

Bölüm 17

MAKİNE ÖĞRENİMİNDE ÖZİNİTELİK SEÇME YÖNTEMLERİNİN KULLANIMI: GÜNCEL PYTHON UYGULAMALARI

Muhammed KOTAN¹
Ömer Faruk SEYMEN²

GİRİŞ

Veri ön işlemede önemli adımlardan biri olan öznelik seçimi (Feature Selection - FS), model performansını iyileştirmek ve hesaplama karmaşıklığını azaltmak için bir veri kümesinden en uygun özneliklerin belirlenmesini içermektedir. FS için önerilen teknikler, en ilgili öznelikleri seçerek doğruluğu artırmayı, aşırı uyumu azaltmayı ve makine öğrenimi modellerinin verimliliğini artırmayı amaçlamaktadır.

FS teknikleri, makine öğrenmesi modellerinin performansının ve yorumlanabilirliğinin geliştirilmesinde önemli bir rol oynamaktadır. Ancak verinin boyutu ve karmaşıklığı arttıkça, FS problemleri daha zor hale gelir ve kullanılan yöntemlerin doğruluğunu etkiler (Zaimoğlu vd., 2023). Güncel çalışmalar, çok çeşitli FS yöntemleri ve bunların farklı alanlardaki uygulamaları hakkında değerli bilgiler sağlayarak, bu alanda devam eden araştırma ve geliştirmeyi vurgulamaktadırlar.

En iyi öznelik alt kümesini bulma süreci, FS yöntemlerinde zorlu bir görevdir. Üç arama stratejisi yaygın olarak kullanılır: (i) tam arama, (ii) en kötü durumda tam arama olan rastgele arama ve (iii) sezgisel teknikler (Elminaam vd., 2021). FS için literatürde kullanılan teknikler üç kategoriye ayrılabilir: (i) filtre yöntemleri, (ii) sarmalayıcı yöntemler ve (iii) gömülü yöntemler (Elminaam vd., 2021; Liu ve Setino, 1995).

Öznelikler arasındaki iç ilişkilere dayanarak sıralama yapan, öğrenme sürecine bağlı olmadan çalışan tekniklere filtre tabanlı (filter) teknikler denir ve hızlı teknikler olarak değerlendirilirler. Sarmalayıcılar (wrapper), her seçilen öznelik alt kümesini değerlendirmek için belirli bir öğrenme algoritması kullanır ve sınıflandırma doğruluğunu artıran bir alt küme seçer (Elminaam vd.,

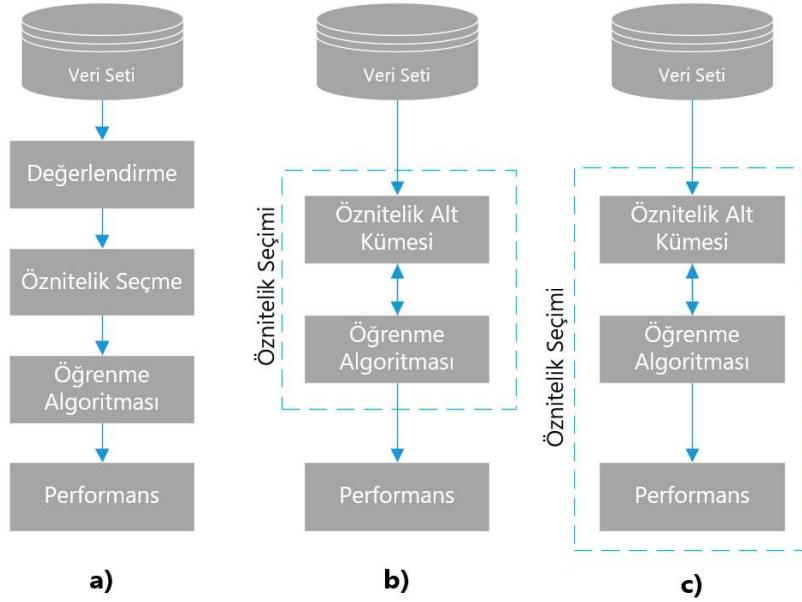
¹ Dr. Öğr. Üyesi; Sakarya Üniversitesi Bilgisayar ve Bilişim Bilimleri Fakültesi Bilişim Sistemleri Mühendisliği Bölümü. mkotan@sakarya.edu.tr ORCID No: 0000-0002-5218-8848

