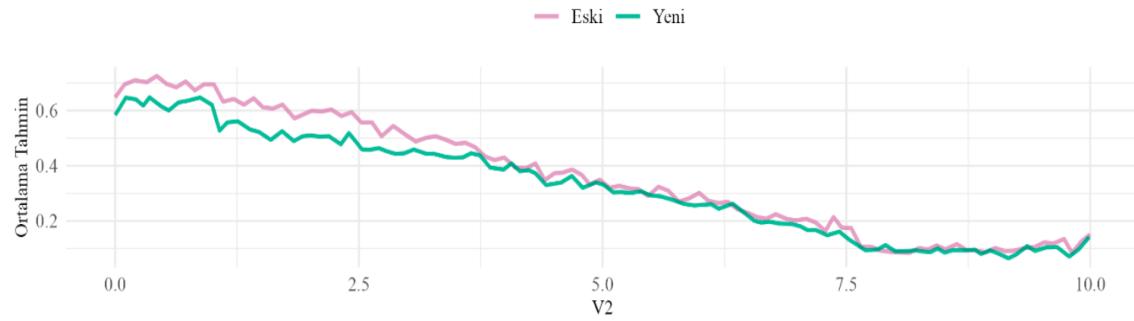
Test - Doğruluk: 0.89 - PDI: 0.19 - L2Der: 0.12 - L2: 0.04



İndeks: 2 - Doğruluk: 0.85 - PDI: 0.2 - L2Der: 0.18 - L2: 0.14



İndeks: 19 - Doğruluk: 0.82 - PDI: 0.24 - L2Der: 0.18 - L2: 0.16



Y1ğın-20

Hyperplane

- Sentetik veri seti.
- 20.000 gözlem ve 2 bağımsız değişken.
- Bu veri seti ile yapılan deneyde, kavram kayması tespiti için PDD'de iki farklı yöntem izlenmiştir;

1 - PDD, diğer veri setlerinde olduğu gibi yeni gelen yiğında hesaplanan 3 metriğin 3'nün de bir önceki hesaplanan metriklerden büyük olduğu durum.

2 - PDD2, hesaplanan 3 metrikten 2'sinin bir önceki eşik değerden yüksek olması kavram kayması olarak değerlendirilmiştir.

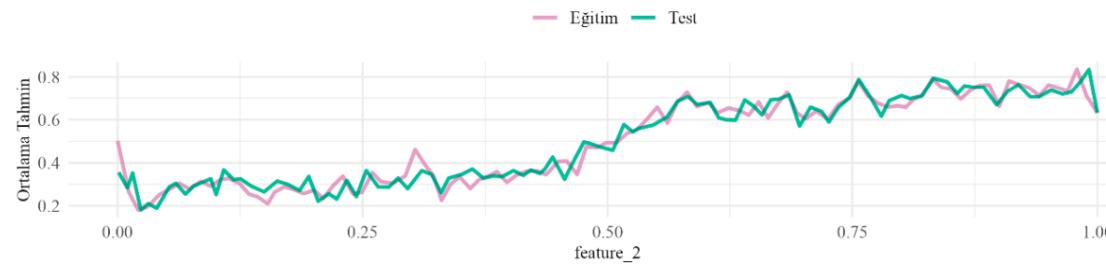
Yöntem	Yığın Sayısı							
	3		5		10		15	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,7930	1	0,6521	0	0,7834	0	0,7108	1
HDDM-W	0,7930	1	0,8188	1	0,7794	2	0,7955	2
KSWIN	0,8383	3	0,8614	5	0,8219	10	0,8170	12
PH	0,7930	1	0,8188	1	0,7834	0	0,7108	1
DDM	0,7930	1	0,8427	2	0,7834	0	0,7256	1
EDDM	0,8383	3	0,8614	5	0,7834	0	0,8164	12
PDD	0,7927	0	0,8371	1	0,7831	1	0,7320	2
PDD2	0,8378	3	0,8527	3	0,8064	3	0,7320	2
Test	0,7912		0,8306		0,6524		0,6883	
Eğitim	0,9985		0,9986		0,9993		0,9985	
Eğitim-Test	10666	2667	8000	2000	5332	1334	4000	1000

PDD yöntemi düşük yığın sayılarında diğer istatistiksel testlere benzer yer yer daha iyi sonuçlar verse de yığın sayısı 20, 25 ve 30'da diğer istatistiksel testlere göre kavram kayması tespiti ve model performansı bakımından düşük sonuçlar vermiştir.

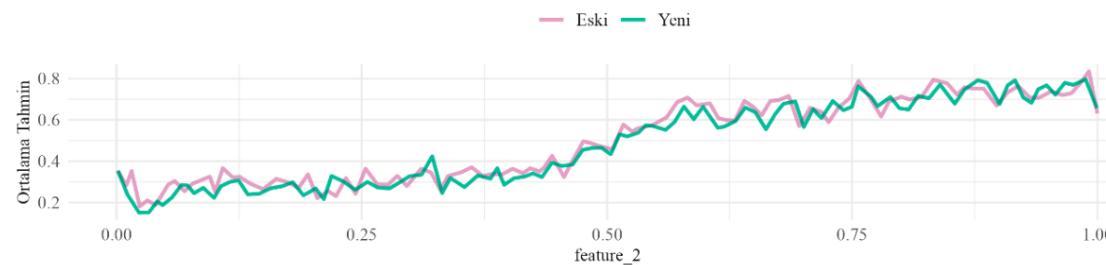
Yöntem	Yığın Sayısı					
	20		25		30	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,6809	1	0,6795	1	0,6267	1
HDDM-W	0,7615	3	0,7342	3	0,7536	6
KSWIN	0,8144	17	0,8023	19	0,7948	22
PH	0,7033	1	0,6949	1	0,6267	1
DDM	0,7290	3	0,6870	2	0,6267	1
EDDM	0,8072	16	0,8006	24	0,7957	24
PDD	0,6639	1	0,6203	1	0,5099	0
PDD2	0,7562	4	0,7537	4	0,7473	6
Test	0,6692		0,5584		0,4799	
Eğitim	0,9984		0,9985		0,9978	
Eğitim-Test	3200	800	2666	667	2285	572

