

**T.C
HACETTEPE ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ÇOK KATEGORİLİ HASTALIK DURUMLARINDA TANISAL
MODELE YENİ BİR BELİRTEÇ EKLENMESİNİN TANI
PERFORMANSINDAKİ DEĞİŞİME ETKİSİNİN İNCELENMESİ**

Hande ŞENOL

**Biyoistatistik Programı
DOKTORA TEZİ**

**ANKARA
2019**

**T.C
HACETTEPE ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ÇOK KATEGORİLİ HASTALIK DURUMLARINDA TANISAL
MODELE YENİ BİR BELİRTEÇ EKLENMESİNİN TANI
PERFORMANSINDAKİ DEĞİŞİME ETKİSİNİN İNCELENMESİ**

Hande ŞENOL

**Biyoistatistik Programı
DOKTORA TEZİ**

**TEZ DANIŞMANI
Prof. Dr. A. Ergun KARAAĞAOĞLU**

ANKARA

2019

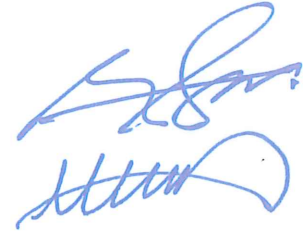
**ÇOK KATEGORİLİ HASTALIK DURUMLARINDA TANISAL MODELE
YENİ BİR BELİRTEÇ EKLENMESİNİN TANI PERFORMANSINDAKİ
DEĞİŞİME ETKİSİNİN İNCELENMESİ**

Hande ŞENOL

Danışman: Prof. Dr. A. Ergun KARAAĞAOĞLU

Bu tez çalışması 05.07.2019 tarihinde jürimiz tarafından “Biyoistatistik Programı”
nda doktora tezi olarak kabul edilmiştir.

Jüri Başkanı: Prof. Dr. C. Reha ALPAR
(Hacettepe Üniversitesi)
Üye: Prof. Dr. Mehtap AKÇİL OK
(Başkent Üniversitesi)
Üye: Prof. Dr. Meriç YAVUZ ÇOLAK
(Başkent Üniversitesi)
Üye: Prof. Dr. Erdem KARABULUT
(Hacettepe Üniversitesi)
Üye: Dr. Öğr. Üye. Sevilay KARAHAN
(Hacettepe Üniversitesi)



Bu tez, Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin
ilgili maddeleri uyarınca yukarıdaki jüri tarafından uygun bulunmuştur.

23 Temmuz 2019



Prof. Dr. Diclehan ORHAN

Enstitü Müdürü

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan **“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”** kapsamında tezimin aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. ⁽²⁾
- Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

26 / 07 / 2019

(İmza)

Hande ŞENOL

ⁱ“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir *. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir. Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

ETİK BEYAN

Bu çalışmadaki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi, görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu, kullandığım verilerde herhangi bir tahrifat yapmadığımı, yararlandığım kaynaklara bilimsel normlara uygun olarak atıfta bulunduğumu, tezimin kaynak gösterilen durumlar dışında özgün olduğunu, Prof. Dr. A. Ergun KARAAĞAOĞLU danışmanlığında tarafımdan üretildiğini ve Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Yönergesine göre yazıldığını beyan ederim.

(İmza)
Öğretim Görevlisi Hande ŞENOL

TEŞEKKÜR

Bu tezin hazırlanmasında, her zaman yanımda olan değerli danışman hocam sayın Prof. Dr. A. Ergun KARAAĞAOĞLU'na değerli zamanını bana ayırdığı için, özellikle de sabrını ve yardımlarını esirgemediği için;

Tez izleme komitesinde bulunan sayın Prof. Dr. C. Reha ALPAR ve sayın Prof. Dr. Mehtap AKÇİL OK hocalarıma bu teze olan katkılarından dolayı;

Tezin, başta benzetim ve uygulama aşamalarında olmak üzere hemen her kısmında vermiş olduğu destek ve yardımlardan ötürü ayrıca manevi olarak da önemli destek sağlayan, Mimar Sinan Güzel Sanatlar Üniversitesi'nden sayın Dr.Öğretim Üyesi E. Özge ÖZDAMAR hocama;

Her zaman yanımda olan, beni her konuda destekleyen ve yalnız bırakmayan sevgili aileme ve sevgili arkadaşım Fulya SARMAŞIK'a;

Tezin her aşamasında, yaşadığım sıkıntılarda yardımlarını esirgemeyen tüm Hacettepe Üniversitesi Biyoistatistik Anabilim Dalı öğretim üyeleri, araştırma görevlileri ve sayın Menekşe TARLA'ya;

Katkılarından ötürü Pamukkale Üniversitesi Biyoistatistik Anabilim Dalı'nda bulunan hocam sayın Prof. Dr. Beyza ŞAHİN'e;

Çok teşekkür ederim.

ÖZET

Şenol, H., Çok Kategorili Hastalık Durumlarında Tanısal Modele Yeni Bir Belirteç Eklenmesinin Tanı Performansındaki Değişime Etkisinin İncelenmesi, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Doktora Tezi, Ankara, 2019. Sağlık alanında, hasta ve sağlıklı bireylerin belirlenmesinde, tanı kestirim modellerinden elde edilen risklerin doğru değerlendirilmesi önemlidir. Bazı durumlarda sınıflamalar ikiden fazla kategoriden oluşabilir. İki kategorili sınıflandırmalarda en sık kullanılan yöntem olan ROC analizi yönteminden yola çıkarak çok kategorili sınıflandırma problemleri için HUM (Hypervolume Under The ROC Manifold) ve CCP - Doğru Sınıflandırma Olasılığı (Correct Classification Probability) yöntemleri geliştirilmiştir. Ancak, bu yöntemlerin var olan bir modele yeni bir belirteç eklendiğinde meydana gelen performans artışını ölçümlemede başarılı olmadığı düşünülmektedir. Bu nedenle, yeni bir belirteçin modele eklendiğinde oluşturulan yeni sınıflama sonucunun performansını gösteren iki performans ölçüsü geliştirilmiştir. Bunlar; NRI (Net Reclassification Improvement) ve IDI (Integrated Discrimination Improvement)'dır. Bu çalışmada, çok kategorili sınıflandırma modellerinde kullanılan performans ölçüleri HUM ve CCP'nin, model performansına etki eden belirteçlerin etkisini ölçen NRI ve IDI ile ilişkilerinin incelenmesi amaçlanmıştır. Bağımlı değişkenin farklı türde olduğu durumlar göz önüne alınarak (sıralı ve sırasız), belirteçlerin modele eklenme sırasının model performansına etkisi araştırılmıştır. İncelemeler, gerçek veri setleri ve bir benzetim çalışması ile yapılmıştır. Benzetim çalışmasında, hem belirteçlerin kendi arasında hem de belirteçler ve bağımlı değişken arasında olmak üzere; pozitif, negatif, düşük, orta ve kuvvetli seviyelerde ilişki yapıları ve sıralı yapıda bir bağımlı değişken tasarlanmıştır. Sonuçlarda, negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin performans sonuçlarında, HUM ve CCP yöntemleri ile NRI ve IDI arasında pozitif yönde ilişkiler olduğu görülmüştür. NRI ve IDI yöntemlerinin, HUM ve CCP yöntemlerinin aksine modele eklenen belirtecin sırasından etkilenmesinden dolayı, doğru sırada eklenen daha az sayıda belirteç ile doğru sınıflandırma performansının arttırılabileceği sonucu elde edilmiştir.

Anahtar Kelimeler: Çok kategorili sınıflandırma, HUM, CCP, NRI, IDI

ABSTRACT

Şenol, H. Investigation of the Change in Diagnostic Accuracy of Adding a New Marker to the Diagnostic Model in Multiclass Disease Status, Hacettepe University Graduate School of Health Sciences Doctor of Philosophy Thesis in Biostatistics, Ankara, 2019. Accurate assessment of the risks in the field of health, which are derived from diagnostic prediction models is important in identifying patients and healthy people. In some cases, classifications may consist of more than two categories. By using the ROC analysis method, which is the most commonly method in two-category classifications, some methods have been developed for multi-category classification problems. These are HUM (Hypervolume Under the ROC Manifold) and CCP (Correct Classification Probability). However, these methods are considered to be unsuccessful in measuring improvement in the performance of the model that occurs when a new biomarker is added to an existing model. Therefore, when a new biomarker is added to the model, two performance measures have been developed to show the performance of the new classification result. These are NRI (Net Reclassification Improvement) and IDI (Integrated Discrimination Improvement) methods. In this study, we aim to investigate the relationship between HUM, CCP measures, which are the performance measures used in multi-category classification models, and NRI, IDI which measure the effect of markers affecting model performance. Considering the type of dependent variable (ordinal and nominal), the effect of the order of adding the markers to the model performance was investigated. Investigations were made with real data sets and a simulation study. In the simulation study, we simulated data based on the correlation structure with low, medium and strong magnitude in both positive and negative directions for relationships between ordinal dependent variable and the markers and within the markers. According to the simulation results of negative correlation structures, it was seen that there was a positive relationship between HUM, CCP and NRI, IDI. As the NRI and IDI methods are affected by the order of the marker added to the model as opposed to the HUM and CCP methods, it is concluded that the correct classification performance can be improved with fewer markers added in the correct order.

Keywords: Multi-category classification, HUM, CCP, NRI, IDI

İÇİNDEKİLER

ONAY SAYFASI	iii
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI	iv
ETİK BEYAN SAYFASI	v
TEŞEKKÜR	vi
ÖZET	vii
ABSTRACT	viii
İÇİNDEKİLER	ix
SİMGELER ve KISALTMALAR	xi
ŞEKİLLER	xii
TABLolar	xiii
1. GİRİŞ	1
2. GENEL BİLGİLER	4
2.1. Üç ve Üzeri Kategori durumunda ROC Eğrisi Yöntemi (VUS ve HUM)	4
2.2. Doğru Sınıflandırma Olasılığı (Correct Classification Probability – CCP)	6
2.3. Net Yeniden Sınıflandırma İyileştirmesi (Net Reclassification Improvement - NRI)	7
2.4. Birleştirilmiş Ayrımsama İyileştirmesi (Integrated Discrimination Improvement - IDI)	7
2.5. NRI ve IDI'nın Gelişimi	7
3.GEREÇ VE YÖNTEM	11
3.1. HUM Değerlerinin Kestirimi	12
3.2. CCP Değerlerinin Kestirimi	13
3.3. NRI Değerlerinin Kestirimi	14
3.3.1. Kategorik NRI	15
3.3.2. Sürekli NRI	17
3.3.3. Ağırlıklandırılmış NRI (wNRI)	20
3.4. IDI kestirimi	23
3.5. Uygulamalar	24
3.5.1. Uygulama 1	25
3.5.2. Uygulama 2	30

3.5.3. Uygulama 3	31
4. BULGULAR	32
4.1. Uygulama 1 - Bağımlı değişkenin sıralı yapıda olduğu durum için örnek uygulama (Benzetim çalışması)	32
4.2. Uygulama 2 - Bağımlı değişkenin sırasız yapıda olduğu durum için örnek uygulama (gerçek veri seti)	61
4.3. Mikrodizin verisi için örnek uygulama	72
5.TARTIŞMA	80
6. SONUÇ VE ÖNERİLER	83
7. KAYNAKLAR	86
8. EKLER	
EK-1: Tez Çalışması Orijinallik Raporu	
EK-2: Dijital Makbuz	
9. ÖZGEÇMİŞ	

SİMGELER VE KISALTMALAR

ALL	Akut Lenfoblastik Lösemi
AML	Akut Miyeloid Lösemi
AUC	ROC Eğrisi Altında Kalan Alan
CCP	Correct Classification Probability - Doğru Sınıflandırma Olasılığı
FTI	Serbest Tiroksin İndeksi
HUM	Hypervolume Under The ROC Manifold – ROC Manifold Altında Kalan Çoklu Hacim
IDI	Integrated Discrimination Improvement - Birleştirilmiş Ayrımsama İyileştirmesi
NRI	Net Reclassification Improvement - Net Yeniden Sınıflandırma İyileştirmesi
ROC	Receiver Operating Characteristic - Alıcı İşlem Karakteristiği
T3	Triiyodotreonin
T4U	Tiroksin Kullanım Oranı
TSH	Tiroit Uyarıcı Hormon
TT4	Toplam Troksin
UCI	University of California, Irvine
VUS	Volume Under Surface - Yüzey Altında Kalan Hacim

ŞEKİLLER

Şekil	Sayfa
4.1. Pozitif ilişki yapısında benzetim yapılan verilerde, n=100 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	49
4.2. Negatif ilişki yapısında benzetim yapılan verilerde, n=100 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	50
4.3. Pozitif ilişki yapısında benzetim yapılan verilerde, n=500 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	51
4.4. Negatif ilişki yapısında benzetim yapılan verilerde, n=500 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	52
4.5. Pozitif ilişki yapısında benzetim yapılan verilerde, n=1000 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	53
4.6. Negatif ilişki yapısında benzetim yapılan verilerde, n=1000 örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler	54
4.7. Pozitif ilişki yapısında benzetim yapılan verilerde, n=100 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	55
4.8. Negatif ilişki yapısında benzetim yapılan verilerde, n=100 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	56
4.9. Pozitif ilişki yapısında benzetim yapılan verilerde, n=500 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	57
4.10. Negatif ilişki yapısında benzetim yapılan verilerde, n=500 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	58
4.11. Pozitif ilişki yapısında benzetim yapılan verilerde, n=1000 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	59
4.12. Negatif ilişki yapısında benzetim yapılan verilerde, n=1000 örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler	60

TABLOLAR

Tablo	Sayfa
3.1. Sürekli NRI Bileşenleri	19
3.2. Olay Görülen ve Görülmeyen Kişilerde Ağırlıklandırılmış Yeniden Sınıflama	21
3.3. Benzetim çalışması için OrdNor paketine yüklenen ve incelenmesi planlanan ilişki yapıları	26
3.4. Benzetim yoluyla elde edilmiş, pozitif ilişki yapısına sahip veri setinde elde edilmiş olan ilişkiler (Spearman ilişki katsayıları, r)	27
3.5. Benzetim yoluyla elde edilmiş, negatif ilişki yapısına sahip veri setinde elde edilmiş olan ilişkiler (Spearman ilişki katsayıları, r)	28
3.6. İncelenen modeller	29
4.1. Modelde bulunan belirteç sayısı ile incelenen modellerin performans sonuçları arasındaki ilişkiler	32
4.2. Pozitif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin, modele eklenen belirteç sayılarına göre performansları	34
4.3. Negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin, modele eklenen belirteç sayılarına göre performansları	35
4.4. İki belirteçli modellerin performans sonuçları arasındaki ilişkiler	37
4.5. Üç belirteçli modellerin performans sonuçları arasındaki ilişkiler	39
4.6. Dört belirteçli modellerin performans sonuçları arasındaki ilişkiler	40
4.7. Modelin ilk belirtecinin X1 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler	43
4.8. Modelin ilk belirtecinin X2 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler	44
4.9. Modelin ilk belirtecinin X3 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler	45
4.10. Modelin ilk belirtecinin X4 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler	46
4.11. Modelin ilk belirtecinin X5 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler	47
4.12. Tek belirteç ile kurulmuş olan modellerin performans sonuçları	61
4.13. İki belirteç ile kurulmuş olan modellerin performans sonuçları	63
4.14. Üç belirteç ile kurulmuş olan modellerin performans sonuçları	65
4.15. Dört belirteç ile kurulmuş olan modellerin performans sonuçları	67

4.16.	Beş belirteç ile kurulmuş olan tam modellerin performans sonuçları	68
4.17.	Elde edilen performans sonuçlarının modelde bulunan belirteç sayısı ile ve kendi aralarındaki ilişkileri	70
4.18.	Tiroid Verisi ile uygulanan Kuadratik Diskriminant Analizi Sonuçları	71
4.19.	Tek gen ekspresyon değerlerinden elde edilen modellerin performans sonuçları	72
4.20.	İki gen ile kurulan modellerin performans sonuçları	74
4.21.	Veri seti 1 ve Veri seti 2’de bulunan genlerin ekspresyon değerleri arasındaki ilişkiler	75
4.22.	Üç gen ile kurulan modellerin performans sonuçları	76
4.23.	Dört gen ile kurulan modellerin performans sonuçları	77
4.24.	Beş gen ile kurulan modellerin performans sonuçları	78
4.25.	Veri Seti 1 ve Veri Seti 2 ile uygulanan Kuadratik Diskriminant Analizi Sonuçları	79

1. GİRİŞ

Sağlık alanında, testlerin ya da kestirim modellerinden elde edilen risklerin doğru değerlendirilmesi, hasta ve sağlıklıları belirlemede önemlidir. Testlerin performanslarını değerlendirmede birçok ölçü bulunmakla birlikte, en sık kullanılan ROC (Receiver Operating Characteristic) eğrisinden elde edilen eğri altında kalan alandır. Bazı durumlarda sınıflamalar ayrı incelenmesi gereken ikiden fazla kategoriden oluşabilir. Örneğin Sinovit, skrolama yaklaşımı ile hastalık şiddetinin beş ayrı kategoride (sağlıklı, post travmatik artropati, osteoartrit, psöriatik artrit ve romatoid artrit) incelendiği bir hastalıktır (1). Tanısal performans, biyobelirteçler kullanılarak uygun yöntemlerle oluşturulmuş modellerde, hastanın hastalık sınıfının doğru saptanabilmesi ile belirlenir. ROC analizi yönteminden yola çıkarak çok kategorili sınıflama problemleri için bazı yeni yöntemler geliştirilmiştir.