² Dr. Öğr. Üyesi; Sakarya Üniversitesi İşletme Fakültesi İşletme Bölümü. ofseymen@sakarya.edu.tr ORCID No: 0000-0003-2224-5546

2021; Nie vd., 2010). Doğruluk artmasıyla birlikte yavaşlık söz konusudur. Öğrenme hızı ve model performansını bir araya getiren gömülü yöntem ise FS'yi eğitim sürecine dahil eder (Abd Elaziz vd., 2022; Kotan, 2022). Bazı çalışmalar dördüncü bir hibrit modeli tanıtır (Abd Elaziz vd., 2022; Kotan, 2022; Salcedo-Sanz vd., 2018). Bu üç FS yöntemi ve genel yapıları sırasıyla Şekil 1. ve Şekil 2.'de gösterilmiştir.



Şekil 1: Öznitelik Seçimi Yöntemleri



Şekil 2: Öznitelik Seçimi: a) Filtre, b) Sarmalayıcı, c) Gömülü

FS, model için daha az öznelikle daha iyi sonuçlara yol açabilir. Bu noktada seçilen özneliklerin doğruluğu, ilgisiz özneliklerin model performansını olumsuz etkileyebileceğinden dolayı kritik bir öneme sahiptir. Günümüzde, büyük ve yüksek boyutlu veri kümelerine kolay erişim olmasına rağmen, işleme zorlukları da model performansını olumsuz etkileyebilmektedir. Bu nedenle, FS gerçekleştirirken veri kümesinin detaylı bir anlayışına sahip olmak ve doğru yöntemi seçmek esastır.

Veri işlemeyle birlikte karşımıza çıkan FS süreci geniş bir uygulama alanı çeşitliliğine sahiptir. Görüntü işleme (Pham vd., 2020), finansal veri analizi (Simumba vd., 2022), doğal dil işleme (Ahmad vd., 2015), tıbbi veri (Oğur vd., 2023), müşteri kaybı tahmini (Al-Shourbaji vd., 2022), enerji ile ilgili çalışmalar (Salcedo-Sanz vd., 2018) gibi çeşitli alanlar bunlardan bazılarıdır. Ayrıca FS tekniklerinin incelenmesine dayalı son yıllarda birçok inceleme çalışmaları da yapılmaktadır (Venkatesh ve Anuradha, 2019; Khaire ve Dhanalakshmi, 2022).

Son yıllarda sıklıkla artan yeni teknikler ve bu tekniklerin kullanımını içeren çalışmalardan dolayı geliştiriciler genellikle kodlarını açık erişim yapmakta ve araştırmacıların kullanımına sunmaktadırlar. FS tekniklerinin kullanımını ve test edilebilirliğini kolaylaştırmaya yönelik yeni kütüphaneler ve modüller ise özellikle kod karmaşasını basitleştirdiğinden ve kolay kullanımından dolayı tercih edilmektedir. Çalışmanın devamında Python tabanlı Mealpy (Van Thieu ve Mirjalili, 2023) ve Mafese (Metaheuristic Algorithms for Feature Selection) (Van Thieu vd., 2023) kütüphanelerinin FS amaçlı kullanımlarından bahsedilecek ve örnek sonuçlar gösterilecektir.