PDD2 ise HDDM-W'ye benzer sayıda kavram kayması tespit etmiştir

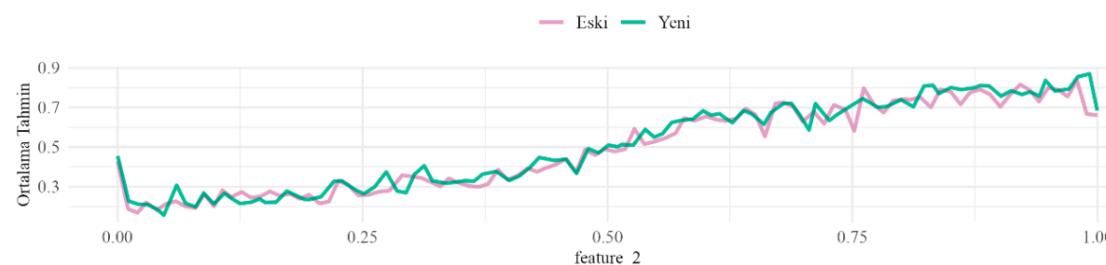
Test - Doğruluk: 0.83 - PDI: 0.28 - L2Der: 1.42 - L2: 0.04



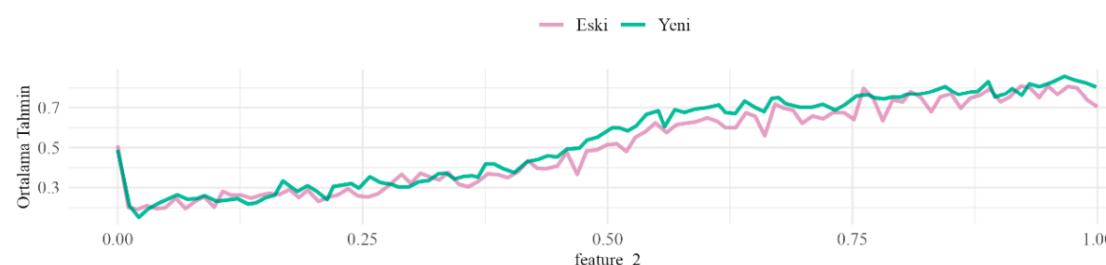
İndeks: 1 - Doğruluk: 0.79 - PDI: 0.32 - L2Der: 0.4 - L2: 0.04



İndeks: 3 - Doğruluk: 0.83 - PDI: 0.37 - L2Der: 1.07 - L2: 0.04

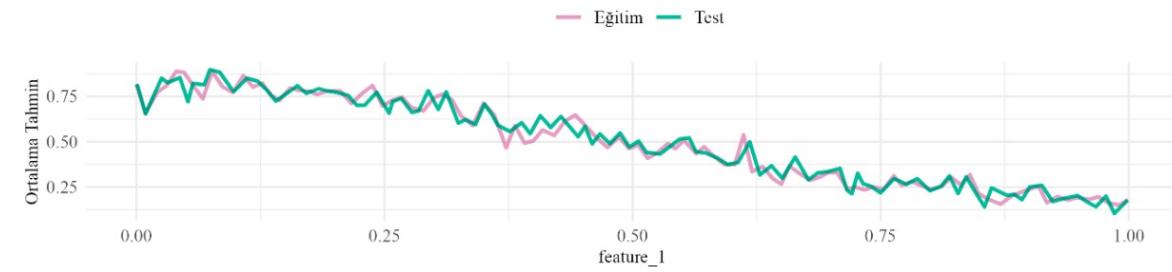


İndeks: 4 - Doğruluk: 0.86 - PDI: 0.32 - L2Der: 1.09 - L2: 0.05

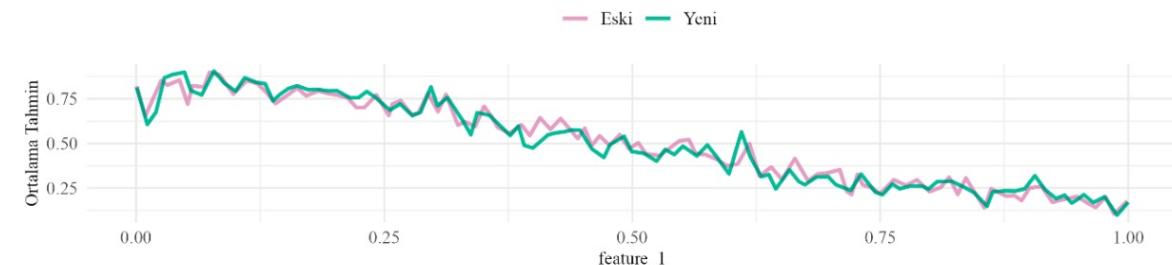


Y1ğın-5

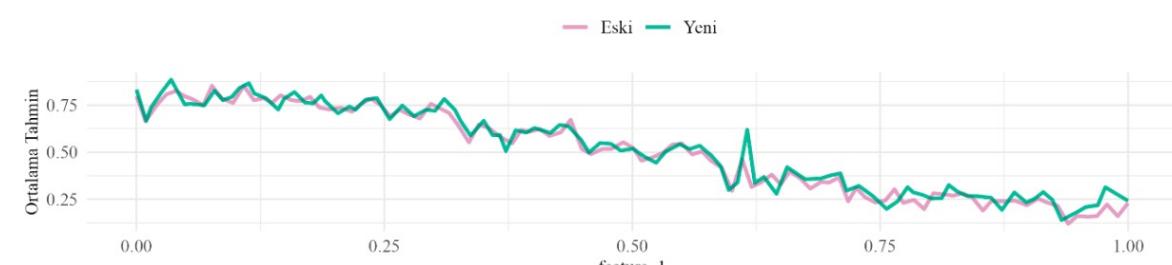
Test - Doğruluk: 0.65 - PDI: 0.26 - L2Der: 1.41 - L2: 0.04