Bu yöntemler, Mossman (2) tarafından üç kategorili sınıflamalarda ROC yüzeyi altında kalan hacim “Volume Under Surface (VUS)” ve üçten fazla kategoriye sahip durumlar için geliştirilmiş olan “Hypervolume Under The ROC Manifold (HUM)”dur. HUM, Nakas ve Yiannouyos tarafından 2004 yılında sıralanmış çoklu yanıtlar (sıralı yanıt değişkenleri) için çıkarılmıştır (2,3). Daha sonra Li ve Fine tarafından 2008 yılında sıralı olmayan çoklu yanıtlar (sırasız yanıt değişkenleri) için geliştirilmiştir (4). ROC tabanlı bu yöntemler bir çok araştırmacı tarafından kullanılmaya başlanmıştır.

“Doğru Sınıflandırma Olasılığı (Correct Classification Probability (CCP))”, incelenen olayların sınıflandırma başarılarına dayanmaktadır. Oldukça temel bir ölçü olan CCP, farklı araştırmacılar tarafından amaçlarına uygun şekillerde uyarlanarak farklı yaklaşımlarla kullanılmaktadır. 2002 yılında Zhou ve arkadaşları tarafından iki kategorili incelemeler için “duyarlılık” ve “seçicilik” değerlerinden elde edilen bir ölçü olarak da incelenmiş ve bir çok çalışmada kullanılmıştır (5). Li ve arkadaşları (1,4) CCP ölçüsünü; çok kategorili sınıflandırma modellerinde her sınıf için sınıf prevalansları ile ağırlıklandırma yaparak incelemişlerdir.

Ancak bazı araştırmacılar bu yöntemlerin, var olan bir modele yeni bir biyobelirteç eklendiğinde meydana gelen tanı performans artışını ölçümlemede

başarılı olmadığını ileri sürmüşlerdir. Pencina ve arkadaşları (6) 2008 yılında yeni bir biyobelirteç modele eklendiğinde oluşturulan yeni sınıflama sonucunun performansını gösteren iki yeni performans ölçüsü sunmuşlardır. Bunlar; NRI (Net Reclassification Improvement) “Net Yeniden Sınıflandırma İyileştirmesi” ve IDI (Integrated Discrimination Improvement) ”Birleştirilmiş Ayrımsama İyileştirmesi” dir (6).

Bu iki yeni ölçü tıp alanyazınına (literatürüne) yeni kazandırılmış olmasına karşın ROC tabanlı yöntemlere göre sunum ve uygulama kolaylığı açısından bir çok araştırmacı tarafından tercih edilen ölçüler haline gelmişlerdir. Ancak çoklu yanıt durumlarında yapılmış olan çalışma sayısı oldukça azdır ve son yıllarda üzerinde sıkça durulmaktadır.

Bu çalışmada, geliştirilmiş olan NRI ve IDI yöntemlerinin klasik performans ölçüleri olan HUM ve CCP yöntemleri ile olan ilişkilerinin incelenmesi amaçlanmıştır.

Bu ilişkiler incelenirken ele alınan hipotezler;

- Üç ve daha fazla sınıf sayısına sahip hastalık durumlarında tanısal modele yeni bir belirteç eklenmesi durumunda sınıflandırma yöntemleri olan HUM ve CCP yöntemlerine ek olarak NRI ve IDI yöntemleri ile, yeni eklenen belirteçin model performansında meydana getirdiği artışın incelenerek gösterilmesi sınıflama doğruluğu açısından önemlidir.
- Örneklem genişliğine göre model performansları değişkenlik göstermektedir. Modele eklenen belirteçler ile model performanslarında meydana gelen artışların istatistiksel olarak anlamlılığını incelemek tanısal modele karar vermek açısından önemlidir.
- Bağımlı değişkenin türüne göre model performansları değişkenlik göstermektedir. Üç ve daha fazla sınıf sayısına sahip hastalık durumlarında, bağımlı değişkenin sıralı ve sırasız olması durumlarına göre model performanslarında meydana gelen artışların istatistiksel olarak anlamlılığını incelemek tanısal modele karar vermek açısından önemlidir.
- Bağımlı değişken ve tanısal model için kullanılan belirteçler arasındaki ilişkiler model performansını etkilemektedir.

- Tanısal model kurulurken, modele eklenen belirteçlerin eklenme sırasına göre model performansına etkisinin bilinmesi model başarısı açısından önemlidir. Bu hipotezlerin incelenebilmesi için;
- Klasik yöntemlerin (HUM ve CCP) ve yeni yöntemlerin (NRI, IDI) tanıtılması,
- Bağımlı değişkenin sıralı ve sırasız olduğu durumlarda kurulan tanısal model performanslarının bu yöntemlerle incelenmesi,
- Tanısal modele eklenen belirteçlerin, bağımlı değişken ile olan ilişkilerinin ve kendi aralarında olan ilişkilerinin incelenmesi ve bu ilişkilerin model performansına etkisinin incelenmesi,
- Tanısal modele eklenen belirteçlerin, modele eklenme sırasının model performansında ortaya çıkardığı sonuçların incelenmesi,
- İncelenen yöntemlerden elde edilen sonuçlar arasındaki ilişkilerin incelenmesi amaçlanmıştır.

2. GENEL BİLGİLER

2.1. Üç ve Üzeri Kategori Olması Durumunda ROC Yüzeyi Yöntemi (VUS ve HUM)

Bir bireyin gerçekten hasta olup olmadığını belirlemede altın standart testler (gold standard tests) ya da belirteçler sıklıkla kullanılmaktadır. Fakat bu testler pahalı, hasta için riskli ya da uzun zamanda sonuç alınan testler olabilmektedir. Bazı hastalıkların tanısında ise altın standart test bulunmayabilir ya da bulunsa da yukarıda söz edilen sakıncalardan ötürü uygulanmayabilir. Bu nedenlerden dolayı alternatif olarak ucuz, hastaya zarar vermeyen ve kısa sürede sonuç alınabilen ancak sonuçları kesin olmayan testler sıklıkla tercih edilmektedir (7).

ROC analizi, tanı testlerinin uygulanabilirlik, doğruluk ve ayırt edicilik güçlerinin değerlendirilmesinde en yaygın kullanıma sahip olan grafiksel bir yöntemdir. ROC analizi ile bir testin ayırt ediciliğinin, yani performansının değerlendirilmesinde, ROC eğrisi altında kalan alan (AUC) incelenmektedir. AUC, biri hasta diğeri sağlıklı olduğu bilinen iki kişinin test (belirteç) sonuçlarının doğru sıralanması olasılığını vermektedir. Doğru sıralanma; hasta bireylerin test sonuçlarının sağlıklı bireylere göre daha yüksek olması durumunda, bir hasta bireyin test sonucunun sağlıklı bireyin test sonucundan daha büyük olması durumudur. AUC değeri 0,5'e eşit olan testler ayırtıcılığı olmayan testlerdir (8).

İncelenen bir testte, belirlenen farklı kesim noktaları için duyarlılık ve seçicilik değerleri farklılık gösterecektir. ROC eğrisi, iki sınıflı bir sınıflama sistemi içinde değişen kesim noktalarına göre elde edilen duyarlılığa (doğru pozitif oran) karşı 1-Seçiciliğin (1-Doğru negatif oran) çizildiği bir grafikdir. Bu grafik, görsel açıdan test sonuçlarının kolaylıkla yorumlanabilmesini sağlamakla birlikte, sonuçların performansı ROC eğrisi altında kalan alanın (AUC) hesaplanması ile de değerlendirilebilir. Elde edilen bu alan testin performansını gösteren tümel bir ölçü olarak tanımlanmaktadır (2, 8).

AUC ölçüsünün, üç sınıflı hastalık durumları için uygulanamaması üzerine, ilk olarak 1996 yılında Scurfield (9), ikiden fazla sınıf için uygun ROC ölçümlerini öneren

bir çalışma gerçekleştirmiştir. Bu çalışma sonrasında, ROC eğrilerinin üç ve üzeri kategorili sınıflandırmalar için uyarlanması üzerine araştırmalar çoğalmıştır. Bunun sonucu olarak, üç kategori sınıflandırmalar için ROC yüzeyleri ve çok kategorili sınıflandırmalar için ise ROC manifoldları önerilmiş, iki kategorili sınıflandırmada kullanılan AUC yerine ise sırasıyla ROC Yüzeyinin Altındaki Hacim (VUS) ve ROC Manifoldunun Altındaki Hipervolum (HUM) değerleri önerilmiştir. Mossman (3) ise, üç kategorili sınıflandırmalar için “Üç Yönlü (boyutlu) ROC Analizi” kavramını tıp ile ilgili karar süreçlerinde ilk kez kullanmıştır.

Üç yönlü ROC analizinde hastalık durumu üç sınıfa sahip olduğu için ROC grafiği üç boyutlu olarak ele alınmaktadır. Analiz sonucunda testin performans göstergesi olarak ROC yüzeyi (surface) çizilir ve ROC yüzeyi altında kalan hacim (Volume Under the ROC Surface – VUS) hesaplanır (10). VUS, testin ayırt ediciliği hakkında tüm olası eşik değerleri için tümel bir performans ölçüsü olarak kullanılır ve sıfır ile bir değerleri arasında değişir. Bu değer, üç sınıflı tanı problemlerinde her bir gruptan rastgele seçilen bireylerin test sonuçlarının doğru olarak sıralanması olasılığını verir. Üç kategorili sınıflandırma durumunda doğru sıralanma; hasta bireylerin test sonuçlarının sağlıklı ve şüpheli bireylere göre daha yüksek, sağlıklı bireylerin düşük ve şüpheli grubun test sonuçlarının ortada yer aldığı durum olarak ifade edilmektedir.

Y1, Y2, Y3 sırası ile rasgele seçilen sağlıklı, şüpheli ve hasta üç kişinin test sonuçlarını göstermek üzere, üç yönlü ROC analizinde yüzey altında kalan hacim $P(Y3 > Y2 > Y1)$ ile gösterilir. Testin yüksek değerlerinin hastalığı gösterdiği durum için böyle tanımlanırken düşük değerler hastalığı gösterdiğinde tam tersi bir tanımlama $P(Y1 > Y2 > Y3)$ geçerlidir. Yüzey altında kalan hacim bire yaklaştıkça testin ayırt edicilik gücü artar (2,3,5).

Üç kategorili sınıflandırmada, kategorilerin sıralı olması durumu için Nakas ve Yiannoutsos (3) VUS kestiriminde U-istatistiği teorisini kullanan bir yaklaşım geliştirmiştir. Bu çalışmada, iki kategorili ROC analizinin genelleştirilmesi olan üç yönlü ROC analizi, k kategorili duruma genellenerek, HUM yöntemi önerilmiş ve HUM değeri $1/(k!)$ değerine eşit olan testler ayıricılığı olmayan testler olarak tanımlanmıştır (3).

Li ve Fine (4), Nakas ve Yiannoutsos (3)'un yaklaşımlarından yola çıkarak, sıralı olmayan çok kategorili sınıflama problemleri için VUS ve HUM kestirimlerinde, sınıflara atanma olasılıklarının kestirilmesine dayalı bir yaklaşım önererek, araştırmalarında HUM ölçüsünü mikrodizin çalışmaları için bir model seçim kriteri olarak uygulamıştır (4). Xiong ve arkadaşları da (11) üç sınıflı ve sıralı bağımlı değişken olması durumunda HUM ölçüsünü karşılaştırabilmek için bir test yöntemi önermiştir.

Mossman'ın (3) 1999 yılında önermiş olduğu parametrik ROC yüzeyi yaklaşımı sonrasında, Li ve Zhou (12), 2009 yılında yapmış oldukları çalışmada asimptotik ROC yüzeyleri yaklaşımıyla parametrik olmayan ve yarı parametrik ROC yüzey tahmin yöntemleri önermiştir. Dreiseiltl ve arkadaşları (13), HUM ölçüsünü AUC'ye benzer bir yaklaşımla yorumlamışlar ve yüksek bir HUM değerinin yüksek bir sınıflama doğruluğunu gösterdiğini belirtmişlerdir.

2.2. Doğru Sınıflandırma Olasılığı (Correct Classification Probability – CCP)

Kategorilere ait doğruluk ölçüleri, sınıflandırma doğruluğunun genel bir değerlendirmesini verir. Ancak, farklı belirteçlerin ayırıcılık başarıları sıralanırken, incelenen belirteç ya da belirteç kombinasyonlarının toplam başarısının ölçülebilmesi için tek bir ölçüye gereksinim vardır (14).

Doğru Sınıflandırma Olasılığı, incelenen testlerle elde edilen sonuçların sınıflandırma başarılarına dayanmaktadır. Oldukça temel bir ölçü olan CCP, farklı araştırmacılar tarafından amaçlarına uygun şekillerde uyarlanarak farklı yaklaşımlarla kullanılmıştır. 2002 yılında Zhou ve arkadaşları tarafından iki kategorili incelemeler için “duyarlılık” ve “seçicilik” değerlerinden elde edilen bir ölçü olarak da incelenmiş ve bir çok çalışmada kullanılmıştır (5). Li ve arkadaşları (1,4) CCP ölçüsünü; çok kategorili sınıflandırma modellerinde her sınıf için sınıf prevalansları ile ağırlıklandırma yaparak incelemişlerdir. Bu incelemeye göre, incelenen her kişinin ait olma olasılığı en yüksek olan sınıfa atandığını ve doğru sınıfa atanan kişilerin oranının ise CCP değerini verdiğini belirtmişlerdir. Li ve arkadaşlarının (1) yapmış olduğu

çalışmada, çoklu yanıt durumları için, iki model arasında elde edilen CCP farklılığı incelenmiş ve bu farka yeniden sınıflandırma iyileştirmesi ismini vermişlerdir.

2.3 Net Yeniden Sınıflandırma İyileştirmesi (Net Reclassification Improvement – NRI)

NRI, elimizdeki bir risk kestirim modelinden (eski model) elde edilen sınıflandırma sonucu ile, bu modele yeni bir belirteç eklenerek elde edilen sınıflandırma sonucunun (yeni model), doğru ve hatalı sınıflanma durumlarını ne yönde değiştirdiği ile ilgilenir. Modele eklenen yeni belirteç ile oluşan sınıf değişimlerini göz önüne alarak bu belirtecin model performansında meydana getirdiği değişimin eski modelde edilen sınıflandırma sonuçlarına göre modele ne oranda katkı sağladığını inceler. Olayların ayrı ayrı sınıflandığı yeniden sınıflandırma tablolarına odaklanmaktadır ve kategoriler arasındaki doğru geçişleri hesaplamaktadır (1,6).

2.4. Birleştirilmiş Ayrımsama İyileştirmesi (Integrated Discrimination Improvement - IDI)

NRI'da olduğu gibi IDI da, eski modele yeni bir belirteç eklendiğinde elde edilen risk kestirim değerlerinin olayın görüldüğü ve görülmediği kişileri birbirinden ne kadar iyi ayırdığını ölçmektedir. Burada eski model ve yeni model arasındaki farklılık iki modelin ayrımsama eğimleri ile belirlenmektedir. Ayrımsama eğimlerinin hesaplanması ise model sonuçlarından elde edilen “duyarlılık” ve “1 – seçicilik” değerlerinin integralleri ile elde edilmektedir (1, 15).

2.5. NRI ve IDI’nın Gelişimi

Pencina ve arkadaşlarının (6), bu iki yeni ölçüyü tanıtmış olduğu 2008 yılında yapılmış olan çalışmalarında, koroner kalp hastalığı için oluşturulmuş olan bir risk kestirim modeline HDL kolesterol değişkenini ekleyerek elde ettikleri sonuçları incelemişlerdir. Elde ettikleri sonuçlarda, AUC değerindeki ufak sayılabilecek bir artışın bile NRI ile belirtilen yeniden sınıflama sonuçlarında önemli bir artışa yol açabileceği ve IDI değerinde daha küçük bir yükseliş olabileceğini söylemişlerdir. HDL kolesterol değişkeninin eklenmesi ile oluşturulan yeni model sonucunda elde edilen NRI ve IDI değerlerinin istatistiksel anlamlılığını z test istatistiğini kullanarak

incelemişlerdir. NRI ve IDI performans ölçüm yöntemi olarak önerilmektedir, ancak AUC'deki artışın da her durumda öncelikli kriter olması gerektiği belirtilmektedir (6).