PYTHON TABANLI MEALPY VE MAFESE KÜTÜPHANELERİ

Mafese, FS amacıyla Python dilinde geliştirilmiş çeşitli strateji ve yaklaşımları içeren bir dizi öznelik seçimi algoritması sağlayan kütüphanelerden biridir. Kütüphane, FS tekniklerini ana hatlarıyla dört başlık altında gruplandırmaktadır:

- Filtre tabanlı FS: İstatistiksel yöntemler kullanarak ilgisiz öznelikleri filtreler.
- Sarmalayıcı tabanlı FS: Seçilen özneliklerin performansına göre iteratif olarak öznelik ekleyip veya çıkararak öznelik alt kümeleri oluşturur. Metasegisel teknikler de bu başlık altında değerlendirilmektedir.
- Gömülü tabanlı FS: Eğitim sürecinin kendisine öznelik seçimi mekanizmasını dahil eder. Lasso tabanlı ve ağaç tabanlı yöntemler içerir.
- Denetimsiz tabanlı FS: Özneliklerin sınıflandırma bilgisi olmadan birbirleriyle olan ilişkilerine göre seçim yapılır.

Mafese, 200'den fazla metasezgisel algoritma sunarak çeşitli yöntemlerin kullanımını kolaylaştırmaktadır. Kütüphane, açık kaynaklıdır ve GNU General Public License (GPL) V3 lisansına sahiptir. Aktif olarak geliştirilmekte ve yeni metasezgisel algoritmalar ve özellikler eklenmektedir. Tablo 1.'de Mafese kütüphanesinin bazı özellikleri özetlenmiştir.

Tablo 1: Mafese Kütüphanesi Özellikleri

Özellik	Değer
Sarmalayıcı Tabanlı Yöntem Sayısı	>200
Filtre Tabanlı Yöntem Sayısı	>15
Gömülü Tabanlı Yöntem Sayısı	>10
Denetimsiz Tabanlı Yöntem Sayısı	>= 4
Toplam veri seti sayısı	47 sınıflandırma ve 7 regresyon
Toplam Performans Metriği Sayısı	16 Sınıflandırma ve 45 Regresyon
Toplam Objective Fonksiyon Sayısı(fitness)	16 Sınıflandırma ve 45 Regresyon
Bağımlılıklar	numpy, scipy, scikit-learn, pandas, mealpy, permetrics, plotly, kaleido

Kütüphane içerisinde en çok bulunan teknikler FS için en iyi öznitelik arama stratejilerinden biri olan metasezgisel olanları içeren sarmalayıcı tekniklerdir. Metasezgisel teknikler, sezgisel bilgilere dayalı olarak arama sürecini yönlendirerek optimal öznitelik alt kümesini keşfetme konusunda etkili ve verimli yöntemler sunar (Demirci ve Yurtay, 2021). Metasezgisel yaklaşımlar, başlangıçta rastgele çözüm kümesini iteratif olarak değiştirerek yöntemin yapısına bağlı olarak en iyi çözümleri arar (Oğur vd., 2023).

Mealpy ise karmaşık optimizasyon problemlerini çözmek için tasarlanmış metasezgisel optimizasyon algoritmaları koleksiyonu sunar. Son yıllarda, metasezgisel yaklaşımlar yaygın bir şekilde benimsenmiştir ve en iyi öznitelik alt kümesini bulma süreci bir optimizasyon problemi olarak kabul edilebilir. Mealpy içerisinde sunulan tekniklerin kategorizasyonu Şekil 3.'de sunulmuştur.



Şekil 3: Mealpy Optimizer Modülü

Metasezgisel algoritmalar, benchmark fonksiyonlarından makine öğrenimi modellerinin parametre optimizasyonuna kadar çeşitli pratik problemleri çözmeye oldukça etkilidir (Van Thieu ve Mirjalili, 2023; Demirci vd., 2023). Mealpy, kullanıcıların bireysel ihtiyaçlarına göre adaptasyon ve özelleştirme yapmalarını kolaylaştırarak, soruna en uygun algoritmayı seçmelerine imkân tanır. Bu sayede, makine öğrenimi, mühendislik ve iş dünyası gibi çeşitli alanlarda karmaşık optimizasyon problemlerini çözmek için kullanışlı bir kütüphanedir.