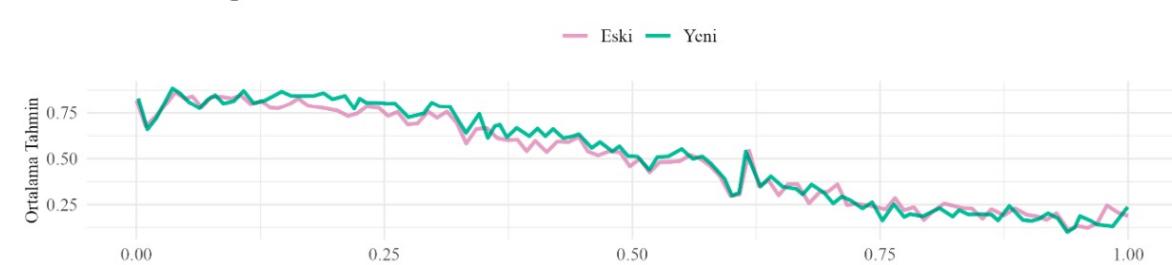
İndeks: 5 - Doğruluk: 0.81 - PDI: 0.22 - L2Der: 2 - L2: 0.04



İndeks: 6 - Doğruluk: 0.83 - PDI: 0.28 - L2Der: 1.54 - L2: 0.04



İndeks: 8 - Doğruluk: 0.86 - PDI: 0.31 - L2Der: 0.99 - L2: 0.04



Y1ğın-10

NOA (Weather)

- 1949-1999 döneminde Nebraska'daki Offutt Hava Üssü'nde sıcaklık, basınç, rüzgar hızı gibi sekiz farklı özellik ölçülmüştür. Amaç, belirli bir günde yağmur yağıp yağmayacağı tahmin etmektir (Elwell & Polikar., 2011).
- 18159 gözlem ve 8 bağımsız değişken.

Yığın Sayısı

Yöntem	3		5		10		15	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,7699	0	0,8170	2	0,7790	0	0,7898	1
HDDM-W	0,7739	2	0,8255	4	0,7883	4	0,7999	10
KSWIN	0,7741	3	0,8241	5	0,7916	10	0,8019	15
PH	0,7699	0	0,8170	2	0,7869	3	0,7898	2
DDM	0,7699	0	0,8241	5	0,7818	1	0,7957	2
EDDM	0,7741	3	0,8241	5	0,7916	10	0,8019	15
PDD	0,7704	1	0,8202	1	0,7839	1	0,7876	1
Test	0,8217		0,7722		0,7723		0,7744	
Eğitim	0,9998		0,9996		0,9996		1	
Eğitim-Test	9684	2422	7263	1816	4842	1211	3631	908

Yığın Sayısı

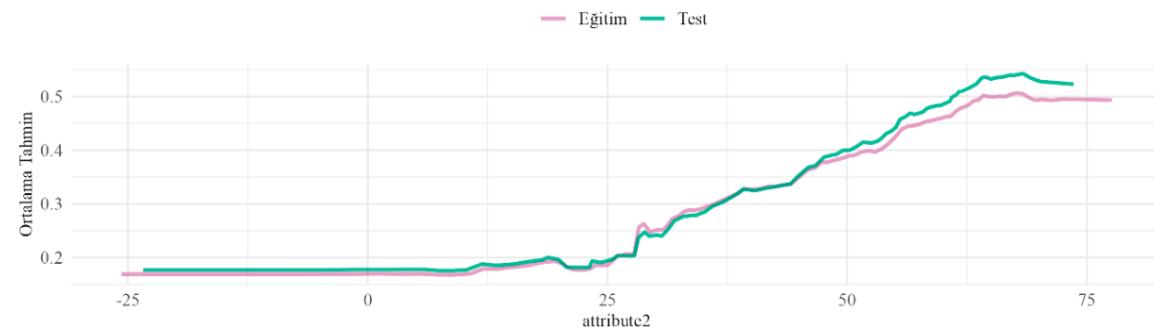
Yöntem	20		25		30	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,7783	4	0,7786	6	0,7780	2
HDDM-W	0,7835	10	0,7826	9	0,7826	9
KSWIN	0,7916	19	0,7912	24	0,7945	29
PH	0,7798	3	0,7719	1	0,7772	3
DDM	0,7914	20	0,7837	6	0,7831	8
EDDM	0,7914	20	0,7931	25	0,7940	30
PDD	0,7676	0	0,7750	1	0,7723	2
Test	0,8338		0,8320		0,8154	
Eğitim	0,9997		0,9996		0,9995	
Eğitim-Test	2904	727	2420	606	2075	519

PDD yöntemi ile genel olarak PH testine benzer sonuçlar elde edilmiştir. Bu iki test diğer testlere göre daha az sayıda kavram kayması tespit etmiştir.