Steyerberg ve arkadaşlarının (16), 2011 yılında yapmış oldukları çalışmada, risk sınıfları için belirlenen farklı kesim noktalarının NRI performanslarını değiştirdiği sonucuna ulaşmışlardır. Pencina ve arkadaşlarının (17), 2011 yılında yapmış olduğu çalışmada, NRI ölçümü için önceden saptanmış ve kabul edilmiş kesim noktalarının olmasının NRI performansı açısından önemli olduğunu belirtmişlerdir. Muhlenbruch ve arkadaşlarının (18), 2013 yılında yapmış oldukları çalışmada, Tip 2 diyabet hastalığı verileri ile farklı biyobelirteçler kullanılarak üç model oluşturulmuş ve modellerin performans sonuçları incelenmiştir. NRI sonuçlarının belirlenmiş olan kesim noktalarına bağlı olarak değiştiği gösterilmiştir.

Risk sınıflarının kesim noktalarına göre NRI değerlerinin çok fazla değişmesinden dolayı NRI yöntemini eleştiren çalışmalar üzerine 2009 yılında Pencina ve arkadaşları bu yöntem için bazı özel çıkarımlar yapmıştır (19). Yapmış oldukları bu çalışmada, “ağırlıklandırılmış NRI” ve “sürekli NRI” yöntemlerini tanıtmışlardır.

Sürekli NRI, belirlenmiş bir kesim noktası bulunmayan uygulamalarda kesim noktası seçme sorununun üstesinden gelmek için önerilmiştir. Bu yöntem, incelenen olayın görüldüğü ve görülmediği tüm durumlar için herhangi bir risk sınıfına gerek duymadan öngörülen riskteki tüm değişiklikleri dikkate alır (20).

Kerr ve arkadaşlarının (21), 2014 yılında yapmış olduğu derleme çalışmasında, yaklaşık olarak 500 civarında çalışmada NRI yönteminin kullanıldığını ve bu çalışmaların özellikle kardiyovasküler alanda yapılmış olduğunu belirtmişlerdir. Ayrıca NRI kullanılan çalışmalarda yapılmış olan genel hataların; NRI'nın gösterimi, yorumlanması ve risk sınıfı belirlenmesi ile ilgili olduğunu göstermişlerdir (21).

Li ve arkadaşları (1), 2013 yılında yapmış oldukları çalışmada, NRI ve IDI ölçümlerini hem benzetim yaparak hem de iki ayrı gerçek veri seti ile incelemişlerdir. Benzetim çalışmasında, NRI ve IDI ölçümlerinin, yanıt değişkeninin kategori sayısından ve tanısız modelde bulunan belirteç sayısından ne şekilde etkilendiğini gösterebilmek için çok değişkenli normal dağılıma uygun veriler üretmişlerdir. Ayrıca

örneklem büyüklüğünün NRI ve IDI performanslarına olan etkisini inceleyebilmek için, $n=150$ ve $n=300$ olacak şekilde incelemeler yapmışlardır. Bu verilere, üç kategorili yanıt değişkeni için iki bağımsız değişken ile multinominal lojistik regresyon; üç kategorili yanıt değişkeni için beş bağımsız değişken ile multinominal lojistik regresyon; beş kategorili yanıt değişkeni için üç bağımsız değişken ile multinominal lojistik regresyon yöntemlerini 1000 tekrar olacak şekilde uygulamışlardır.

Elde edilen tüm benzetim sonuçlarında "Kapsama Oranı (Coverage Rate)" kullanılarak NRI ve IDI sonuçlarının tutarlılığını incelemişlerdir. Bu değer $\%95$ ve üzerinde olması başarılı bir modelleme yapıldığının göstergesi olarak kabul edilmiştir. Li ve arkadaşları tarafından yapılmış olan bu benzetim çalışmasında, NRI ve IDI sonuçlarının kapsama oranlarının $\%95$ civarında olduğu ve örneklem genişliği büyüdükçe kapsama oranlarının yükseldiği gösterilmiştir. Benzetim çalışması sonuçlarını izleyerek gerçek veri setleri üzerinde iki uygulama yapmışlardır.

İlk olarak; sinovit hastalığının şiddetini belirten skorlama sonucunu beş kategoride incelemişlerdir. Bu kategoriler: 1-Sağlıklı kontroller 2-Post travmatik artropati 3-Osteoartrit 4-Psöriatik artrit 5- Romatoid artrit şeklinde sınıflanmıştır. Sinovit hastalığının tanısında kullanılan üç farklı belirteç ile öncelikle tek belirteçli lojistik regresyon modelleri kurmuş ve daha sonra bu belirteçleri ikili olarak modellemişlerdir. En yüksek performans sonucunu veren ikili modele üçüncü belirteç de eklenerek tekrar incelemişlerdir. Üç değişkenli modelden elde edilen performans sonucuna göre üçüncü belirteçin modelde bulunmasının anlamlı bir etki yaratmadığı görülmüştür. Lojistik regresyon modeli kullanılmış olduğu için NRI ve IDI değerlerinin istatistiksel anlamlılığını Wald testi kullanarak elde etmişlerdir.

İncelemiş oldukları diğer bir örnekte ise, çok kategorili bir sınıflandırma modeli incelenmiş ve akut lösemi hastalığı sınıflaması kullanılmıştır. Bu akut lösemi tipleri, T hücrelerinden kaynaklanan akut lenfoblastik lösemi (ALL T), B hücrelerinden kaynaklanan akut lenfoblastik lösemi (ALL B) ve akut miyeloid lösemi (AML)'dir. Tek bir gen ile doğru sınıflama olasılığı (CCP) 0,84 elde edilmiştir. Doğruluk artışını maksimize eden diğer geni bulabilmek için NRI ve IDI yöntemlerini kullanarak yeni bir sınıflandırma sonucu elde edilmiş ve bu iki gen ile yapılan

sınıflandırmanın sonucunda ise tek gen ile elde edilen sınıflandırmada daha önce doğru sınıflandırılmayan gözlemlerde %20 civarında daha doğru sınıflara atandığı hesaplanmıştır. Elde edilen HUM sonuçları, klasik yöntem olarak kullanmış oldukları CCP yönteminden elde edilen sonuçlarla da karşılaştırma yapılarak (bir gen ile %81 ve iki gen ile %100) incelenmiş ve oldukça yakın değerler elde edildiği gösterilmiştir. Hem HUM hem de CCP incelemelerinde, modele eklenmiş olan ikinci genin belirlenmesinde NRI ve IDI yöntemleri kullanılmıştır (1).

Ayrıca bu çalışmada, R yazılımında “MCCA” isimli bir paket hazırlamışlardır. Paket içeriğinde, 2013 yılında yapmış oldukları çalışmada kullandıkları formülasyonları tanımlamışlar ve “HUM”, “CCP”, “NRI” ve “IDI” gibi ölçüleri kullanıma sunmuşlardır (1, 22).

3. GEREÇ VE YÖNTEM

Bu çalışmada, çok kategorili sınıflandırma modellerinde yaygın olarak kullanılan performans ölçüleri olan HUM ve CCP'nin, model performansına etki eden belirteçlerin etkisini ölçen NRI ve IDI ile ilişkilerinin incelenmesi amaçlanmıştır. Ayrıca, çok kategorili sınıflandırma modellerinde bağımlı değişkenin farklı tipte olduğu durumlar göz önüne alınarak (sıralı ve sırasız), belirteçlerin modele eklenme sırasının model performansına etkisi araştırılmıştır. Bu amaçlar doğrultusunda, iki kategorili sınıflandırma modellerinde oldukça sık kullanılan ve çok kategorili sınıflandırma modelleri için henüz gelişmekte olan NRI ve IDI yöntemlerinin çok kategorili sınıflandırma modellerinde de, modele eklenen belirteç/lerin belirlenebilmesi için kullanılan yöntemler arasında sayılabileceğinin önerisinin yapılması hedeflenmiştir.

Bu bağlamda, tezin köken aldığı hipotezleri sınavabilmek ve daha genellenebilir bir hale getirmek amacıyla hem gerçek veri setleri ile incelemeler yapılmış hem de bir benzetim çalışması kurgulanmıştır. Bağımlı değişkeni sıralı çok kategorili olan ve incelenen belirteçlerle farklı ilişki yapılarında veri setleri üretilmiştir. Benzetim çalışmasında tasarlanan farklı ilişki yapıları, hem belirteçlerin kendi arasında hem de belirteçler ve bağımlı değişken arasında olmak üzere; pozitif, negatif, düşük, orta ve kuvvetli seviyelerde tasarlanmıştır. Ayrıca modelde performans artışına etki eden önemli belirteçlerin belirlenmesinde ilişki yapılarının ve örneklem genişliklerinin ne şekilde sonuç verdiği incelenmiştir.

Ek olarak, bağımlı değişkenin sıralı olmayan çok kategorili olduğu (sırasız) iki ayrı açık erişimli veri seti ile farklı model kombinasyonları farklı örneklem genişliklerinde incelenerek belirteçlerin sırasının, model performans sonuçlarına nasıl bir katkı yaptığı araştırılmıştır. Bu veri setleri sırasıyla, UCI (University of California, Irvine) makine öğrenmesi veri tabanından (23) alınmış tiroid hastalığına ait veri seti ve bir çok sınıflandırma çalışmasında kullanılmış olan Golub ve arkadaşlarının (24) kullandıkları akut lösemi hastalığına ait mikrodizin veri setidir.

3.1. HUM Değerlerinin Kestirimi

Dreiseitl ve arkadaşları (13), Mossman'ın geliştirmiş olduğu üç yönlü ROC analizi yöntemini U istatistiği kullanarak düzenlemişlerdir. Bu düzenleme ile VUS kestiricisi için bir varyans kestirimi geliştirmiş ve bu kestirim sayesinde bir hipotez test metodu sunmuşlardır. Nakas ve Yiannoutsos (3) ise, sıralı çoklu yanıtlar için bir HUM kestirimi önermişlerdir. Daha sonra Li ve Fine (4), sıralı olmayan çoklu yanıtlar için bir HUM kestirimi önermiştir. Bunun dışında, literatürde bir çok araştırmacının önermiş olduğu farklı HUM kestirimleri bulunmaktadır. Yapmış olduğumuz çalışmada Li ve Fine tarafından sıralı olmayan çoklu yanıtlar için geliştirilmiş olan HUM kestirimi üzerinde durulacaktır (4).

Li ve Fine (4), Dreiseitl ve arkadaşlarının (13), önerdiği HUM kestirimini $M \geq 3$ için genelleştirmişlerdir. $i = 1, \dots, M$ ve p_{ij} $j = 1, \dots, n_i$ adet sınıf sayılı n_i gözleme ait olasılık vektörü olmak üzere, HUM kestiricisi Eşitlik 3.1. ile gösterilmiştir.

$$\frac{1}{\prod_{h=1}^M n_h} \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \dots \sum_{k_M=1}^{n_M} CR(p_{1_{k_1}}, p_{2_{k_2}}, \dots, p_{M_{k_M}}) \quad (3.1)$$

Dreiseitl ve arkadaşları U istatistiği kullanarak bu kestiricinin asimptotik olarak normal dağılıma uyduğunu ve varyansın hesaplanabilir olduğunu göstermişlerdir. Sınıf sayısının 3'den büyük olma durumunda ($M > 3$) varyansın kestirimi oldukça karmaşık hale gelmekte, p_{ij} değerlerinin bilinmemesi durumunda ise kestirici hesaplanamamaktadır.

Li ve Fine (4), p_{ij} değerlerinin bilinmemesi durumunu incelemişlerdir. p_{ij} değerlerini tanısal testler kullanılarak kestirmiş ve bu değerleri Eşitlik 3.1'de gösterilmiş olan HUM kestiricisinde kullanmışlardır. Li ve Fine (4)'ün yaklaşımı şu şekildedir:

Her gözlem, $i = 1, \dots, M$, ve $j = 1, \dots, n_i$ olmak üzere, q sayıda tanısal test sonucuna sahip olsun. Bu durumda olasılık vektörü Eşitlik 3.2. ile gösterilebilir.

$$T_{ij} = (T_{ij1}, T_{ij2}, \dots, T_{ijq})^T \quad (3.2.)$$

Sınıf olasılıkları olan p_{ij} ' ler; $i = 1, \dots, M$, $j = 1, \dots, n_i$, ve $c = 1, \dots, M$ olmak üzere, multinomial lojistik regresyon analizi kullanılarak Eşitlik 3.3. ile kestirilebilir

$$P_{ij,c}(\beta) = \Pr(c \text{ sınıfında bulunan } ij \text{ kişisi} | T_{ij}) = \frac{I(c>1) \exp(\beta_{c-1}^T T_{ij}) + I(c=1)}{1 + \sum_{k=1}^{M-1} \exp(\beta_k^T T_{ij})} \quad (3.3.)$$

Bu hesaplamadaki $\beta = (\beta_1^T, \beta_2^T, \dots, \beta_{M-1}^T)^T$ parametresi Eşitlik 3.4. ile gösterilen Log-olabilirlik fonksiyonunu en büyük yaparak kestirilebilir.

$$\sum_{i=1}^M \sum_{j=1}^{n_i} \log \{P_{ij,i}(\beta)\} \quad (3.4.)$$

Eşitlik 3.4.'te bulunan, β yerine $\hat{\beta} = (\hat{\beta}_1^T, \dots, \hat{\beta}_{M-1}^T)^T$ kullanılırsa, sınıf olasılıklarının kestirilmiş değerleri olan Eşitlik 3.5. elde edilir.

$$\hat{p}_{ij} = (\hat{p}_{ij,1}, \hat{p}_{ij,2}, \dots, \hat{p}_{ij,M}), \quad i = 1, \dots, M, \quad j = 1, \dots, n_i \quad (3.5.)$$

Bu süreçte lojistik regresyon dışında farklı istatistiksel yöntemler de kullanılabilir. Elde edilecek olan M boyutlu setler Eşitlik 3.6. ile gösterilen HUM kestiricisinde bulunan p_{ij} sınıf olasılık vektör setleri yerine yerleştirildiğinde,

$$\widehat{HUM} = \frac{1}{\prod_{h=1}^M n_h} \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \dots \sum_{k_M=1}^{n_M} CR(\hat{p}_{1k_1}, \hat{p}_{2k_2}, \dots, \hat{p}_{Mk_M}) \quad (3.6.)$$

eşitliği elde edilecektir.

3.2. CCP Değerlerinin Kestirimi

M_1 modelinin, Ω sayıdaki belirteç içerisinde Ω_1 sayıdaki belirteç ile kurulduğunu düşünelim. Böyle bir M_1 modeli; $\sum_{m=1}^M p_m(M_1) = 1$ olduğu durumda her gözlem için Eşitlik 3.7. gibi olan bir sınıf olasılık vektörü elde edilecektir.

$$p(M_1) = (p_1(M_1), \dots, p_M(M_1)) \quad (3.7.)$$

Karar vericiler, bir kişiyi, sınıf olasılık vektöründe bulunan en yüksek olasılıklı bileşene göre M sayıda sınıftan bir sınıfa atayacaklardır. Bu durumda, Ω_1 sayıda belirteç ile kurulmuş olan M_1 modelinden elde edilen doğru sınıflandırma olasılığı Eşitlik 3.8. gibi olacaktır.

$$CCP = \sum_{m=1}^M p_m CCP_m \quad (3.8.)$$

CCP_m , m. kategori için CCP değeri olduğunda M_1 modeli için CCP Eşitlik 3.9. gibi elde edilir.

$$CCP_m(M_1) = P\{p_m(M_1) = \max p(M_1 | Y = m)\}, \quad m = 1, \dots, M. \quad (3.9.)$$

Daha fazla belirteç ekleyerek bir model daha elde ettiğimizi düşünelim. Yeni elde edilen M_2 modelinde; $\sum_{m=1}^M p_m(M_2) = 1$ olduğu durumda her gözlem için olasılık vektörü,

$$p(M_2) = (p_1(M_2), \dots, p_M(M_2)) \quad (3.10.)$$

şeklinde olacaktır. Burada M_2 modeli için Eşitlik 3.11. elde edilecektir.

$$CCP_m(M_2) = P\{p_m(M_2) = \max p(M_2 | Y = m)\}, \quad m = 1, \dots, M. \quad (3.11.)$$

Eşitlik 3.9. ve Eşitlik 3.11. ile gösterilen modellerden elde edilen doğruluk oranlarında meydana gelen toplam doğruluk artışı Eşitlik 3.12. ile gösterilebilir. Bu formülde ω_m değeri m. kategori için pozitif ağırlık değeridir.

$$T = \sum_{m=1}^M \omega_m \{CCP_m(M_2) - CCP_m(M_1)\} \quad (3.12.)$$

$M=2$ olduğu durumlarda, altın standart testlere göre yapılan incelemelerde, $CCP_m(M_1)$ ve $CCP_m(M_2)$ değerleri, incelenen testlerin duyarlılık ve seçicilik değerleri olarak ifade edilir. Burada T ölçüsü, duyarlılık ve seçicilik değerlerinin ağırlıklandırılmış toplamalarında meydana gelen artışı göstermektedir (1, 14).