Mafese, metasezgisel tabanlı teknikleri FS amaçlı kullanım için Mealpy kütüphanesinde bulunan yöntemlerden kullanıcılara belirli algoritmaları seçme, değerlendirme metriklerini özelleştirme ve ayarlamayı kolaylaştırma gibi özelliklerle çeşitli FS teknikleriyle deneme yapma imkânı tanır. Dahası, popüler makine öğrenimi algoritmalarını FS sürecine entegre ederek makine öğrenimi yeteneklerini kapsar. Bu entegrasyon, seçilen özneliklerin farklı makine öğrenimi modelleri üzerindeki etkisini değerlendirerek, optimal öznelik alt kümelerini seçmede yardımcı olur. Popüler makine öğrenmesi yöntemleri ve çeşitli puanlama metrikleri kullanarak araştırmacıların çalışmalarını hızlı ve kolay gerçekleştirmelerine olanak tanır.

DENEYİMSEL SONUÇLAR

Bu bölümde Mafese kütüphanesinden örnek uygulama ve sonuçlar gösterilecektir. Metasezgisel tekniklere yönelik çok sayıda algoritma yer aldığından dolayı içeriklerinden bazı güncel olanlar seçilecektir. Araştırmacılar, veri setlerini ve parametreleri belirleyebilirler. Ayrıca, eğitim-test oranlarını ve popülasyon boyutlarını belirleme seçenekleri de sunulmaktadır. Seçilen öznelikler ve elde edilen küçültülmüş veri seti de ekrana yazdırılmaktadır.

Örnekler kütüphanesi içi varsayılan parametreler kullanılarak çalıştırılmıştır. Veri seti ve tekniklerin yapısına göre farklı çalıştırmalarda sonuçlar değişkenlik gösterebilir.

Geliştiricileri, deneysel çalışmalara imkân sağlamak amacıyla önceden hazırlanmış çeşitli veri setlerini Mafese bünyesine katmıştır. Bu veri setleri, kaynak kodun "mafese.utils.data_loader" bölümü altında tanımlanan "get_dataset" fonksiyonu kullanılarak kolayca koda entegre edilebilir. Tablo 2.'de bu veri setlerinin kapsamlı bir özeti sunulmaktadır. Bu veri setleri, Mafese GitHub sayfasından indirilebilir (Van Thieu, 2023a). Ayrıca, kullanıcılar kendi veri setlerini değerlendirme esnekliğine de sahiptirler.

Tablo 2: Kütüphane İçinden Erişilebilir Örnek Veri Setleri

Sınıflandırma				Regresyon
Arrhythmia	Lymphography	Zoo	jain	boston-housing computer-hardware diabetes gauss-100-20 gauss-50-12 gauss-75-17 linnerud
Blood	Madelon	aggregation	liver	
BreastCancer	Monk1	aniso	moons.	
BreastEW	Monk2	appendicitis	mouse	
CongressEW	Monk3	balance	pathbased	
Digits	Sonar	banknote	seeds	
Glass	Soybean-small	blobs	smiley	
HeartEW	SpectEW	circles	varied	
Hill-valley	Tic-tac-toe	diagnosis_II	vary-density	
Horse	Vowel	ecoli	vertebral2	
Ionosphere	WaveformEW	flame	wdbc	
Iris	Wine	heart		

Mafese kütüphanesi dokümantasyonundan (Van Thieu, 2023b) örnek kodların kullanımı ile çeşitli FS tekniklerine yönelik sınıflandırma ve regresyon çalışmalarında kullanılabilir örnek sonuçlar aşağıda gösterilmiştir.

Filtre Tabanlı

Filtre tabanlı FS için kullanılabilir yöntemler ve parametrik gösterimleri Tablo 3.'te gösterilmiştir.