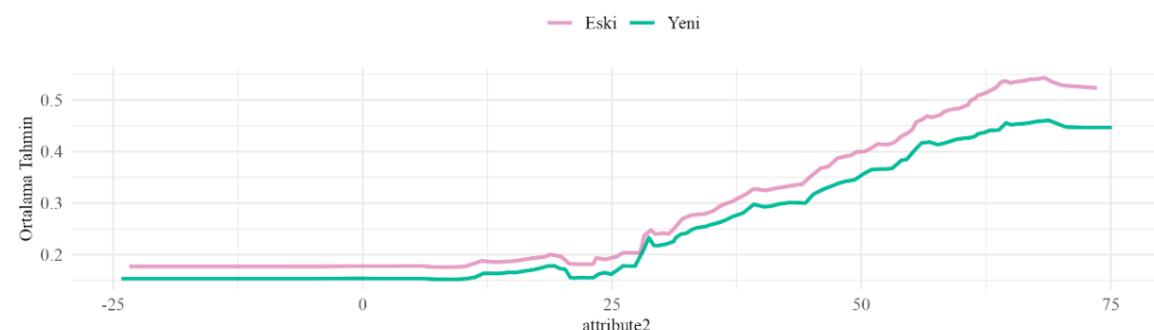
KSWIN ve EDDM testleri her yığında kavram kayması tespit etmiştir.

Ancak modelin çok daha az güncellendiği diğer istatistiksel yöntemlerden yüksek model performansına ulaşılmış olsa da bu fark maksimum 0,02 seviyelerinde kalmıştır.

Test - Doğruluk: 0.77 - PDI: 0.06 - L2Der: 0.01 - L2: 0.15

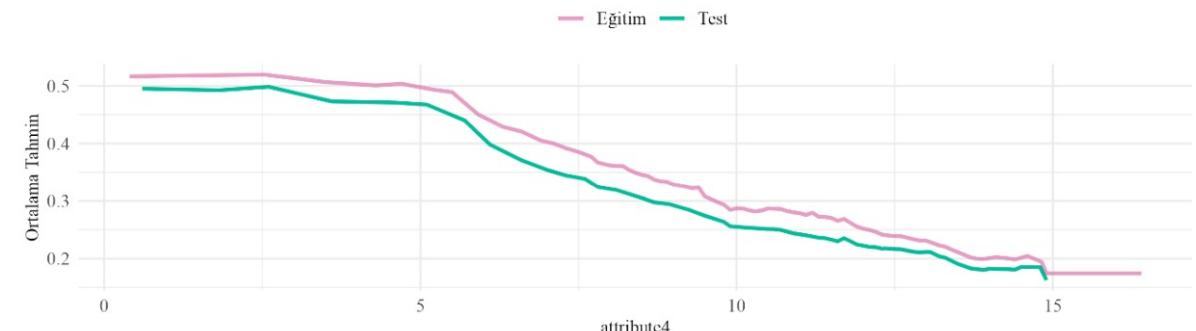


İndeks: 2 - Doğruluk: 0.79 - PDI: 0.08 - L2Der: 0.03 - L2: 0.42

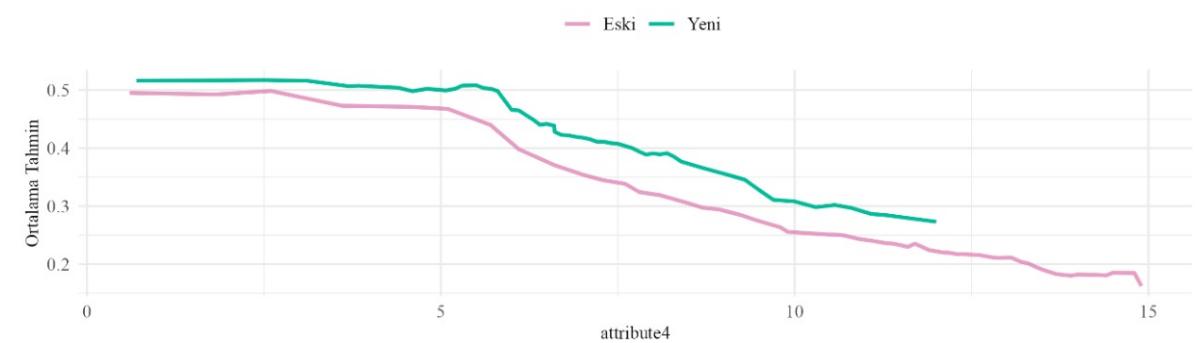


Y1ğın-5

Test - Doğruluk: 0.77 - PDI: 0.05 - L2Der: 0.04 - L2: 0.12



İndeks: 15 - Doğruluk: 0.74 - PDI: 0.11 - L2Der: 0.07 - L2: 0.16



Y1ğın-15

Ozone

- Ozon Seviyesi Tespit veri setine dayanmaktadır (Zhang & Fan, 2008)
- Houston bölgesindeki çeşitli hava kirleticileri ve meteorolojik bilgileri tanımlayan bağımsız değişkenler bulunmaktadır.
- 2536 gözlem ve 72 bağımsız değişken.
- Yanıt değişkeninde 2374 adet 0 ve 160 adet 1 bulunmaktadır ve veri oldukça dengesizdir.