3.3. NRI Değerlerinin Kestirimi

NRI, amaca göre kategorik ve sürekli olarak iki ayrı şekilde kullanılmaktadır. Risk sınıfları için kesinleştirilmiş sınıflar mevcut olduğunda kategorik NRI kullanımının doğru sonuçlar verdiği birçok çalışmada gösterilmiştir. Belirlenmiş risk sınıfları mevcut olmadığında sürekli NRI kullanımı ön plana çıkmaktadır.

3.3.1. Kategorik NRI

Risk kestirim modelleri, hasta ve sağlıklı olan kişilerin ilgilenilen hastalık açısından sahip oldukları risk düzeylerinin sınıflandırılmasında oldukça yaygın olarak kullanılmaktadır. Kullanılan risk kestirim modeline yeni bir belirteç eklendiğinde elde edilen yeni risk kestirim modelinin sınıflama sonuçlarının, eski modelden elde edilen sonuçlara göre daha yüksek doğruluk oranında kestirim yapabildiğinin belirlenebilmesi risk kestirim modelleri oluşturabilmek için oldukça önemlidir. Burada önem kazanan nokta ise, yeni bir belirteç eklenerek kurulan modelin, eski modelde bulunan bir kişiyi hangi sınıfa geçirdiğidir. Örneğin; hasta olduğu bilinen bir kişinin yeni belirteçle kurulan model sonucunda, daha önce bulunduğu sınıftan daha yüksek riskli bir sınıfa geçmesi (yukarı yönde hareket) ya da sağlıklı olduğu bilinen bir kişinin yeni belirteçle kurulan model sonucunda, daha önce bulunduğu sınıftan daha az riskli bir sınıfa geçmesi (aşağı yönde hareket) model performansı açısından bir kazançtır. Tam tersini düşünecek olursak; hasta olduğu bilinen bir kişinin yeni belirteçle kurulan model sonucunda daha az riskli bir sınıfa geçmesi (aşağı yönde hareket) ya da sağlıklı olduğu bilinen bir kişinin yeni belirteçle kurulan model sonucunda daha yüksek riskli bir sınıfa geçmesi (yukarı yönde hareket) model performansı açısından istenmeyen bir durumdur.

Kategorik NRI incelemesinde, hasta kişiler yani olayın (event) görüldüğü kişiler ve sağlıklı kişiler (olayın görülmediği kişiler) ayrı tablolar ile incelenmektedir. Yeni modelden elde edilen sınıflama sonuçları eski modeldeki sınıflama sonuçlarına göre hesaplanmaktadır. Yeni model ile elde edilen sınıflama sonucunda eski modele göre kaç kişinin doğru sınıfa atandığı belirlenmektedir. Olayın görüldüğü ve görülmediği kişilerden elde edilen sonuçlar birleştirilerek tek bir sonuç elde edilmektedir. NRI'nın kavramını geliştirerek, yukarı yönde oluşan hareketler için “1” tanımlayalım. Aşağı yönde oluşan hareketler için de “-1” tanımlayalım. Kategoriler arasında hareket olmaması durumunu ise “0” olarak tanımlarsak; NRI, Eşitlik 3.13. ile gösterilebilir (1,6);

$$\frac{\sum_i \text{olgu } v(i)}{\# \text{ olgular}} - \frac{\sum_j \text{sağlıklı } v(j)}{\# \text{ sağlıklılar}} \quad (3.13.)$$

Burada $v(i)$ daha riskli bir sınıfa geçişin (üst sınıfa geçiş, yukarı yönde hareket) göstergesidir. Yeni kurulan model ile daha önce bulunduğu risk sınıfından daha riskli bir sınıfa geçen bir “olgu” için $v(i) = 1$ olacaktır. Aynı şekilde “olgu”nun aşağı yönde bir hareketi için de $v(i) = -1$ olacaktır. Olasılığın negatif yönde olması durumu model performansını düşürmesi açısından tanımlanmıştır. Olay görülen bir kişinin yeni kurulan model ile daha az riskli bir sınıfa geçmesi istenmeyen bir durumdur. Yani yeni modele eklenen belirtecin eski modele göre daha iyi bir sınıflama kestirimi sunmadığını göstermektedir. Bu durumda, her kişi için 0, 1 ya da -1 yerine, öngörülen olasılıklarıdaki farkı atamak daha mantıklı olacaktır. $\hat{p}_{yeni}$ ile olguların model bazlı yeni öngörülen olasılıkları ve eski model bazlı olasılıklarını da $\hat{p}_{eski}$ ile gösterecek olursak Eşitlik 3.14. elde edilecektir.

$$\frac{\sum_i olgu(\hat{p}_{yeni}(i) - \hat{p}_{eski}(i))}{\# olgular} - \frac{\sum_j sağlıklı(\hat{p}_{yeni}(j) - \hat{p}_{eski}(j))}{\# sağlıklılar} \quad (3.14.)$$

Belirli bir olgu için, eski model ile yeni belirtecin eklenmesi ile elde edilen yeni modelden hesaplanan olayın gerçekleşmesi olasılıklarını kullanarak oluşturulan sınıflandırma durumunu düşünelim. Yeni belirteç ile kurulan ikinci model ile olgunun görüldüğü grupta yukarı doğru meydana gelen bir hareket ve sağlıklı grupta aşağı doğru meydana gelen bir hareket için D tanımlamasını yapalım. NRI kestirimi için Eşitlik 3.15., Eşitlik 3.16., Eşitlik 3.17. ve Eşitlik 3.18. ile gösterilen dört olasılık tahmin edicisi tanımlarız:

$$\hat{P}(yukarı | D = 1) = \hat{p}_{yukarı,olgu} = \frac{\# yukarı hareket olan olgular}{\# olgular} \quad (3.15.)$$

$$\hat{P}(aşağı | D = 1) = \hat{p}_{aşağı,olgu} = \frac{\# aşağı hareket olan olgular}{\# olgular} \quad (3.16.)$$

$$\hat{P}(yukarı | D = 0) = \hat{p}_{yukarı,sağlıklı} = \frac{\# yukarı hareket sağlıklılar}{\# sağlıklılar} \quad (3.17.)$$

$$\hat{P}(aşağı | D = 0) = \hat{p}_{aşağı,sağlıklı} = \frac{\# aşağı hareket olan sağlıklılar}{\# sağlıklılar} \quad (3.18.)$$

Bu durumda NRI, Eşitlik 3.19 ile gösterilebilir ve NRI kestirimi de Eşitlik 3.20. gibi hesaplanacaktır.

$$NRI = [P(yukarı | D = 1) - P(aşağı | D = 1)] - [P(yukarı | D = 0) - P(aşağı | D = 0)] \quad (3.19.)$$

$$\widehat{NRI} = (\hat{p}_{yukarı,olgu} - \hat{p}_{aşağı,olgu}) - (\hat{p}_{yukarı,sağlıklı} - \hat{p}_{aşağı,sağlıklı}) \quad (3.20.)$$

Olgu ve sağlıklı sınıflandırmalarındaki gelişimin değerlendirilmesi NRI'nın bileşenleri kullanılarak test edilebilir. NRI'nın test edilebilmesi için gerekli bileşenlerin testleri, Eşitlik 3.21. ile olgular için, Eşitlik 3.22. ile de sağlıklı kişiler için gösterilmiştir.

$$Z_{olgu} = \frac{\hat{p}_{yukarı,olgu} - \hat{p}_{aşağı,olgu}}{\sqrt{\frac{\hat{p}_{yukarı,olgu} + \hat{p}_{aşağı,olgu}}{\# olgular}}} \quad (3.21.)$$

$$Z_{sağlıklı} = \frac{\hat{p}_{aşağı,sağlıklı} - \hat{p}_{yukarı,sağlıklı}}{\sqrt{\frac{\hat{p}_{aşağı,sağlıklı} + \hat{p}_{yukarı,sağlıklı}}{\# sağlıklı}}} \quad (3.22.)$$

İki model arasında meydana gelen performans artışını inceleyebilmek için, sıfır hipotezinin NRI değerinin 0 olması durumu yani iki model arasında herhangi bir değişim olmadığı ($H_0: NRI = 0$) durum olarak düşünecek olursak; McNemar Anlamlılık testi formülü ile sıfır hipotezini test edebilmek için Eşitlik 3.23. kullanılabilir.

$$Z = \frac{\widehat{NRI}}{\sqrt{\frac{\hat{p}_{yukarı,olgu} + \hat{p}_{aşağı,olgu}}{\# olgular} + \frac{\hat{p}_{yukarı,sağlıklı} + \hat{p}_{aşağı,sağlıklı}}{\# sağlıklı}}} \quad (3.23.)$$

3.3.2. Sürekli NRI

Kesinleştirilmiş sınıfı bulunmayan durumlarda, sınıf tanımındaki farklılıkların üstesinden gelebilmek için 2011 yılında Pencina ve arkadaşları tarafından önerilmiştir. Herhangi bir risk sınıflaması gerektirmeyen sürekli NRI, olayın görüldüğü ve görülmediği tüm durumlar için kestirilen riskte meydana gelen tüm değişiklikleri dikkate alır (19). Sürekli NRI sadece birkaç durumda önerilmektedir. Bu durumlar;

- Eğer öncelikli hedef model performansından çok modele eklenen belirteçin gücünü incelemek ise sürekli NRI kullanılabilir. Ancak, sürekli NRI'nın büyüklüğünün aşırı yorumlanmamasına dikkat edilmesi gerekir. Sürekli NRI değeri kategorik NRI'ya göre oldukça yüksektir ve incelenen modellerin iyi

ayarlanmış olması gereklidir. Modellerin iyi ayarlanması ise, incelenen modellerin klinik olarak da kullanılabilir olması, kullanılan belirteçlerin model performansında ne gibi etkilere neden olabileceğinin öngörüsünün olması gibi bilgiler içermektedir ve bu bilgiler sürekli NRI değerinin doğru elde edilmesi ve yorumlanması açısından oldukça önemlidir.

- İkidenden fazla kategoriye sahip bir bağımlı değişken için kestirim yapılması amacıyla incelemeler yapıldığında, incelenen modele eklenen belirtecin gücünü incelemek için kullanılabilir. NRI yönteminin çıkarsandığı formül gereği, ikiden çok kategoriye sahip bir bağımlı değişken ile inceleme yapıldığında yeniden sınıflandırma tabloları oluşturulamamaktadır. Bu nedenle sürekli NRI yaklaşımı ile inceleme yapılması tek çözüm olmaktadır.

Sürekli NRI bileşenleri ve açıklamaları Tablo 3.1.'de gösterilmiştir. Burada, Olgu NRI; olgu görülen grupta, risk kestirim modeline yeni bir belirteç eklenmesi durumunda elde edilen olasılık değerinin artışı ya da azalışını hesaplamaktadır. Olgu görülen grupta, yeni belirteç eklenen model sonucunda elde edilen olasılık değerinin yükselmesi “yükselen risk olasılığı”, düşmesi ise “düşen risk olasılığı” durumunu göstermektedir. Aynı şekilde, normal NRI da olgunun görülmediği kişilerde yapılan hesaplamalardır. Olgu NRI’den farkı, normal NRI’da yükselen ve düşen risk durumunun ters olmasıdır. Toplam NRI ise, olgu ve normal NRI değerlerinden elde edilen değerlerin toplamıdır. İki ayrı yüzde değerinin toplanması ile elde edilen bir değer olduğu için yüzde olarak belirtilmesi yanlıştır (20).

Tablo 3.1. Sürekli NRI Bileşenleri

Olgu NRI (Olay Görülen Kişilerde NRI)	$\text{Pr}(\text{yükselen risk olasılığı} \mid \text{olgu}) - \text{Pr}(\text{düşen risk olasılığı} \mid \text{olgu})$ $= (\text{Tahmin edilen riski yükselen olay görülen kişiler} - \text{Tahmin edilen riski düşen olay görülen kişiler}) / \text{Tüm olay görülen kişiler}$
	Tahmin edilen riski yükselen olay görülen kişilerin net yüzdesi, -%100 ve +%100 arasında bir yüzde değeri alabilir
Normal NRI (Olay Görülmeyen Kişilerde NRI)	$\text{Pr}(\text{düşen risk olasılığı} \mid \text{normal}) - \text{Pr}(\text{yükselen risk olasılığı} \mid \text{normal})$ $= (\text{Tahmin edilen riski düşen olay görülmeyen kişiler} - \text{Tahmin edilen riski yükselen olay görülmeyen kişiler}) / \text{Tüm olay görülen kişiler}$
	Tahmin edilen riski düşen olay görülmeyen kişilerin net yüzdesi , -%100 ve +%100 arasında bir yüzde değeri alabilir
Toplam NRI	$[\text{Pr}(\text{yükselen risk olasılığı} \mid \text{olgu}) - \text{Pr}(\text{düşen risk olasılığı} \mid \text{olgu})] + [\text{Pr}(\text{düşen risk olasılığı} \mid \text{normal}) - \text{Pr}(\text{yükselen risk olasılığı} \mid \text{normal})]$ $= \text{Olgu NRI} + \text{Normal NRI}$
	Doğru şekilde farklı risk kestirimi hesaplanan olgu görülen ve normal kişilerin net yüzdelерinin toplamı. Yüzde olarak ifade edilemez, Teorik aralığı -2 ve +2 arasındadır

3.3.3. Ağırlıklandırılmış NRI (wNRI)

Greenland, orijinal NRI çalışmasını yorumladığı çalışmasında, maliyetlerin dikkate alındığı durumda bu ölçünün daha anlamlı olabileceğini önermiştir (25). Bu öneri üzerine Pencina ve arkadaşları 2011 yılında yapmış oldukları çalışmada ağırlıklandırılmış NRI yöntemini önermiştir (19).

Bu çalışmada; üst sınıfa geçmesi gereken bir kişinin (olgu), kurulan model ile üst sınıfta yeniden sınıflanması ile elde edilen kazanım s_1 olarak; alt sınıfa geçmesi gereken bir kişinin (normal), kurulan model ile alt sınıfta yeniden sınıflanması ile elde edilen kazanım s_2 olarak tanımlanmıştır.

En basit haliyle, s_1 : olayın görüleceği kişiye başlanan tedavinin getirdiği kazanım; s_2 : olayın görülmeyeceği kişiye gereksiz bir tedavi uygulamadan kaçınmanın getireceği kazanım olarak örneklenmiştir (26).

$$wNRI = s_1 \cdot (P(olgu|üst) \cdot P(üst) - P(olgu|alt) \cdot P(alt)) + s_2 \cdot (P(normal|alt) \cdot P(alt) - P(normal|üst) \cdot P(üst)) \quad (3.24.)$$

Eşitlik 3.24. ile gösterilen bu hesaplamada; $P(üst) = \frac{n_{üst}}{n}$ ve $P(alt) = \frac{n_{alt}}{n}$ olarak özetlenebilir.

Eğer maliyetler kazanımlardan azsa bu maliyet kazanımlı bir belirteç denilebilir. Risk sınıflamasının amacı doğru tedavi kararlarına varabilmektir. Örneğin, kardiyovasküler hastalıklar için: Düşük risk sınıfı, girişim olmadığını; Orta risk sınıfı, yaşam tarzı değişikliklerini; Yüksek risk sınıfı, hem yaşam tarzı değişikliklerini hem de girişimleri gösterir. Olay görülen bir kişinin sınıfının “yüksek riskten düşük risk sınıfına” geçmesi, “yüksek risk sınıfından orta risk sınıfına” geçmesine oranla daha önemli bir hatadır. Üç ya da daha fazla sayıda kategori olduğu durumda, modele yeni eklenen belirtecin kişiyi geçirebileceği tüm risk sınıflarının incelenmesi gerekir. Üç risk kategorisi için; üst risk sınıfına geçmenin 3 ayrı durumu vardır:

- Düşük riskten orta risk sınıfına
- Düşük riskten yüksek risk sınıfına
- Orta riskten yüksek risk sınıfına
- Her kategori arasındaki geçişe eşit ağırlık verir. Örneğin düşük risk grubundan yüksek risk grubuna geçmek ya da düşük risk grubundan orta risk grubuna geçmenin ağırlıkları aynıdır.
- Farklı tipteki yeniden sınıflandırmaları ağırlıklandırmak oldukça zordur. Ama bu zorluk eşit ağırlıklar kullanmayı haklı çıkarmaz.
- Ağırlıklandırma yapabilmek için ortaya atılmış görüşlerden birisi, her yeniden sınıflandırmanın ağırlığını kişinin atladığı kategori sayısı ile ağırlıklandırmadır.