Tablo 3: Filtre Tabanlı FS İçin Kütüphane İçinden Erişilebilir Yöntemler

	Parametre	Yöntem
Sınıflandırma	CHI	Ki Kare (Chi Square)
Sınıflandırma veya Regresyon	ANOVA MI KENDALL SPEARMAN POINT RELIEF RELIEF-F VLS-RELIEF-F	ANOVA F-skor Ortak Bilgi (Mutual Information) Kendall Tau korelasyon Spearman Rho korelasyon Nokta Çift Serili Korelasyon (Point-biserial correlation) Orijinal Relief Ağırlıklı ortalama Relief Çok Yüksek Ölçekli Relief (Very Large Scale Relief)
Regresyon	PEARSON	Pearson korelasyon

Filtre tabanlı FS için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 4.'te sunulmuştur.

Tablo 4: Filtre Tabanlı Örnek FS Sonuçları

Veri Seti	Yöntem	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
Iris	KENDALL	4	2	[2 3]
Arrhythmia	ANOVA	279	5	[111 148 237 240 250]
BreastCancer	CHI	10	5	[0 2 3 6 8]

Sarmalayıcı Tabanlı

Sarmalayıcı tabanlı yöntemler kütüphane içerisinde üç başlık altında toplanmıştır. Her başlık için açıklamalar ve örnek sonuçlar aşağıda sunulmuştur.

1) Özyineli (Recursive)

Özyineli modülünde geçerli desteklenen tahminleyici (estimator) parametresi için kullanılabilir modeller ve parametrik gösterimleri Tablo 5.'te sunulmuştur.

Tablo 5: Özyineli Tabanlı FS İçin Kütüphane İçinden Kullanılabilir Modeller

Parametre	Yöntem
svm	Destek Vektör Makineleri (Support Vector Machine kernel = 'linear')
rf	Rastgele Ormanlar (Random Forest)
adaboost	AdaBoost
xgb	Gradient Boosting
tree	Ekstra Ağaçlar (Extra Trees)

Özyineli modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 6’da sunulmuştur.

Tablo 6: Özyineli tabanlı Örnek FS Sonuçları

Veri Seti	Yöntem	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
BreastCancer	rf	10	5	[2 3 5 6 7]
BreastCancer	svm	10	5	[1 3 6 7 9]
Ionosphere	adaboost	34	3	[5 20 27]

2) Sıralı (İleriye Doğru-Geriye Doğru)

Sıralı seçim modülü ileri ve geri yönlerde olmak üzere iki farklı yön (direction) parametresi içermektedir. Sıralı modülünde geçerli desteklenen tahminleyici (estimator) parametresi için kullanılabilir modeller ve parametrik gösterimleri Tablo 7’de sunulmuştur.

Tablo 7: Sıralı Tabanlı FS İçin Kütüphane İçinden Kullanılabilir Modeller

Parametre	Yöntem
knn	k en yakın komşu (k-nearest Neighbors)
svm	Destek Vektör Makineleri (Support Vector Machine)
rf	Rastgele Ormanlar (Random Forest)
adaboost	AdaBoost
xgb	Gradient Boosting
tree	Ekstra Ağaçlar (Extra Trees)
ann	Yapay Sinir Ağları (Artificial Neural Network (Multi-Layer Perceptron))

Sıralı modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 8’de sunulmuştur.

Tablo 8: Sıralı Tabanlı Örnek FS Sonuçları

Veri Seti	Yöntem	Yön	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
Ionosphere	knn	İleriye Doğru	34	3	[2 4 5]
Ionosphere	ann	İleriye Doğru	34	3	[0 3 4]
HeartEW	xgb	Geriye Doğru	13	5	[1 2 7 10 12]
Sonar	Tree	Geriye Doğru	60	2	[36 48]

3) Metasezgisel (Mha)

Mevcut desteklenen metasezgisel tekniklerin listesi (Van Thieu, 2023c) adresinden görüntülenebilir. Burada desteklenen sınıflandırıcılar Tablo 7 ile aynıdır. Güncel olarak 200'den fazla klasik ve yeni algoritmaları bünyesinde barındırmaktadır. Örnek kullanım için rastgele olarak seçilmiştir. Metasezgisel (mha) modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 9'da sunulmuştur.