Yıgın Sayısı

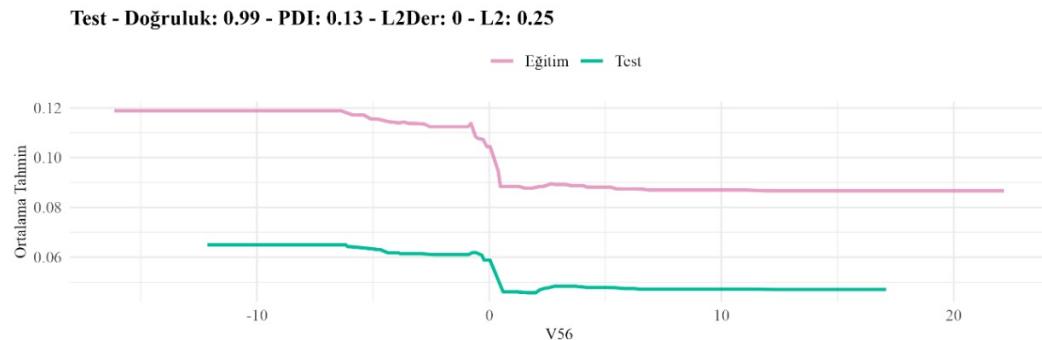
Yöntem	Yıgın Sayısı							
	3		5		10		15	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,9548	0	0,9607	0	0,9508	0	0,9463	1
HDDM-W	0,9548	0	0,9607	0	0,9508	0	0,9463	1
KSWIN	0,9548	0	0,9599	1	0,9529	1	0,9463	3
PH	0,9548	0	0,9607	0	0,9508	0	0,9457	0
DDM	0,9538	2	0,9607	0	0,9513	2	0,9494	8
EDDM	0,9548	1	0,9599	1	0,9524	2	0,9452	2
PDD	0,9548	1	0,9607	0	0,9502	1	0,9383	1
Test	0,9764		0,9569		0,9941		0,8516	
Eğitim	1,0000		1,0000		1,0000		1,0000	
Eğitim-Test	1351	338	1013	254	675	169	506	127

İstatistiksel testler diğer veri setlerinde olduğundan daha az sayıda kavram kayması tespit etmiştir. Bunun bir nedeni de verinin dengesiz olmasıdır.

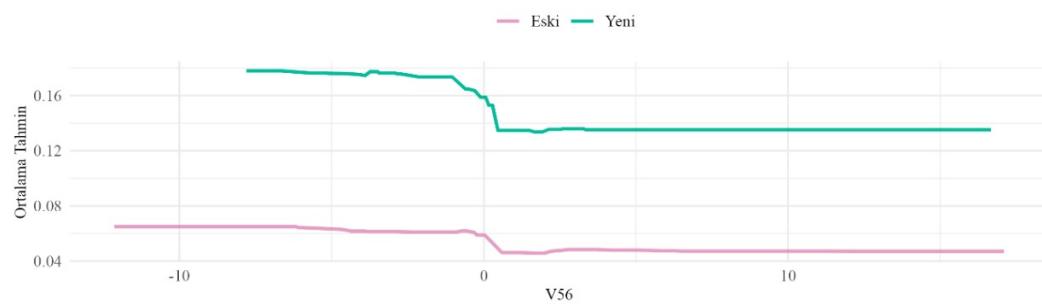
Yıgın Sayısı

Yöntem	Yıgın Sayısı					
	20		25		30	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,9427	1	0,9391	1	0,9374	0
HDDM-W	0,9427	1	0,9391	1	0,9401	1
KSWIN	0,9383	4	0,9395	6	0,9437	3
PH	0,9422	0	0,9395	0	0,9374	0
DDM	0,9427	2	0,9409	3	0,9437	3
EDDM	0,9441	3	0,9353	3	0,9437	3
PDD	0,9422	0	0,9405	1	0,9445	2
Test	0,9806		1		1	
Eğitim	1		1		1	
Eğitim-Test	404	102	337	85	289	73

PDD yöntemi diğer yöntemlerle benzer model performansına sahip olduğu gözlemlenmektedir.



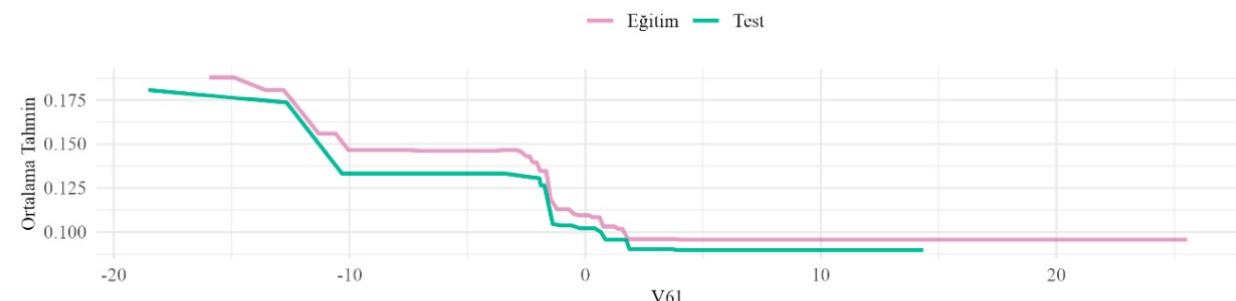
İndeks: 1 - Doğruluk: 0.79 - PDI: 0.35 - L2Der: 0.01 - L2: 0.48



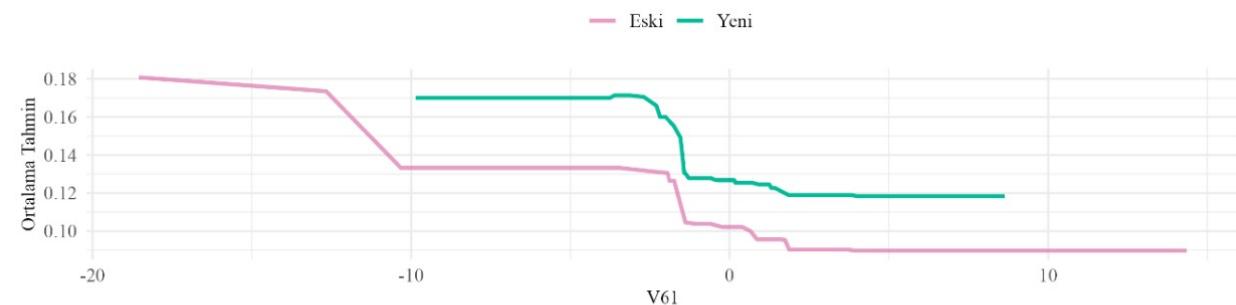
Yıgın-5

Farklı yıkın sayılarının hepsinde, farklı bağımsız değişkenler en önemli değişken olarak bulunmuştur.