Tablo 3.2. Olay Görülen ve Görülmeyen Kişilerde Ağırlıklandırılmış Yeniden Sınıflama

	Eski / Yeni	Düşük Risk	Orta Risk	Yüksek Risk	Toplam
Olay Görülen Kişilerde Yeniden Sınıflama (event)	Düşük Risk	e_{11}	e_{12}	e_{13}	e_{1*}
	Orta Risk	e_{21}	e_{22}	e_{23}	e_{2*}
	Yüksek Risk	e_{31}	e_{32}	e_{33}	e_{3*}
	Toplam	e_{*1}	e_{*2}	e_{*3}	e_{**}
	Eski / Yeni	Düşük Risk	Orta Risk	Yüksek Risk	Toplam
Olay Görülmeyen Kişilerde Yeniden Sınıflama (non event)	Düşük Risk	n_{11}	n_{12}	n_{13}	n_{1*}
	Orta Risk	n_{21}	n_{22}	n_{23}	n_{2*}
	Yüksek Risk	n_{31}	n_{32}	n_{33}	n_{3*}
	Toplam	n_{*1}	n_{*2}	n_{*3}	n_{**}

Tablo 3.2.'de, örnek bir yeniden sınıflandırma tablosu görülmektedir. Burada, olay görülen ve görülmeyen kişilerde ağırlıklandırılmış bir yaklaşım ile olgu ve sağlıklı NRI değerlerinin ne şekilde hesaplanabileceği Eşitlik 3.25. ve Eşitlik 3.26. ile gösterilmiştir.

$$\begin{aligned}
\text{Olgu NRI} &= \frac{e_{12} + 2e_{13} + e_{23}}{e_{**}} - \frac{e_{21} + 2e_{31} + e_{32}}{e_{**}} \\
&= 2 * P(2 \text{ sınıf üst} | D=1) + P(1 \text{ sınıf üst} | D=1) - P(1 \text{ sınıf alt} | D=1) - 2 * P(2 \text{ sınıf alt} | D=1) \\
&= \sum_{k=1}^{K-1} k. \{P(k \text{ kategori üst} | D = 1) - P(k \text{ kategori alt} | D = 1)\} \quad (3.25.)
\end{aligned}$$

$$\begin{aligned}
\text{Normal NRI} &= \frac{n_{21} + 2n_{31} + n_{32}}{n_{**}} - \frac{n_{12} + 2n_{13} + n_{23}}{n_{**}} \\
&= 2 * P(2 \text{ sınıf alt} | D=0) + P(1 \text{ sınıf alt} | D=0) - P(1 \text{ sınıf üst} | D=0) - 2 * P(2 \text{ sınıf üst} | D=0) \\
&= \sum_{k=1}^{K-1} k. \{P(k \text{ kategori alt} | D = 0) - P(k \text{ kategori üst} | D = 0)\} \quad (3.26.)
\end{aligned}$$

Toplam NRI = Olgu NRI + Normal NRI olarak tanımlanmıştır.

Üç kategorili sınıflama için genelleştirilmiş NRI Eşitlik 3.27. ile gösterilmiştir.

Burada;

- D: Düşük; O: Orta; Y: Yüksek olmak üzere;
- Üç kategori için; üst sınıfa geçiş 3 anlama gelebilir:
düşükten ortaya geçiş, düşükten yükseğe geçiş, ortadan yükseğe geçiş

$$\begin{aligned}
\text{NRI} &= P(DO | olgu) + P(DY | olgu) + P(OY | olgu) \\
&\quad - P(OD | olgu) - P(YD | olgu) - P(YO | olgu) \\
&\quad + P(OD | normal) + P(YD | normal) + P(YO | normal) \\
&\quad - P(DO | normal) - P(DY | normal) - P(OY | normal) \quad (3.27) \\
&= [P(olgu | DO) P(DO) + P(olgu | DY) P(DY) + P(olgu | OY) P(OY)] / P(olgu) \\
&\quad - [P(olgu | OD) P(OD) + P(olgu | YD) P(YD) + P(olgu | YO) P(YO)] / P(olgu) \\
&\quad + [P(normal | OD) P(OD) + P(normal | YD) P(YD) + P(normal | YO) P(YO)] / P(normal) \\
&\quad - [P(normal | DO) P(DO) + P(normal | DY) P(DY) + P(normal | OY) P(OY)] / P(normal)
\end{aligned}$$

wNRI, olgu NRI ve normal NRI'nın ayrı ayrı gösterilmesinden ya da bir yeniden sınıflandırma tablosundan elde edilebilir. En geniş biçimde, ağırlıklandırılmış NRI, eski model yerine yeni model kullanılarak elde edilen sonuçların ortalama kazanımları

(maddi ya da kaliteli yaşam yılı) olarak yorumlanabilir. Ağırlıklandırılmamış kategorik NRI, yeni belirteç/belirteçlerin modele olan katkısının belirlenmesi işleminde erken aşama olarak kabul edilir (19).

3.4. IDI kestirimi

NRI'da olduğu gibi IDI da, eski modele yeni bir belirteç eklendiğinde elde edilen risk kestirimi olasılıklarının olayın görüldüğü ve görülmediği kişileri birbirinden ne kadar iyi ayırdığını ölçmektedir (27). Burada eski model ve yeni model arasındaki farklılık iki modelin ayırmsama eğimleri ile belirlenmektedir. Ayırmsama eğimlerinin hesaplanmasında kullanılan iki kavram bulunmaktadır. “Duyarlılık” ve “1 – Seçicilik” kavramları 3.28. ve 3.29.’da bulunan eşitliklerle gösterilebilir;

$$\text{Duyarlılık (u)} = S(u) = P(X > u | D = 1) = \int_u^1 f(x | D = 1) dx \quad (3.28.)$$

$$1 - \text{Seçicilik (u)} = P(u) = P(X > u | D = 0) = \int_u^1 f(x | D = 0) dx \quad (3.29.)$$

X , T zamanından önce olayın ortaya çıkmasının kestirilen olasılığını ve D de olayla ilgili durumu gösterdiğinde, $0 < u < 1$ aralığında u gibi herhangi bir kesim noktası için f , x 'in olasılık yoğunluk fonksiyonu olduğunda Eşitlik 3.28. ve Eşitlik 3.29. elde edilecektir.

Birleştirilmiş duyarlılık (IS) ve birleştirilmiş 1 – seçicilik (IP) değerleri ise eşitlik 3.30. ve eşitlik 3.31.’de görüldüğü gibi tanımlanabilir (6).

$$IS = \int_0^1 S(u) du = \int_0^1 \int_u^1 f(x | D = 1) dx du \quad (3.30.)$$

$$IP = \int_0^1 P(u) du = \int_0^1 \int_u^1 f(x | D = 0) dx du \quad (3.31.)$$

Eşitlik 3.30. ve Eşitlik 3.31. de gösterilmiş olan birleştirilmiş duyarlılık ve birleştirilmiş 1- seçicilik değerlerinin formüllerinde yer alan integrallerin sırası değiştirildiğinde ise olay görülmesi ve görülmemesi durumlarının örneklem ortalamaları kullanılarak koşullu beklenen değerler kestirilebilir. Bu durum Eşitlik 3.32. ve Eşitlik 3.33.’te görülebilir.

$$IS = \int_0^1 \int_0^x f(x|D = 1) du dx = \int_0^1 xf(x|D = 1) dx = E(X|D = 1) \quad (3.32.)$$

$$IP = \int_0^1 \int_0^x f(x|D = 0) du dx = \int_0^1 xf(x|D = 0) dx = E(X|D = 0) \quad (3.33.)$$

0 ve 1 aralığında, duyarlılık değerlerinin ve 1-seçicilik değerlerinin integralleri ortalama duyarlılık ve ortalama 1-seçicilik olarak görülebilir. IDI, ortalama duyarlılıkta meydana gelen artış ve 1-seçicilik değeri ortalamasında meydana gelecek potansiyel artış arasındaki fark olarak da görülebilir (6).

$$IDI = (IS_{yeni} - IS_{eski}) - (IP_{yeni} - IP_{eski}) \quad (3.34.)$$

- $\bar{p}_{yeni,olgular}$: olayı geliştiren kişilerin yeni modele dayalı tahmin edilen olasılıklarının ortalaması
- $\bar{p}_{eski,olgular}$: eski modele göre karşılık gelen değer
- $\bar{p}_{yeni,sağlıklılar}$: olayı geliştirmeyen kişilerin yeni modele dayalı tahmin edilen olay görülme olasılıklarının ortalaması
- $\bar{p}_{eski,sağlıklılar}$: eski modele göre karşılık gelen değer

IDI kestirimi Eşitlik 3.35. ile gösterilmiştir. Bu eşitlik, yeni model ve eski model arasındaki ayırmsama eğrileri arasındaki farktır (28).

$$\widehat{IDI} = (\bar{p}_{yeni,olgular} - \bar{p}_{yeni,sağlıklılar}) - (\bar{p}_{eski,olgular} - \bar{p}_{eski,sağlıklılar}) \quad (3.35.)$$

3.5. Uygulamalar

Bu çalışmada kullanılan tüm veri setlerinde incelenen modellemeler çok kategorili (multinomial) lojistik regresyon analizi yöntemi ile yapılmıştır. Ayrıca mikrodizin veri seti ile incelenen modellerden elde edilen sınıflandırma performansları Kuadratik Diskriminant Analizi yöntemi ile de incelenmiş ve elde edilen sonuçlar kullanmış olduğumuz modellemeler sonucunda elde edilen sonuçlarla karşılaştırılmıştır. Kuadratik Diskriminant Analizi yöntemi, verilerin normal dağılım gösterdiği ancak varyans – kovaryans matrislerinin benzer olmaması durumunda kullanılabilecek farklı bir ayırma analizi yaklaşımıdır (29).

İlk olarak, modellerin sınıflandırma performansı sonuçları klasik yöntemler olan HUM ve CCP ile incelenmiştir. İkinci olarak da incelenen tanı tahmin modeline yeni bir belirteç eklendiğinde tanı performansında meydana gelen artışlar yeni incelenen yöntemler olan NRI ve IDI yöntemlerinden elde edilen sonuçlarla karşılaştırılmıştır. Bağımlı değişken ile modelleme için kullanılan belirteçler arasındaki ilişkiler ve klasik yöntemlerden elde edilen sonuçlar ile yeni yöntemlerden elde edilen sonuçlar arasındaki ilişkiler Spearman Korelasyon Analizi ile incelenmiştir. Ayrıca yapılmış olan incelemelerin tümünün örneklem genişliğinden ne şekilde etkilendiği 100, 500 ve 1000 birimlik veri setleri ile incelenmiştir.

Verilerin analizi ve benzetim çalışmaları R Studio (Version 1.1.453 – © 2009-2018 RStudio, Inc.) (30) programı kullanılarak gerçekleştirilmiştir. Bu süreçte R Studio programında, verilerin üretilmesi ve analizlerin yapılması aşamasında “OrdNor”, “MASS”, “bcROCsurface”, “mcca”, “PredictABEL” ve “gtools” paketleri kullanılmıştır (22, 31 - 35). Ayrıca, Kuadratik Diskriminant Analizi için IBM SPSS (IBM SPSS Statistics 24 software (Armonk, NY: IBM Corp.)) paket programı kullanılmıştır (36).

3.5.1. Uygulama 1

Bu uygulamada, önemli belirteçlerin belirlenmesinde, bazı ilişki yapılarının ve örnek genişliklerinin ne şekilde sonuç verdiği incelenmeye çalışılmıştır.

“OrdNor” paketi çok değişkenli normal dağılım varsayımı altında veri üretmektedir. Tablo 3.3.’te gösterilmiş olan ilişki yapıları pakete yüklenmiştir. Üretim aşamasında, üretilen verilerin ilişki matrisinin determinantının pozitif olması gerektiğinden, pakete yüklenen ilişki yapısına en yakın yapıya sahip ve determinantı pozitif olan veriler üretilmiştir. Üretilmiş olan verilerden elde edilen ve bu uygulamada kullanılan ilişki yapıları ise Tablo 3.4. ve Tablo 3.5.’te gösterilmiştir.

Tablo 3.3. Benzetim çalışması için OrdNor paketine yüklenen ve incelenmesi planlanan ilişki yapıları

		y	X1	X2	X3	X4	X5
Pozitif İlişki Yapısı	y	1	0.930	0.910	0.310	0.110	0.110
	X1		1	0.920	0.810	0.150	0.180
	X2			1	0.80	0.170	0.120
	X3				1	0.110	0.120
	X4					1	0.170
	X5						1
		y	X1	X2	X3	X4	X5
Negatif İlişki Yapısı	y	1	-0.930	-0.910	-0.310	-0.110	-0.110
	X1		1	-0.920	-0.810	-0.150	-0.180
	X2			1	-0.80	-0.170	-0.120
	X3				1	-0.110	-0.120
	X4					1	-0.170
	X5						1

Benzetim yöntemi ile 3 kategorili ve sıralı yapıda ($y=1,2,3$) olan bir bağımlı değişken elde edilmiştir. Ayrıca bu bağımlı değişken ile farklı düzeylerde ilişki yapılarına sahip 5 sürekli bağımsız değişken yaratılarak tüm tanı modelleri farklı örnek genişlikleri ile incelenmiştir. Bağımlı değişken ile bağımsız değişkenler arasındaki ilişkilerin tanı performansını ne şekilde etkilediği açıklanmaya çalışılmıştır.

Ek olarak, bağımsız değişkenlerin kendi aralarında olan ilişkilerin de tanı performansı üzerindeki etkileri elde edilmeye çalışılmıştır.

İncelenen ilişki yapıları:

- Pozitif ilişkiler:
 - X1 ve X2, Bağımlı değişken (Y) ile kuvvetli kendi aralarında da kuvvetli bir ilişkiye sahip
 - X3, Y ile orta düzey, X1 ve X2 ile kuvvetliye yakın bir ilişkiye sahip
 - X4 ve X5 ise hem Y ile hem de diğer bağımsız değişkenlerle ilişkisiz ya da çok düşük bir ilişkiye sahip

Tablo 3.4. Benzetim yoluyla elde edilmiş, pozitif ilişki yapısına sahip veri setinde elde edilmiş olan ilişkiler (Spearman ilişki katsayıları, r)

n=100	Y	X1	X2	X3	X4	X5
Y	1	0,813	0,816	0,436	0,121	0,072
X1		1	0,997	0,655	0,190	0,211
X2			1	0,673	0,207	0,144
X3				1	0,169	-0,019
X4					1	0,178
X5						1
n=500	Y	X1	X2	X3	X4	X5
Y	1	0,811	0,809	0,371	0,095	0,050
X1		1	0,998	0,640	0,139	0,099
X2			1	0,650	0,155	0,042
X3				1	0,144	0,023
X4					1	0,188
X5						1
n=1000	Y	X1	X2	X3	X4	X5
Y	1	0,799	0,795	0,355	0,134	0,112
X1		1	0,998	0,652	0,170	0,133
X2			1	0,661	0,186	0,076
X3				1	0,132	0,040
X4					1	0,190
X5						1

- Negatif ilişkiler:

- X1 ve X2, Bağımlı değişken (Y) ile orta düzeyde kendi aralarında ilişkisiz ya da çok düşük bir ilişkiye sahip

- X3, Y ile ilişkisiz, X1 ve X2 ile orta düzey bir ilişkiye sahip

- X4 ve X5 ise hem Y ile hem de diğer bağımsız değişkenlerle ilişkisiz ya da çok düşük bir ilişkiye sahip

Tablo 3.5. Benzetim yoluyla elde edilmiş, negatif ilişki yapısına sahip veri setinde elde edilmiş olan ilişkiler (Spearman ilişki katsayıları, r)

n=100	Y	X1	X2	X3	X4	X5
Y	1	-0,408	-0,400	0,059	0,058	-0,168
X1		1	-0,170	-0,182	-0,049	-0,115
X2			1	-0,192	-0,045	0,037
X3				1	0,045	-0,269
X4					1	-0,157
X5						1
n=500	Y	X1	X2	X3	X4	X5
Y	1	-0,324	-0,390	0,063	-0,048	0,019
X1		1	-0,220	-0,263	0,049	-0,188
X2			1	-0,244	-0,034	0,041
X3				1	-0,011	-0,071
X4					1	-0,124
X5						1
n=1000	Y	X1	X2	X3	X4	X5
Y	1	-0,334	-0,390	0,072	-0,024	0,022
X1		1	-0,150	-0,307	-0,042	-0,082
X2			1	-0,319	-0,042	-0,001
X3				1	-0,010	-0,097
X4					1	-0,126
X5						1

Her ilişki yapısı için;

n= 100

n=500

n=1000

örneklemlik veri setleri üretilmiş ve Tablo 3.6.'da belirtilmiş olan model kombinasyonlarının sonuçları incelenmiştir.