Tablo 9: Metasezgisel Tabanlı Örnek FS Sonuçları

Veri Seti	Yöntem	Sınıflandırıcı	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
Arrhythmia	BaseGA	knn	279	26	[7 21 24 42 46 54 58 69 75 89 90 91 93 97 111 117 129 142]
Arrhythmia	OriginalMPA	svm	279	9	[172 177 178 179 186 218 219 233]
Arrhythmia	OriginalWOA	svm	279	4	[10 60 69 89 90 92 111 199 247]
BreastCancer	OriginalWOA	knn	10	2	[27 39 92 111]
Sonar	OriginalARO	rf	60	9	[1 3]

Gömülü Tabanlı

Gömülü teknikler altında iki farklı yaklaşım kullanılabilir.

1) Lasso

FS için Lasso tabanlı yöntemleri içermektedir. Lasso modülünde geçerli desteklenen tahminleyici (estimator) parametresi için kullanılabilir modeller ve parametrik gösterimleri Tablo 10'da sunulmuştur.

Tablo 10: Gömülü Tabanlı (Lasso) FS İçin Kütüphane İçinden Erişilebilir Yöntemler

Amaç	Parametre	Yöntem
Sınıflandırma Veya Regresyon	lasso	Lasso
Sınıflandırma	lr svm	Logistic Regression Destek Vektör Makinesi

Lasso modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 11’de sunulmuştur.

Tablo 11: Gömülü Tabanlı (Lasso) Örnek FS Sonuçları

Veri Seti	Yöntem	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
Arrhythmia	Lasso	279	85	[0 2 3 5 6 7 8 9 10 11 12 13 14 16 17 20 27 28 30 31 32 39 40 41 42 43 44 51 52 53 54 55 56 63 64 65 66 68 75 76 77 79 80 87 88 89 90 91 92 99 100 101 102 104 111 113 114 115 116 123 124 125 126 127 128 135 136 137 138 147 149 162 168 171 177 178 181 188 190 198 208 218 220 227 228 231 232 237 238 239 241 242 246 247 248 250 251 255 257 258 267 268 278]
Lymphography	Lr	18	16	[0 1 2 4 5 6 7 8 9 10 11 12 13 14 15 17]
Zoo	svm	16	9	[1 3 4 5 6 7 8 11 13]

2) Ağaç (Tree)

FS için Tree tabanlı yöntemleri içermektedir. Tree modülünde geçerli desteklenen tahminleyici (estimator) parametresi için kullanılabilir modeller ve parametrik gösterimleri Tablo 12’de sunulmuştur.

Tablo12: Gömülü Tabanlı (Tree) FS İçin Kütüphane İçinden Erişilebilir Modeller

Parametre	Yöntem
rf	Rastgele Ormanlar (Random Forest)
adaboost	AdaBoost
xgb	Gradient Boosting
tree	Ekstra Ağaçlar (Extra Trees)

Tree modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 13’de sunulmuştur.

Tablo 13: Gömülü Tabanlı (Tree) Örnek FS Sonuçları

Veri Seti	Yöntem	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
Arrhythmia	tree	279	151	[0 1 2 3 4 5 6 7 8 9 10 11 12 14 16 17 20 27 28 29 32 34 39 40 44 51 52 56 64 65 68 75 76 77 80 87 88 89 90 92 94 99 100 101 102 104 111 112 113 116 123 124 125 128 135 136 137 140 147 148 149 152 159 161 162 165 166 167 168 169 170 171 172 175 176 177 178 179 180 181 185 186 187 188 189 191 195 196 197 198 199 201 202 205 206 207 208 209 210 211 212 215 216 218 219 221 222 223 225 226 227 228 229 230 231 232 233 235 236 237 238 239 240 241 242 245 246 247 248 249 251 252 255 256 257 258 259 260 261 262 265 266 267 268 269 270 271 275 276 277 278]
Lymphography	rf	18	6	[1 9 12 13 14 17]
Zoo	xgb	16	5	[1 3 8 11 12]

Denetimsiz (Unsupervised) Tabanlı

Denetimsiz tabanlı yöntemleri içermektedir.