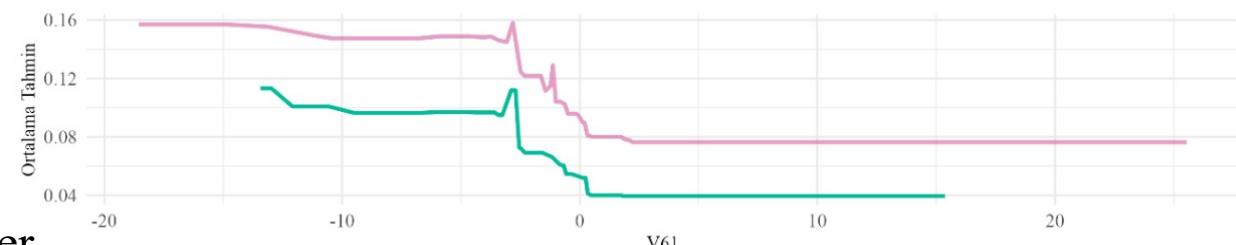
Test - Doğruluk: 1 - PDI: 0.09 - L2Der: 0.01 - L2: 0.05



İndeks: 3 - Doğruluk: 0.92 - PDI: 0.13 - L2Der: 0.01 - L2: 0.14



İndeks: 6 - Doğruluk: 1 - PDI: 0.14 - L2Der: 0.02 - L2: 0.23



Yıgın-30

Elec2

- Arz ve talep tarafından etkilenen fiyatlara sahip Avustralya Yeni Güney Galler Elektrik Piyasası'na ait bilgileri içermektedir.
- 45075 gözlem bulunmaktadır.
- Literatürde modelleme sırasında 4 bağımsız değişken (*vicprice*, *vicdemand*, *nswprice*, *nswdemand*) modele dahil edilmiştir (Riso & Guerzoni,2022). Bu çalışmada da bu 4 bağımsız değişken kullanılmıştır.
- Deneyde PDD2 kullanılmıştır.

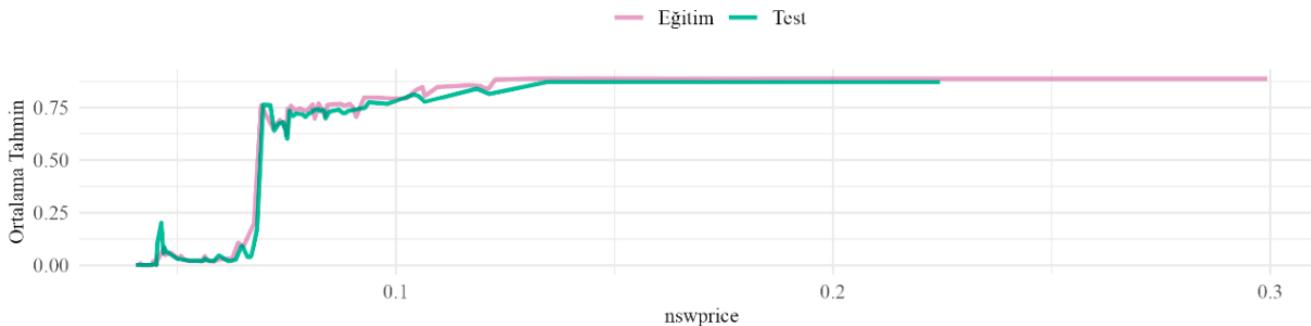
Yöntem	Yığın Sayısı							
	3		5		10		15	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,6780	3	0,675	5	0,7029	10	0,7187	15
HDDM-W	0,6780	3	0,675	5	0,7029	10	0,7187	15
KSWIN	0,6780	3	0,675	5	0,7029	10	0,7187	15
PH	0,6780	3	0,675	5	0,6992	8	0,7122	9
DDM	0,6780	3	0,675	5	0,7029	10	0,7187	15
EDDM	0,6780	3	0,675	5	0,7029	10	0,7187	15
PDD2	0,7018	0	0,6671	2	0,6947	6	0,723	4
Test	0,6485		0,7381		0,8107		0,8368	
Eğitim	0,8912		0,8824		0,8424		0,8391	
Eğitim-Test	24166	6042	18124	4532	12083	3021	9062	2266

PH testi haricinde diğer klasik istatistiksel testler genel olarak her yiğin sayısı için her yiğında kavram kayması tespit etmiş ve modeli güncellemiştir.