Tablo 3.6. İncelenen modeller

	İki belirteçli modeller	Üç belirteçli modeller	Dört belirteçli modeller	Beş belirteçli modeller
	x_1+x_2	$x_1+(x_2+x_3)$	$x_1+(x_2+x_3+x_4)$	$x_1+(x_2+x_3+x_4+x_5)$
Modelin ilk belirtecini	x_1+x_3	$x_1+(x_2+x_4)$	$x_1+(x_2+x_3+x_5)$	
	x_1+x_4	$x_1+(x_2+x_5)$	$x_1+(x_3+x_4+x_5)$	
X1 olduğu durum	x_1+x_5	$x_1+(x_3+x_4)$ $x_1+(x_3+x_5)$ $x_1+(x_4+x_5)$		
	x_2+x_1	$x_2+(x_1+x_3)$	$x_2+(x_1+x_3+x_4)$	$x_2+(x_1+x_3+x_4+x_5)$
Modelin ilk belirtecini	x_2+x_3	$x_2+(x_1+x_4)$	$x_2+(x_1+x_3+x_5)$	
	x_2+x_4	$x_2+(x_1+x_5)$	$x_2+(x_3+x_4+x_5)$	
X2 olduğu durum	x_2+x_5	$x_2+(x_3+x_4)$ $x_2+(x_3+x_5)$ $x_2+(x_4+x_5)$		
	x_3+x_2	$x_3+(x_2+x_1)$	$x_3+(x_2+x_1+x_4)$	$x_3+(x_2+x_1+x_4+x_5)$
Modelin ilk belirtecini	x_3+x_1	$x_3+(x_2+x_4)$	$x_3+(x_2+x_1+x_5)$	
	x_3+x_4	$x_3+(x_2+x_5)$	$x_3+(x_1+x_4+x_5)$	
X3 olduğu durum	x_3+x_5	$x_3+(x_1+x_4)$ $x_3+(x_1+x_5)$ $x_3+(x_4+x_5)$		
	x_4+x_2	$x_4+(x_2+x_3)$	$x_4+(x_2+x_3+x_1)$	$x_4+(x_2+x_3+x_1+x_5)$
Modelin ilk belirtecini	x_4+x_3	$x_4+(x_2+x_1)$	$x_4+(x_2+x_3+x_5)$	
	x_4+x_1	$x_4+(x_2+x_5)$	$x_4+(x_3+x_1+x_5)$	
X4 olduğu durum	x_4+x_5	$x_4+(x_3+x_1)$ $x_4+(x_3+x_5)$ $x_4+(x_1+x_5)$		
	x_5+x_2	$x_5+(x_2+x_3)$	$x_5+(x_2+x_3+x_4)$	$x_5+(x_2+x_3+x_4+x_1)$
Modelin ilk belirtecini	x_5+x_3	$x_5+(x_2+x_4)$	$x_5+(x_2+x_3+x_1)$	
	x_5+x_4	$x_5+(x_2+x_1)$	$x_5+(x_3+x_4+x_1)$	
X5 olduğu durum	x_5+x_1	$x_5+(x_3+x_4)$ $x_5+(x_3+x_1)$ $x_5+(x_4+x_1)$		

3.5.2. Uygulama 2

Bu çalışmada kullanılmış olan tiroit hastalığına ait veriler UCI (University of California, Irvine) makine öğrenmesi veri tabanından alınmıştır (<http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease>) (23). Orijinal veri setinde 166 adet sağlıklı, 368 adet hipertiroidi ve 6666 adet hipotiroidi hastasına ait toplamda 7200 adet örnek bulunmaktadır.

- Bağımlı değişken olarak Tiroid hastalığı incelenmiştir.

İncelenen Tiroid hastalığının sınıfları:

- Sağlıklı, Hipertiroidi ve Hipotiroidi olarak sırasız yapıda kabul edilerek incelenmiştir.
- İncelenen Bağımsız Değişkenler:

TSH: Tiroit Uyarıcı Hormon

T3: Triiyodotreonin

TT4: Toplam Troksin

T4U: Tiroksin Kullanım Oranı

FTI: Serbest Tiroksin İndeksi

Örneklem genişliği ile ilgili incelemeler için;

n= 100

n=500

n=1000

örneklemlik rastgele veri setleri yaratılmıştır ve farklı model kombinasyonlarının sonuçları incelenmiştir. İncelenen veri setlerinde, orijinal verideki grupların oranları dengesiz olduğu için, hesaplamalarda kullanılabilecek ve orijinal veri setine benzer olabilecek şekilde örnekler rastgele seçilmiştir. “n=100” olan küçük örneklem genişliğine sahip veri setinde, 15 adet sağlıklı, 30 adet hipertiroidi ve 55 adet

hipotiroidi hastası, “n=500” olan orta örneklem genişliğine sahip veri setinde, 75 adet sağlıklı, 150 adet hipertiroidi ve 275 adet hipotiroidi hastası ve “n=1000” olan büyük örneklem genişliğine sahip veri setinde, 150 adet sağlıklı, 300 adet hipertiroidi ve 550 adet hipotiroidi hastası yer almaktadır.

Bu veri seti ile olan incelemelerde 325 adet HUM ve CCP hesaplanan model, 320 adet de NRI ve IDI hesaplanan model kullanılmıştır. Her belirtecin modele ilk eklenen belirteç olduğu durum ve bu belirtecin arkasından modele eklenebilecek olan diğer belirteçlerin sırasının performans sonuçlarına nasıl bir katkı yaptığı incelenmiştir.

3.5.3. Uygulama 3

Yapılan çalışmada, Golub ve arkadaşlarının çalışmasından alınan 7129 gen ekspresyonunun verisi kullanılmıştır. Veriler, Üç farklı tip akut lösemide bulunan gen ekspresyonu verilerini içermektedir (24). Bu akut lösemi tipleri, T hücrelerinden kaynaklanan (ALL T hücresi) akut lenfoblastik lösemi, B hücrelerinden kaynaklanan ALL (ALL B hücresi), ve akut miyeloid lösemi (AML) tipleridir. Veri seti, 8 adet ALL T-hücre örneği, 19 adet ALL B-hücre örneği ve 11 adet AML örneği içermektedir ve her örnek, affymetrix yüksek yoğunluklu oligonükleotid mikrodizilerinden elde edilen 7129 gen ekspresyon değerini içerir.

Öncelikle tüm genler için HUM ve CCP değerleri hesaplanmıştır. HUM değerlerinin en yüksek olduğu 5 gen (VERİ SETİ 1) ile ve HUM değerlerinin yüksek, orta ve düşük düzey olduğu 5 gen (VERİ SETİ 2) ile iki ayrı uygulama yapılmıştır. Bu şekilde bir inceleme yapılmasının nedeni, bu genler ile kurulabilecek modelleri tek gen ile kurulacak modeller ile karşılaştırarak inceleyebilmektir. Ayrıca, bu iki ayrı veri setine uygulanan Kuadratik Diskriminant analizi sonuçları da incelenerek karşılaştırmalı olarak inceleme yapılması amaçlanmıştır.

Kuadratik Diskriminant Analizi, alanyazıda kabul görmüş ve sınıflandırma modellerinin incelendiği çalışmalarda sıkça kullanılan bir sınıflandırma yöntemidir. Bu karşılaştırmalarda amaç, en yüksek sınıflandırma performansını veren modellerde bulunan belirteçlerin belirlenmesi ve bu belirteçlerin modele girme sırasının literatürde kabul görmüş bir yöntemle karşılaştırılarak incelenebilmesidir.

4. BULGULAR

4.1. Uygulama 1 - Bağımlı değişkenin sıralı yapıda olduğu durum için örnek uygulama (Benzetim çalışması)

Modellerde kullanılan belirteç sayısı ($i = 1, \dots, 5$) ile bu modellere ait HUM, CCP, NRI ve IDI değerleri arasındaki ilişkilerin örneklem genişliği açısından incelenmesi Tablo 4.1.'de gösterilmiştir. Örneklem genişliğinin düşük olduğu durumlarda, hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerden elde edilen modellerin performans sonuçları ile bu modellerde bulunan belirteç sayıları arasındaki ilişkilerin büyük örneklem genişliğine sahip modellere göre daha düşük olduğu görülmüştür.

Tablo 4.1. Modelde bulunan belirteç sayısı ile incelenen modellerin performans sonuçları arasındaki ilişkiler

n	Pozitif İlişki Yapısında Benzetim Yapılan Veriler				Negatif İlişki Yapısında Benzetim Yapılan Veriler			
	HUM	CCP	NRI	IDI	HUM	CCP	NRI	IDI
100	0,745	0,682	0,292	0,346	0,765	0,302	0,420	0,753
500	0,902	0,917	0,361	0,463	0,692	0,872	0,583	0,744
1000	0,937	0,912	0,511	0,539	0,767	0,887	0,626	0,738

Tüm incelemelerde Spearman Korelasyon Analizi kullanılmıştır.

Tablo 4.1.'den görüleceği gibi, pozitif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerde, HUM değerleri için; örneklem genişliği arttıkça modelde bulunan belirteç sayısı ile ilişkilerin kuvvetlendiği görülmüştür, ancak bu durumun negatif ilişki yapısında benzetim yapılarak oluşturulan modellerde sağlanmadığı belirlenmiştir. Hem pozitif hem de negatif ilişki yapısına sahip modellerde büyük örneklem genişliklerinde, modelde bulunan belirteç sayısı ile CCP değerleri arasında elde edilen ilişkilerin küçük örneklem genişliğine göre daha güçlü olduğu görülmüştür. IDI değerlerinde ise; negatif ilişki yapısına sahip modeller, tüm örnek genişliklerinde modelde bulunan belirteç sayısı ile kuvvetli ilişkili sonuçlar verirken, pozitif ilişki yapısına sahip modeller örneklem genişliğinden etkilenmiştir. Ancak negatif ilişki yapısına sahip modellerde, modelde bulunan belirteç sayısı ile IDI değerleri arasında yüksek ilişki elde edilmiş olmasına rağmen bu düzeyde bir ilişki pozitif ilişki yapısına sahip olan modellerde görülmemiştir.

Tablo 4.2.'de pozitif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerde bulunan belirteç sayılarına göre elde edilen modellerin performans sonuçları gösterilmiştir. Modelde bulunan belirteç sayısının 5 olduğu durumda incelenen performans sonuçlarının en yüksek değerlere ulaştığı görülmüştür. Negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerde de; örneklem genişliğinin düşük olduğu durumda en yüksek değerler 4 belirteç olan modellerden elde edilmiştir. Diğer örneklem genişliklerinde ise, en yüksek performans sonuçları pozitif ilişki yapılarında olduğu gibi 5 belirteçli modellerde elde edilmiştir (Tablo 4.3.). Tek belirteç ile kurulan modellerde, modele ikinci bir belirteç eklenmemiş olduğu için NRI ve IDI değerleri elde edilmemektedir.



Tablo 4.2. Pozitif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin, modele eklenen belirteç sayılarına göre performansları

n	Belirteç Sayısı		HUM	CCP	NRI	IDI
100	1	Minimum	0,1984	0,3397	-	-
		Maximum	0,8071	0,7453	-	-
	2	Minimum	0,2371	0,3654	-0,0003	-0,0004
		Maximum	0,8267	0,7726	0,4250	0,5476
	3	Minimum	0,4157	0,4765	0,0003	0,0118
		Maximum	0,8305	0,7726	0,4329	0,6056
	4	Minimum	0,8062	0,7603	0,0180	0,0204
		Maximum	0,8211	0,7726	0,4363	0,6082
	5	Minimum	0,8218	0,7726	0,0273	0,0734
		Maximum	0,8218	0,7726	0,4343	0,6099
n			HUM	CCP	NRI	IDI
500	1	Minimum	0,1842	0,3442	-	-
		Maximum	0,7637	0,7548	-	-
	2	Minimum	0,2007	0,3659	-0,0086	-0,0001
		Maximum	0,8237	0,7734	0,4176	0,5051
	3	Minimum	0,3493	0,4470	-0,0108	0,0004
		Maximum	0,8239	0,7747	0,4305	0,5752
	4	Minimum	0,7968	0,7619	0,0124	0,0563
		Maximum	0,8232	0,7753	0,4287	0,5762
	5	Minimum	0,8249	0,7766	0,0620	0,0923
		Maximum	0,8249	0,7766	0,4653	0,5906
n			HUM	CCP	NRI	IDI
1000	1	Minimum	0,2141	0,3432	-	-
		Maximum	0,7404	0,7432	-	-
	2	Minimum	0,2262	0,3624	-0,0043	0,0002
		Maximum	0,8067	0,7732	0,3958	0,4977
	3	Minimum	0,3517	0,4434	-0,0110	0,0020
		Maximum	0,8082	0,7781	0,4320	0,5555
	4	Minimum	0,7963	0,7694	0,0261	0,0506
		Maximum	0,8089	0,7755	0,4336	0,5576
	5	Minimum	0,8648	0,8042	0,1301	0,1731
		Maximum	0,8648	0,8042	0,5158	0,6323

Tablo 4.3. Negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin, modele eklenen belirteç sayılarına göre performansları

n	Belirteç Sayısı		HUM	CCP	NRI	IDI
100	1	Minimum	0,1826	0,3333	-	-
		Maximum	0,3356	0,4665	-	-
	2	Minimum	0,1781	0,3377	-0,0119	0,0001
		Maximum	0,4674	0,5465	0,1407	0,1423
	3	Minimum	0,2681	0,3831	-0,0271	0,0007
		Maximum	0,4844	0,5465	0,2132	0,2462
	4	Minimum	0,3378	0,3810	-0,0390	0,0199
		Maximum	0,4945	0,5498	0,2165	0,2987
	5	Minimum	0,4908	0,5422	0,0758	0,2086
		Maximum	0,4908	0,5422	0,2089	0,2997
n			HUM	CCP	NRI	IDI
500	1	Minimum	0,1785	0,3321	-	-
		Maximum	0,3707	0,4605	-	-
	2	Minimum	0,1933	0,3529	-0,0075	0,0005
		Maximum	0,4797	0,5581	0,1378	0,1313
	3	Minimum	0,2114	0,3680	0,0111	0,0017
		Maximum	0,5090	0,5732	0,2261	0,2082
	4	Minimum	0,3142	0,4665	0,0277	0,0040
		Maximum	0,5134	0,5685	0,2257	0,2144
	5	Minimum	0,5072	0,5785	0,1181	0,1405
		Maximum	0,5072	0,5785	0,2464	0,2168
n			HUM	CCP	NRI	IDI
1000	1	Minimum	0,1763	0,3333	-	-
		Maximum	0,3668	0,4731	-	-
	2	Minimum	0,1892	0,3361	-0,0042	0,0002
		Maximum	0,4661	0,5479	0,1483	0,1185
	3	Minimum	0,2037	0,3635	-0,0016	0,0008
		Maximum	0,5076	0,5788	0,2197	0,2087
	4	Minimum	0,3076	0,4395	0,0087	0,0023
		Maximum	0,5116	0,5762	0,2402	0,2141
	5	Minimum	0,5113	0,5836	0,1104	0,1374
		Maximum	0,5113	0,5836	0,2502	0,2174

Her modelden elde edilen performans sonuçlarının kendi aralarındaki ilişkiler, modellerde bulunan belirteç sayıları açısından incelendiğinde, tek belirteçli modellerde NRI ve IDI yöntemlerinden sonuç elde edilememiştir. Bunun nedeni, bu yöntemlerin tanı modeline yeni bir belirteç eklendiği durumda oluşan değişimi gösteren ölçüler olmasıdır.

İki belirteçli modeller incelendiğinde ise; negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellere ait HUM ve CCP değerleri ile NRI ve IDI değerleri arasındaki ilişkilerin sonuçlarının, pozitif ilişki yapısındaki modellere ait HUM ve CCP değerleri ile NRI ve IDI değerleri arasındaki ilişkilere göre daha kuvvetli olduğu görülmüştür. Ayrıca NRI ve IDI ile HUM arasındaki ilişkiler, pozitif ilişki yapısında örneklem genişliği arttıkça düşüş göstermiş, negatif ilişki yapısında ise örneklem genişliği arttıkça yükseliş göstermiştir (Tablo 4.4.).