Tablo 14: Denetimsiz Tabanlı FS İçin Kullanılabilir Yöntemler

Parametre	Yöntem
VAR	Varyans Eşikdeğeri Yöntemi (Variance Threshold)
MAD	Ortalama Mutlak Sapma (Mean Absolute Difference)
DR	Dispersiyon Oranı (Dispersion Ratio)
MCL	Çoklu Bağntı (Multicollinearity)

Unsupervised modül için Mafese kullanılarak elde edilen örnek sonuçlar Tablo 15’te sunulmuştur.

Tablo 15: Denetimsiz Öğrenme Örnek FS Sonuçları

Veri Seti	Yöntem	Toplam Öznitelik Sayısı	Seçilen Öznitelik Sayısı	Seçilen Özniteliklerin İndeksi
diabetes	MAD	10	5	[0 1 2 3 8]
diabetes	MCL	10	5	[0 1 2 3 9]
Iris	MCL	4	2	[0 1]

SONUÇ

Öznitelik seçimi, özellikle büyük ve karmaşık veri setlerinde optimal öznitelik altkümesini seçmek için çözüm bulmada güçlü araçlardır. Bu yöntemler, çeşitli makine öğrenme problemlerinde uygulanabilirlik gösterir ve model performansını iyileştirmeye katkıda bulunur. Bu çalışmada Python tabanlı Mafese kütüphanesini kullanarak farklı FS yaklaşımlarına yönelik uygulamalar sunulmuştur. Böylece, başta metasezgisel teknikler olmak üzere çok fazla sayıda FS yaklaşım ve yöntemlerinin bir yerden sunulmasını ve FS problemlerindeki etkilerini ve performanslarını incelemeyi kolaylaştırmaktadır. Bu FS uygulama kolaylığı, araştırmacılara, uygulayıcılara ve alan uzmanlarına, veri setlerini kolayca yükleme, istenen FS yaklaşımını seçme, optimizasyon parametrelerini belirleme ve elde edilen sonuçları analiz etme imkânı sunar. Ayrıca, kullanıcıların FS sonuçlarını etkin bir şekilde yorumlamalarına ve karşılaştırmalarına yardımcı olacak açıklamalar sağlar.

Gelecekteki çalışmalar, daha geniş bir yöntem yelpazesini kapsayan ve ince ayarlı özelleştirmeye imkân tanıyan daha kapsamlı bir arayüz uygulamasının geliştirilmesini içerebilir. Bu geliştirmeler, araştırmacılara ve uygulayıcılara öznitelik seçimini etkili bir şekilde kullanma ve veri analizinde bilinçli karar verme yeteneği kazandırarak, öznitelik seçimi tekniklerinin çeşitli alanlarda yaygın olarak benimsenmesini teşvik edecektir.