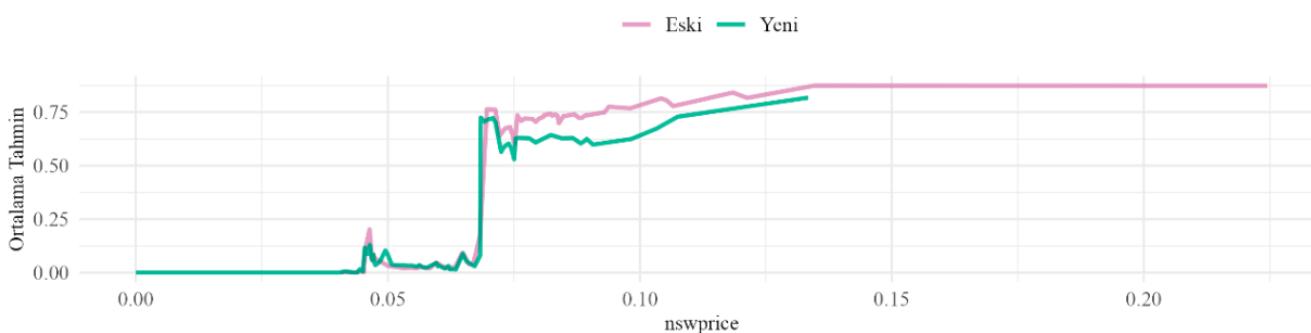
Yöntem	Yığın Sayısı					
	20		25		30	
	Doğruluk	Kayma	Doğruluk	Kayma	Doğruluk	Kayma
HDDM-A	0,7243	20	0,7337	25	0,7374	30
HDDM-W	0,7243	20	0,7337	25	0,7374	30
KSWIN	0,7243	20	0,7337	25	0,7374	30
PH	0,7103	11	0,7219	15	0,7368	16
DDM	0,7250	16	0,7337	25	0,7253	17
EDDM	0,7243	20	0,7337	25	0,7374	30
PDD2	0,7180	4	0,7278	6	0,7413	2
Test	0,8142		0,7612		0,804	
Eğitim	0,8367		0,836		0,8418	
Eğitim-Test	7249	1813	6041	1511	5178	1295

PDD2 ise çok daha az sayıda kavram kayması tespit etmiştir ancak model performansı açısından diğer testlere çok yakın ya da daha yüksek sonuçlar vermiştir.

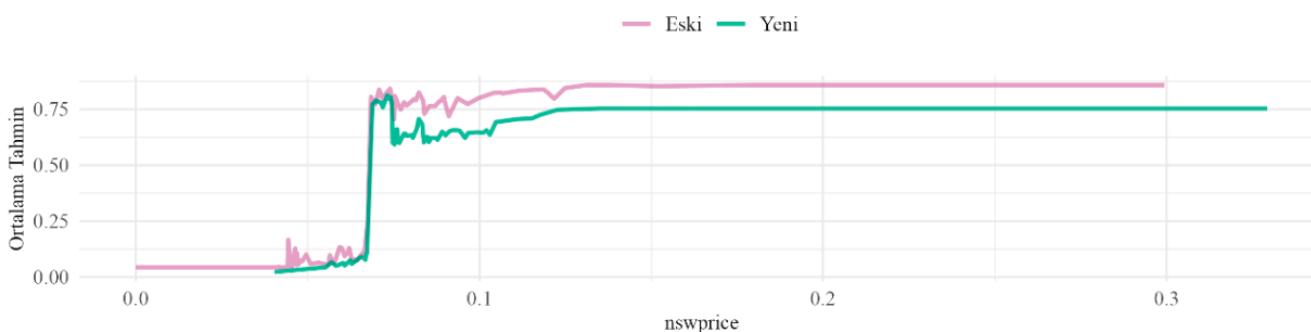
Test - Doğruluk: 0.8 - PDI: 0.26 - L2Der: 2.43 - L2: 0.01



İndeks: 1 - Doğruluk: 0.78 - PDI: 0.47 - L2Der: 11.71 - L2: 0.03



İndeks: 5 - Doğruluk: 0.76 - PDI: 0.55 - L2Der: 2.65 - L2: 0.06



Friedman (Regresyon)

- Sentetik olarak r i v e r paketi kullanılarak oluşturulmuştur (Ikonomovska vd., 2011).
- 20.000 gözlem ve 10 bağımsız değişken.
- PDD ve PDD2 önceki çalışmalarında olduğu gibi burada da benzer şekilde kullanılmıştır.
- Bu örnekte PDI değişkeni genel olarak 0 değerini aldığı için, L2 ve L2Der'de %10'luk bir atış olması kavram kayması olarak değerlendirilmiş ve PDD3 olarak isimlendirilmiştir.

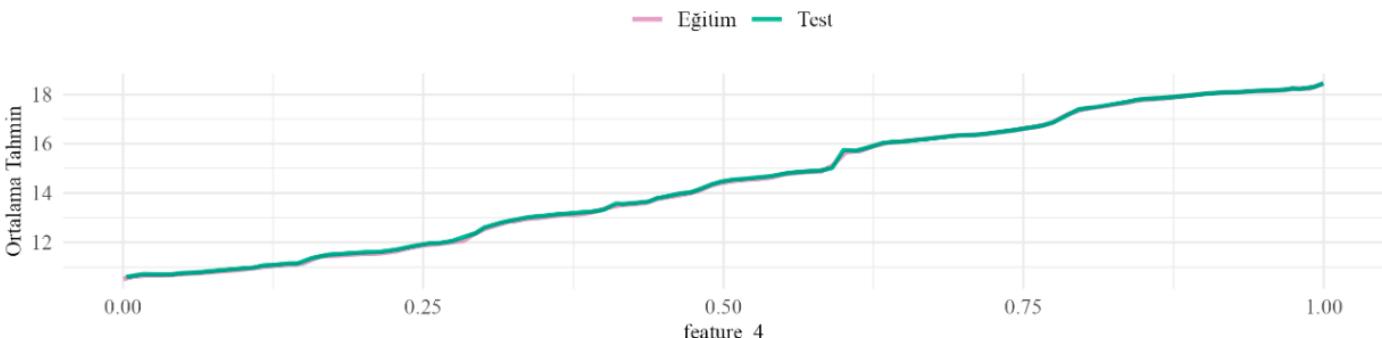
Yöntem	Yıgın Sayısı							
	3		5		10		15	
	RMSE	Kayma	RMSE	Kayma	RMSE	Kayma	RMSE	Kayma
KSWIN	2,637	0	2,049	0	2,082	0	1,964	0
PH	2,637	0	2,049	0	2,082	0	1,964	0
PDD	2,637	0	2,049	0	2,082	0	1,927	1
PDD2	2,606	2	1,988	2	2,039	1	1,927	1
PDD3	2,602	3	1,886	4	2,016	5	1,841	8
Test	1,609		1,583		1,639		1,778	
Eğitim	0,694		0,704		0,727		0,759	
Eğitim-Test	10666	2667	8000	2000	5332	1334	4000	1000