Tablo 4.4. İki belirteçli modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,890 (0,000)	0,353 (0,126)	0,287 (0,220)
	CCP	-	1	0,336 (0,148)	0,047 (0,844)
	NRI	-	-	1	0,762 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,891 (0,000)	0,211 (0,371)	0,432 (0,057)
	CCP	-	1	0,284 (0,225)	0,495 (0,026)
	NRI	-	-	1	0,914 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,964 (0,000)	0,148 (0,534)	0,338 (0,145)
	CCP	-	1	0,214 (0,364)	0,368 (0,110)
	NRI	-	-	1	0,901 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,733 (0,000)	0,298 (0,202)	0,565 (0,010)
	CCP	-	1	0,432 (0,057)	0,468 (0,037)
	NRI	-	-	1	0,705 (0,001)
	IDI	-	-	-	1
500	HUM	1	0,891 (0,000)	0,408 (0,074)	0,522 (0,018)
	CCP	-	1	0,519 (0,019)	0,543 (0,013)
	NRI	-	-	1	0,869 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,997 (0,000)	0,502 (0,024)	0,613 (0,004)
	CCP	-	1	0,506 (0,023)	0,606 (0,005)
	NRI	-	-	1	0,869 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Üç belirteçli modellere bakıldığında; negatif ilişki yapısındaki modellere ait HUM ve CCP değerleri ile NRI ve IDI değerleri arasındaki ilişkilerin pozitif ilişki yapısında elde edilen ilişkilere göre daha kuvvetli olduğu görülmüştür. Hem pozitif hem de negatif ilişki yapılarında yapılan incelemelerde, NRI ve IDI ile HUM değerleri arasındaki ilişkiler, iki belirteçli modellerde görüldüğünün aksine, örneklem genişliği arttıkça yükselmiştir. Negatif ilişki yapısında, IDI ile HUM değerleri arasında görülen ilişkiler örneklem genişliğinden etkilenmemiş, tüm incelemelerde kuvvetli ilişkiler elde edilmiştir. Ancak NRI ile HUM değerleri arasındaki ilişkiler örneklem genişliği arttıkça yükselmiştir (Tablo 4.5.).

Tablo 4.6.'da dört belirteçli modellerden elde edilen ilişkiler gösterilmiştir. Pozitif ilişki yapısında elde edilen HUM ve CCP değerleri ile NRI ve IDI değerleri arasındaki ilişkilerin, negatif ilişki yapısındaki modellerden elde edilen HUM ve CCP değerleri ile NRI ve IDI yöntemleri arasındaki ilişkilere göre daha zayıf düzeyde olduğu görülmüştür. Pozitif ilişki yapısında NRI ve IDI ile HUM arasındaki ilişkiler örneklem genişliğinden etkilenmemiştir ve benzer sonuçlar görülmüştür. Negatif ilişki yapısında ise örneklem genişliği arttıkça HUM ve IDI arasındaki ilişkiler yükselmiştir.

Beş belirteçli modellerde ise, HUM ve CCP değerleri modele eklenen belirtecin sırasından etkilenmediği için ve bu modeller tam model olduğu için, tüm modellerde aynı sonuç elde edilmiştir. Bu sebeple, HUM ve CCP değerlerinin NRI ve IDI değerleri ile olan ilişkileri incelenememiştir.

Tablo 4.5. Üç belirteçli modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,640 (0,000)	0,113 (0,389)	0,216 (0,097)
	CCP	-	1	0,356 (0,005)	0,335 (0,009)
	NRI	-	-	1	0,957 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,754 (0,000)	0,170 (0,193)	0,483 (0,000)
	CCP	-	1	0,344 (0,007)	0,531 (0,000)
	NRI	-	-	1	0,881 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,754 (0,000)	0,172 (0,188)	0,521 (0,000)
	CCP	-	1	0,291 (0,024)	0,545 (0,000)
	NRI	-	-	1	0,888 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,869 (0,000)	0,543 (0,000)	0,856 (0,000)
	CCP	-	1	0,672 (0,000)	0,779 (0,000)
	NRI	-	-	1	0,793 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,976 (0,000)	0,624 (0,000)	0,822 (0,000)
	CCP	-	1	0,646 (0,000)	0,825 (0,000)
	NRI	-	-	1	0,851 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,952 (0,000)	0,628 (0,000)	0,801 (0,000)
	CCP	-	1	0,659 (0,000)	0,791 (0,000)
	NRI	-	-	1	0,884 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Tablo 4.6. Dört belirteçli modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,725 (0,000)	0,105 (0,255)	0,217 (0,017)
	CCP	-	1	0,033 (0,722)	0,260 (0,004)
	NRI	-	-	1	0,903 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,224 (0,014)	0,191 (0,037)	0,331 (0,000)
	CCP	-	1	0,110 (0,231)	0,123 (0,180)
	NRI	-	-	1	0,969 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,359 (0,000)	0,117 (0,205)	0,263 (0,004)
	CCP	-	1	0,264 (0,004)	0,226 (0,013)
	NRI	-	-	1	0,973 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,667 (0,000)	0,553 (0,000)	0,692 (0,000)
	CCP	-	1	0,634 (0,000)	0,849 (0,000)
	NRI	-	-	1	0,850 (0,000)
	IDI	-	-	-	1
500	HUM	1	1	0,356 (0,000)	0,809 (0,000)
	CCP	-	1	0,356 (0,000)	0,809 (0,000)
	NRI	-	-	1	0,738 (0,000)
	IDI	-	-	-	1
1000	HUM	1	1	0,503 (0,000)	0,852 (0,000)
	CCP	-	1	0,503 (0,000)	0,852 (0,000)
	NRI	-	-	1	0,834 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Bağımlı değişkenle modelde bulunan belirteçler arasındaki ilişkilerin modelden elde edilen performans sonuçlarını ne şekilde etkilediğini inceleyebilmek amacıyla, farklı belirteçlerin modelin ilk belirteci olarak modele dahil edildiği incelemelerde elde edilen performans ölçülerinin arasındaki ilişkilerin sonuçları incelenmiştir.

Tablo 4.7.'de modele ilk alınan belirtecin X1 belirteci olduğu modellerden elde edilen performans sonuçları arasındaki ilişkiler gösterilmiştir. X1 belirteci, hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerde, bağımlı değişkenle en yüksek ilişkiye sahip olan iki belirteçten biridir. X1 belirteci ile başlanan modellerden elde edilen performans sonuçları arasındaki ilişkiler incelendiğinde; hem pozitif hem de negatif ilişki yapısında HUM ve CCP değerleri ile NRI ve IDI değerleri arasında pozitif yönde kuvvetli düzeyde ilişkiler elde edilmiştir. Ayrıca örneklem genişliği arttıkça bu ilişkilerin de yükseldiği görülmüştür.

Modele ilk alınan belirtecin X2 belirteci olduğu modellerden elde edilen performans sonuçları arasındaki ilişkiler Tablo 4.8.'de gösterilmiştir. X2 belirteci, hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerde, bağımlı değişkenle en yüksek ilişkiye sahip olan diğer belirteçtir. X2 belirteci ile başlanan modellerden elde edilen performans sonuçları arasındaki ilişkiler incelendiğinde; X1 belirteci ile başlanan modellerde elde edilen sonuçlara benzer sonuçlar elde edilmiştir. Burada dikkat çeken nokta ise; sadece örneklem genişliğinin düşük olduğu ve pozitif ilişki yapısında benzetim yapılan verilerden elde edilen sonuçların, X1 belirteci ile başlanan modellerde elde edilen sonuçlara göre daha düşük düzeyde ilişki göstermiş olmasıdır.

Pozitif ilişki yapısında benzetim yapılan verilerde bağımlı değişken ile orta düzeyde bir ilişkiye, negatif ilişki yapısında benzetim yapılan verilerde ise bağımlı değişken ile zayıf düzeyde bir ilişkiye sahip olan X3 belirtecinin, modelde bulunan ilk belirteç olduğu modellerin performans sonuçları arasındaki ilişkiler Tablo 4.9.'da gösterilmiştir. Pozitif ilişki yapısında elde edilen performans sonuçları arasındaki ilişkilerin negatif ilişki yapısında elde edilen performans sonuçları arasındaki ilişkilere göre daha zayıf düzeyde olduğu görülmüştür. NRI ve IDI yöntemleri modele eklenen belirtecin modele olan katkısını gösteren ölçüler olduğundan bu durumun beklenen bir sonuç olduğu düşünülmüştür. Negatif ilişki yapısında, X3 belirteci bağımlı değişken

ile düşük düzeyde bir ilişkiye sahip olduğundan, modele eklenen diğer belirteçlerin NRI ve IDI değerleri yüksek olacaktır. Bu nedenle, NRI ve IDI değerlerinin HUM ve CCP değerleri ile olan ilişkileri de yükselecektir. HUM ve CCP değerleri, modele eklenen belirtecin sırasından etkilenmediği için, elde edilen ilişkilerin yükselişi NRI ve IDI değerlerinden kaynaklanmaktadır. Bu durum daha net bir şekilde X4 ve X5 belirteçlerinin başlangıç belirteci olduğu modellerden elde edilen sonuçlarda da görülebilmektedir (Tablo 4.10. ve Tablo 4.11.). Hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerde, bağımlı değişkenle en düşük ilişkiye sahip olan belirteçler X4 ve X5 belirteçleridir. Her iki incelemede de, hem pozitif hem de negatif ilişki yapılarında elde edilen modellerin performans sonuçları arasındaki ilişkilerin genel olarak kuvvetli olduğu görülmektedir. Sadece örneklem genişliğinin küçük olduğu ve pozitif ilişki yapısında benzetim yapılan verilerden elde edilen sonuçlarda, X4 belirtecinin başlangıç belirteci olduğu modellerden elde edilen NRI değerleri ile HUM ve CCP değerleri arasındaki ilişkilerin orta düzeyde olduğu görülmüştür (Tablo 4.10.).

Tablo 4.7. Modelin ilk belirtecinin X1 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,754 (0,000)	0,710 (0,000)	0,854 (0,000)
	CCP	-	1	0,976 (0,000)	0,813 (0,000)
	NRI	-	-	1	0,792 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,907 (0,000)	0,926 (0,000)	0,915 (0,000)
	CCP	-	1	0,955 (0,000)	0,935 (0,000)
	NRI	-	-	1	0,926 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,966 (0,000)	0,966 (0,000)	0,971 (0,000)
	CCP	-	1	1,000 (0,000)	0,938 (0,000)
	NRI	-	-	1	0,939 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,442 (0,000)	0,413 (0,001)	0,790 (0,000)
	CCP	-	1	1,000 (0,000)	0,299 (0,016)
	NRI	-	-	1	0,299 (0,016)
	IDI	-	-	-	1
500	HUM	1	0,846 (0,000)	0,838 (0,000)	0,764 (0,000)
	CCP	-	1	1,000 (0,000)	0,936 (0,000)
	NRI	-	-	1	0,936 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,871 (0,000)	0,865 (0,000)	0,881 (0,000)
	CCP	-	1	1,000 (0,000)	0,973 (0,000)
	NRI	-	-	1	0,973 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Tablo 4.8. Modelin ilk belirtecinin X2 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,563 (0,000)	0,417 (0,001)	0,474 (0,000)
	CCP	-	1	0,505 (0,000)	0,729 (0,000)
	NRI	-	-	1	0,164 (0,195)
	IDI	-	-	-	1
500	HUM	1	0,950 (0,000)	0,948 (0,000)	0,971 (0,000)
	CCP	-	1	0,999 (0,000)	0,923 (0,000)
	NRI	-	-	1	0,922 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,938 (0,000)	0,935 (0,000)	0,961 (0,000)
	CCP	-	1	1,000 (0,000)	0,932 (0,000)
	NRI	-	-	1	0,932 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,423 (0,000)	0,399 (0,001)	0,780 (0,000)
	CCP	-	1	1,000 (0,000)	0,277 (0,027)
	NRI	-	-	1	0,277 (0,027)
	IDI	-	-	-	1
500	HUM	1	0,844 (0,000)	0,836 (0,000)	0,748 (0,000)
	CCP	-	1	1,000 (0,000)	0,922 (0,000)
	NRI	-	-	1	0,922 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,872 (0,000)	0,866 (0,000)	0,870 (0,000)
	CCP	-	1	1,000 (0,000)	0,962 (0,000)
	NRI	-	-	1	0,962 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Tablo 4.9. Modelin ilk belirtecinin X3 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,419 (0,001)	0,544 (0,000)	0,582 (0,000)
	CCP	-	1	0,802 (0,000)	0,647 (0,000)
	NRI	-	-	1	0,758 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,794 (0,000)	0,144 (0,258)	0,860 (0,000)
	CCP	-	1	0,570 (0,000)	0,857 (0,000)
	NRI	-	-	1	0,318 (0,010)
	IDI	-	-	-	1
1000	HUM	1	0,864 (0,000)	0,840 (0,000)	0,931 (0,000)
	CCP	-	1	0,997 (0,000)	0,872 (0,000)
	NRI	-	-	1	0,865 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,704 (0,000)	0,698 (0,000)	0,966 (0,000)
	CCP	-	1	0,996 (0,000)	0,681 (0,000)
	NRI	-	-	1	0,692 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,854 (0,000)	0,847 (0,000)	0,775 (0,000)
	CCP	-	1	1,000 (0,000)	0,976 (0,000)
	NRI	-	-	1	0,976 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,879 (0,000)	0,873 (0,000)	0,873 (0,000)
	CCP	-	1	1,000 (0,000)	0,967 (0,000)
	NRI	-	-	1	0,967 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Tablo 4.10. Modelin ilk belirtecinin X4 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,844 (0,000)	0,388 (0,002)	0,938 (0,000)
	CCP	-	1	0,410 (0,001)	0,814 (0,000)
	NRI	-	-	1	0,377 (0,002)
	IDI	-	-	-	1
500	HUM	1	0,969 (0,000)	0,953 (0,000)	0,949 (0,000)
	CCP	-	1	0,998 (0,000)	0,955 (0,000)
	NRI	-	-	1	0,953 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,962 (0,000)	0,960 (0,000)	0,972 (0,000)
	CCP	-	1	1,000 (0,000)	0,932 (0,000)
	NRI	-	-	1	0,932 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,804 (0,000)	0,799 (0,000)	0,868 (0,000)
	CCP	-	1	0,996 (0,000)	0,639 (0,000)
	NRI	-	-	1	0,647 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,897 (0,000)	0,892 (0,000)	0,889 (0,000)
	CCP	-	1	1,000 (0,000)	0,996 (0,000)
	NRI	-	-	1	0,996 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,897 (0,000)	0,892 (0,000)	0,874 (0,000)
	CCP	-	1	1,000 (0,000)	0,978 (0,000)
	NRI	-	-	1	0,978 (0,000)
	IDI	-	-	-	1

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Tablo 4.11. Modelin ilk belirtecinin X5 olduğu durumda elde edilen modellerin performans sonuçları arasındaki ilişkiler

n		Pozitif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,824 (0,000)	0,887 (0,000)	0,859 (0,000)
	CCP	-	1	0,727 (0,000)	0,744 (0,000)
	NRI	-	-	1	0,877 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,948 (0,000)	0,858 (0,000)	0,955 (0,000)
	CCP	-	1	0,778 (0,000)	0,937 (0,000)
	NRI	-	-	1	0,761 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,982 (0,000)	0,981 (0,000)	0,972 (0,000)
	CCP	-	1	1,000 (0,000)	0,953 (0,000)
	NRI	-	-	1	0,953 (0,000)
	IDI	-	-	-	1
n		Negatif İlişki Yapısında Benzetim Yapılan Veriler			
		HUM	CCP	NRI	IDI
100	HUM	1	0,787 (0,000)	0,780 (0,000)	0,805 (0,000)
	CCP	-	1	0,998 (0,000)	0,653 (0,000)
	NRI	-	-	1	0,658 (0,000)
	IDI	-	-	-	1
500	HUM	1	0,995 (0,000)	0,995 (0,000)	0,970 (0,000)
	CCP	-	1	1,000 (0,000)	0,966 (0,000)
	NRI	-	-	1	0,966 (0,000)
	IDI	-	-	-	1
1000	HUM	1	0,997 (0,000)	0,997 (0,000)	0,978 (0,000)
	CCP	-	1	1,000 (0,000)	0,975 (0,000)
	NRI	-	-	1	0,975 (0,000)
	IDI	-	-	-	1