REFERANSLAR

- Abd Elaziz, M., Ewees, A. A., Yousri, D., Abualigah, L., ve Al-qaness, M. A. A. (2022). Modified marine predators algorithm for feature selection: Case study metabolomics. *Knowledge and Information Systems*, 64(1), 261-287.
- Ahmad, S. R., Bakar, A. A., ve Yaakub, M. R. (2015). Metaheuristic algorithms for feature selection in sentiment analysis. *2015 Science and Information Conference (SAI)*, (pp. 222-226). London, United Kingdom: IEEE.
- Al-Shourbaji, I., Helian, N., Sun, Y., Alshathri, S., ve Abd Elaziz, M. (2022). Boosting ant colony optimization with reptile search algorithm for churn prediction. *Mathematics*, 10(7), 1031.
- Demirci, H., ve Yurtay, N. (2021). Effect of the chaotic crossover operator on breeding swarms algorithm. *Sakarya University Journal of Computer and Information Sciences*, 4(1), 120-130.
- Demirci, H., Yurtay, N., Yurtay, Y., ve Zaimoğlu, E. A. (2023). Electrical search algorithm: A new metaheuristic algorithm for clustering problem. *Arabian Journal for Science and Engineering*, 48(8), 10153-10172.
- Elminaam, D. S. A., Nabil, A., Ibraheem, S. A., ve Houssein, E. H. (2021). An efficient marine predators algorithm for feature selection. *IEEE Access*, 9, 60136-60153.
- Khaire, U. M., ve Dhanalakshmi, R. (2022). Stability of feature selection algorithm: A review. *Journal of King Saud University - Computer and Information Sciences*, 34(4), 1060-1073.
- Kotan, M. (2022) Duygu Analizi ve Dijital Dönüşüm Üzerine Etkileri. *Dijital Etkileşimler: Sektörel Yansımaları 2*, (pp. 59:82). Gazi Kitapevi.
- Liu, H., ve Setiono, R. (1995). Chi2: Feature selection and discretization of numeric attributes. *Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence*, (pp. 388-391). Herndon, VA, USA: IEEE Comput. Soc. Press.
- Nie, F., Huang, H., Cai, X., ve Ding, C. (2010). Efficient and robust feature selection via joint $\ell_2, 1$ -norms minimization. *Advances in neural information processing systems*, 23.
- Oğur, N. B., Kotan, M., Balta, D., Yavuz, B. Ç., Oğur, Y. S., Yuvacı, H. U., ve Yazıcı, E. (2023). Detection of depression and anxiety in the perinatal period using Marine Predators Algorithm and kNN. *Computers in Biology and Medicine*, 161, 107003.
- Pham, T. N., Tran, L. V., ve Dao, S. V. T. (2020). Early disease classification of mango leaves using feed-forward neural network and hybrid metaheuristic feature selection. *IEEE Access*, 8, 189960-189973.

- Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., ve García-Herrera, R. (2018). Feature selection in machine learning prediction systems for renewable energy applications. *Renewable and Sustainable Energy Reviews*, 90, 728-741.
- Simumba, N., Okami, S., Kodaka, A., ve Kohtake, N. (2022). Multiple objective metaheuristics for feature selection based on stakeholder requirements in credit scoring. *Decision Support Systems*, 155, 113714.
- Van Thieu, N., (2023a). *Metaheuristic Algorithms for FEature SElection Github*. <https://github.com/thieu1995/mafese/tree/main/mafese/data> adresinden 8 Ağustos 2023 tarihinde erişilmiştir.
- Van Thieu, N., (2023b). *Mafese*. <https://mafese.readthedocs.io/en/latest/index.html> adresinden 8 Ağustos 2023 tarihinde erişilmiştir.
- Van Thieu, N., (2023c). *Mealpy*. <https://github.com/thieu1995/mealpy> adresinden 8 Ağustos 2023 tarihinde erişilmiştir.
- Van Thieu, N., ve Mirjalili, S. (2023). MEALPY: An open-source library for latest meta-heuristic algorithms in Python. *Journal of Systems Architecture*, 139, 102871.
- Van Thieu, N., Nguyen, N. H., ve Heidari, A. A. (2023). Feature Selection using Metaheuristics Made Easy: Open Source MAFESE Library in Python, *Zenodo*.
- Venkatesh, B., ve Anuradha, J. (2019). A review of feature selection and its methods. *Cybernetics and Information Technologies*, 19(1), 3-26. <https://doi.org/10.2478/cait-2019-0001>
- Zaimoğlu, E. A., Yurtay, N., Demirci, H., ve Yurtay, Y. (2023). A binary chaotic horse herd optimization algorithm for feature selection. *Engineering Science and Technology, an International Journal*, 44, 101453.