Yöntem	Yıgın Sayısı					
	20		25		30	
	RMSE	Kayma	RMSE	Kayma	RMSE	Kayma
KSWIN	1,998	0	2,109	0	2,098	0
PH	1,998	0	2,109	0	2,098	0
PDD	1,960	1	2,042	1	2,056	2
PDD2	1,942	2	2,016	2	2,043	2
PDD3	1,885	11	1,944	10	1,943	14
Test	1,775		1,705		1,745	
Eğitim	0,765		0,798		0,810	
Eğitim-Test	3200	800	2666	667	2285	572

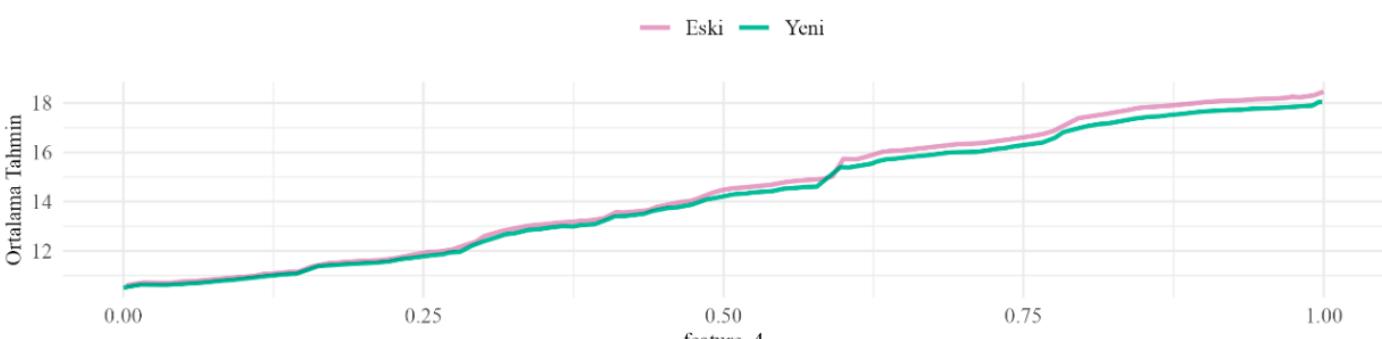
Her istatistiksel test sürekli değişkenler için uygun olmadığından istatistiksel testlerden KSWIN ve PH testleri kullanılmıştır.

RMSE (Root Mean Square Error) metriği üzerinden kayma tespit edilmeye çalışıldığından ise bu iki test herhangi bir kayma tespit etmemişken, PDD'ler kavram kayması tespitinde ve model performansında daha başarılı olmuşlardır.

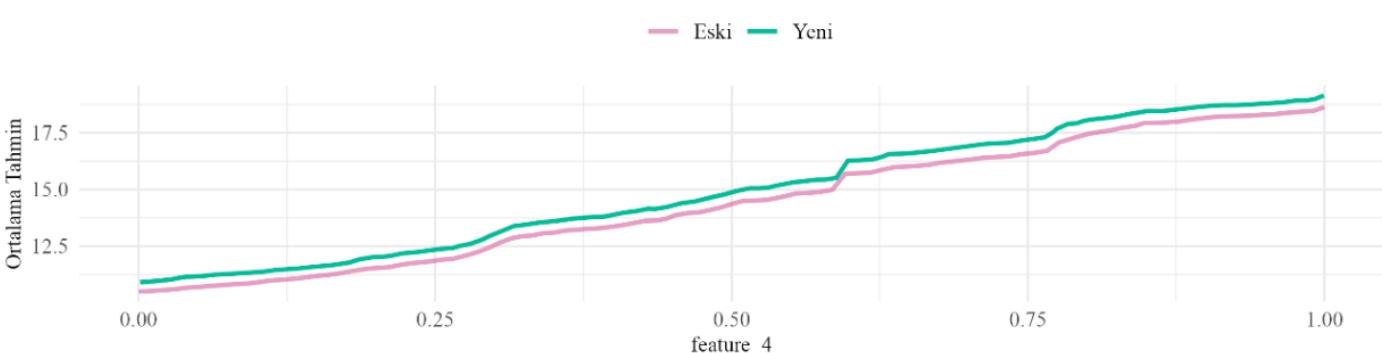
Test - RMSE: 1.78 - PDI: 0 - L2Der: 0.59 - L2: 0.05



İndeks: 1 - RMSE: 1.82 - PDI: 0 - L2Der: 0.68 - L2: 0.26



İndeks: 13 - RMSE: 1.67 - PDI: 0 - L2Der: 1.31 - L2: 0.52



Avantajları

- Açıklanabilirlik.
- Regresyon ve sınıflandırma problemleri için uygun.
- Az sayıda ancak tutarlı kavram kayması tespiti.

Dezavantajları

- Hesaplama maliyeti.
- Kategorik değişkenlerde uygulanamama.
- Tek bir değişkeni baz alma.
- Yanıt değişkenin 2'den fazla seviyeye sahip olması.
- Bağımsız değişkenler arasındaki ilişki.

Referanslar