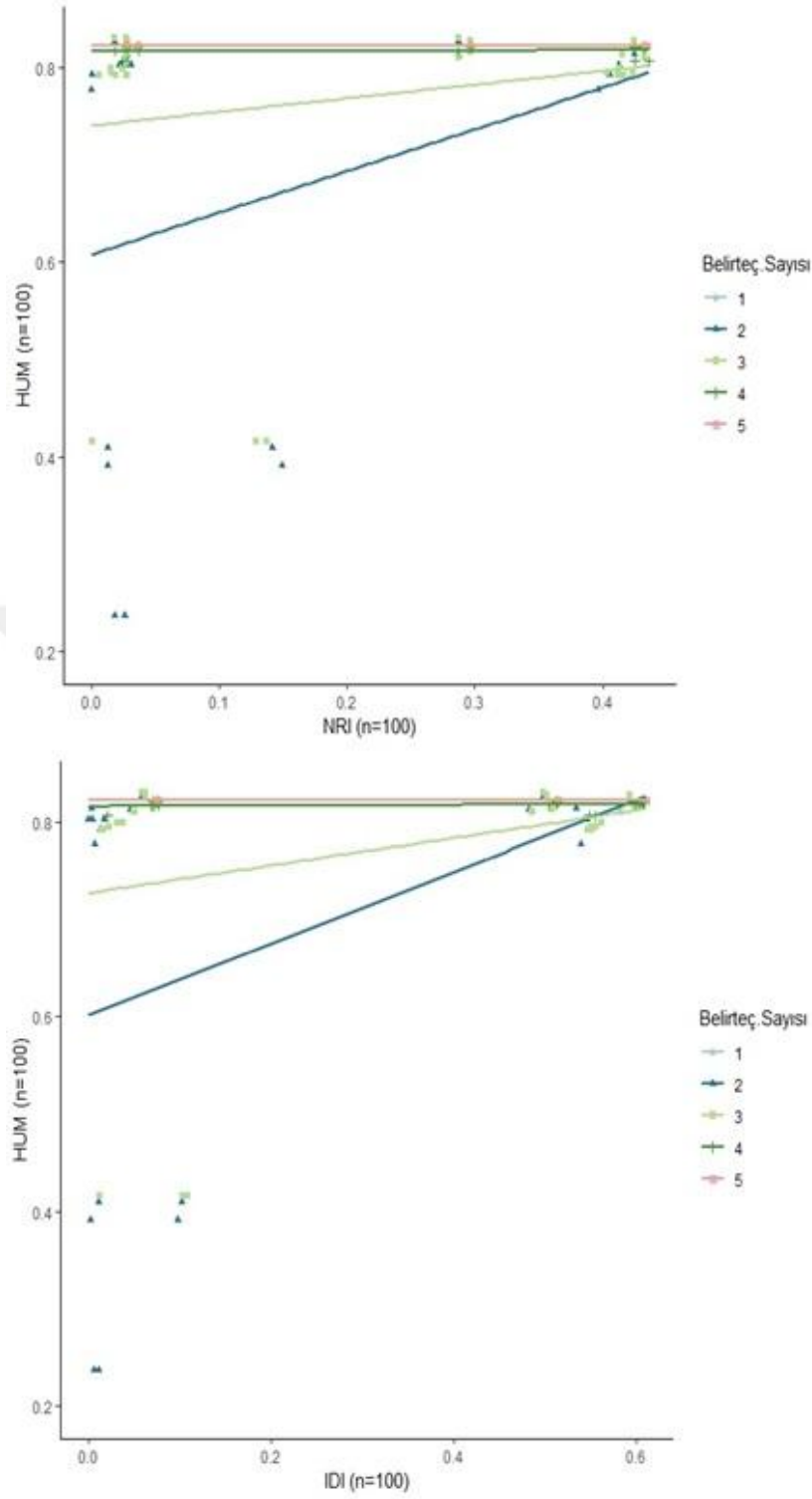
Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır. Gösterimler; korelasyon katsayısı - r (anlamlılık – p değeri) şeklinde ifade edilmiştir.

Genel olarak bakılacak olursa, hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerden elde edilen modellerin performans sonuçları arasındaki ilişkilerde, incelemiş olduğumuz klasik yöntemlerden birisi olan CCP değerlerinin HUM değerlerine göre NRI ve IDI ile daha yüksek ilişkilere sahip olduğu görülmüştür. Bunun en büyük sebebinin ise CCP yönteminin örneklem genişliği açısından HUM yöntemine göre daha hassas olmasından kaynaklandığı düşünülmüştür.

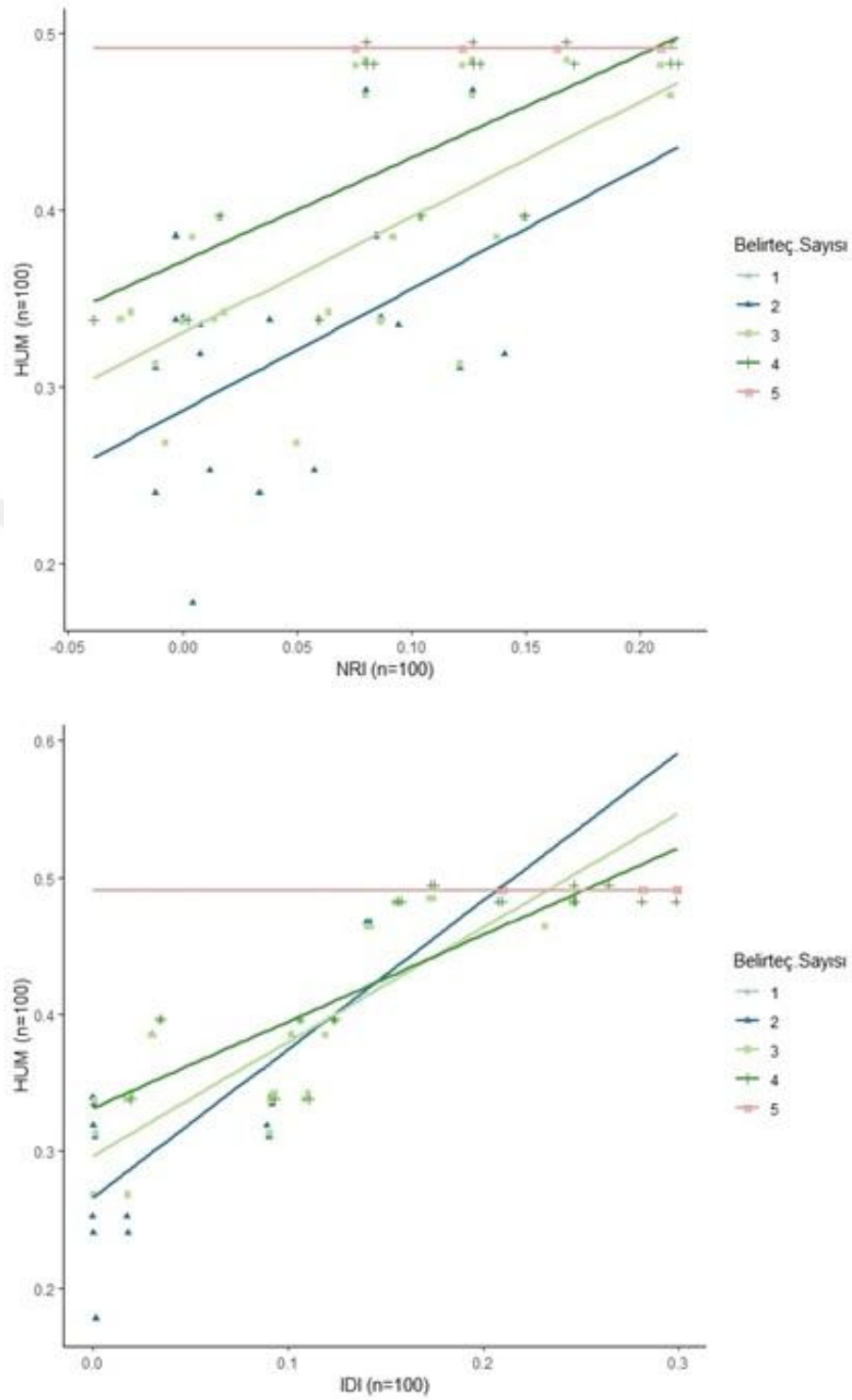
Negatif ilişki yapısında benzetim yapılan verilerden elde edilen modellerin NRI ve IDI değerlerinin klasik yöntemler olan HUM ve CCP değerleri ile daha yüksek ve tutarlı ilişkilere sahip olduğu görülmüştür.

Görsel olarak ilişkileri görebilmek için saçılım grafikleri ile de ilişkiler gösterilmiştir. Pozitif ilişki yapısında ve tüm örneklem genişliklerinde, modelde bulunan belirteç sayısı 2 ve 3 olduğunda incelenen performans sonuçları arasındaki ilişkilerin kuvvetli olduğu görülmektedir. Modelde bulunan belirteç sayısının 5 olduğu durumda ise HUM değerlerinin her modelde aynı sonucu vermesinden dolayı sadece NRI'nın değişimi görülebilmektedir. 4 belirteç ile kurulmuş olan modellerde ise sadece beş model kurulabildiği için, bu modellerin HUM değerlerinde de benzer sonuçlar görülmüştür (Şekil 4.1., Şekil 4.3., Şekil 4.5.). Negatif ilişki yapısında benzetim yapılan verilerle kurulan modellerden elde edilen sonuçlar incelendiğinde; NRI değerlerinin negatif ilişki yapısında olan verilerden etkilenmesi nedeniyle modelde bulunan belirteç sayısı 2, 3 ve 4 olduğunda ilişkilerin kuvvetli olduğu görülmektedir. 5 belirteçli modellerde ise pozitif sonuçlardaki durumun aynısı görülmüştür (Şekil 4.2., Şekil 4.4., Şekil 4.6.).

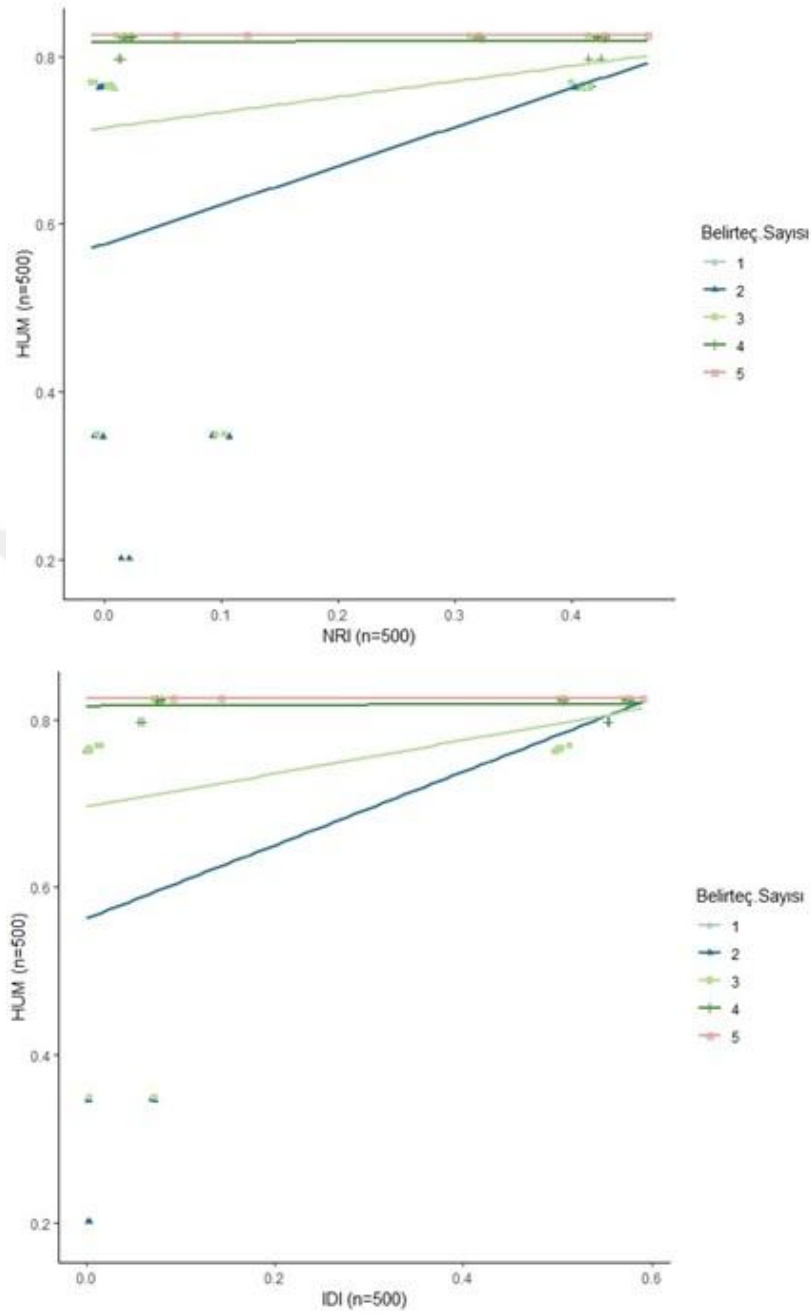
Baz değişkenlerin farklı olduğu durumlarda kurulan modellerin performans sonuçları arasındaki ilişkiler incelendiğinde ise, incelenen modellerin sonuçları aralarındaki ilişkilerin kuvvetli olduğu görülmektedir. Hem pozitif hem de negatif ilişki yapısında benzetim yapılan verilerle kurulan modellerin sonuçlarında, tüm örneklem genişliklerinde modele farklı belirteçler eklenmesi ile elde edilen HUM ve NRI değerlerinin ilişkili olduğu görülebilmektedir (Şekil 4.7., Şekil 4.8., Şekil 4.9., Şekil 4.10., Şekil 4.11., Şekil 4.12.).



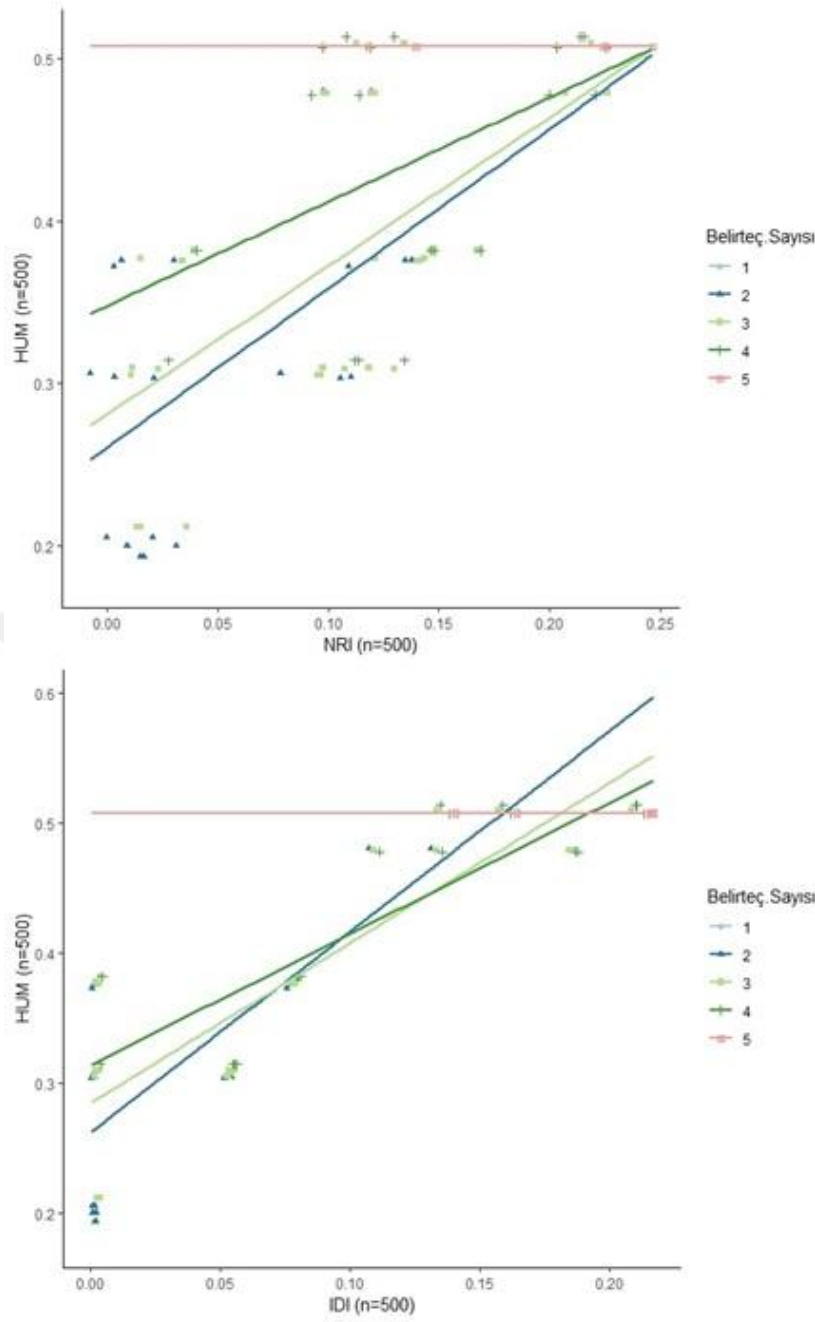
Şekil 4.1. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=100$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



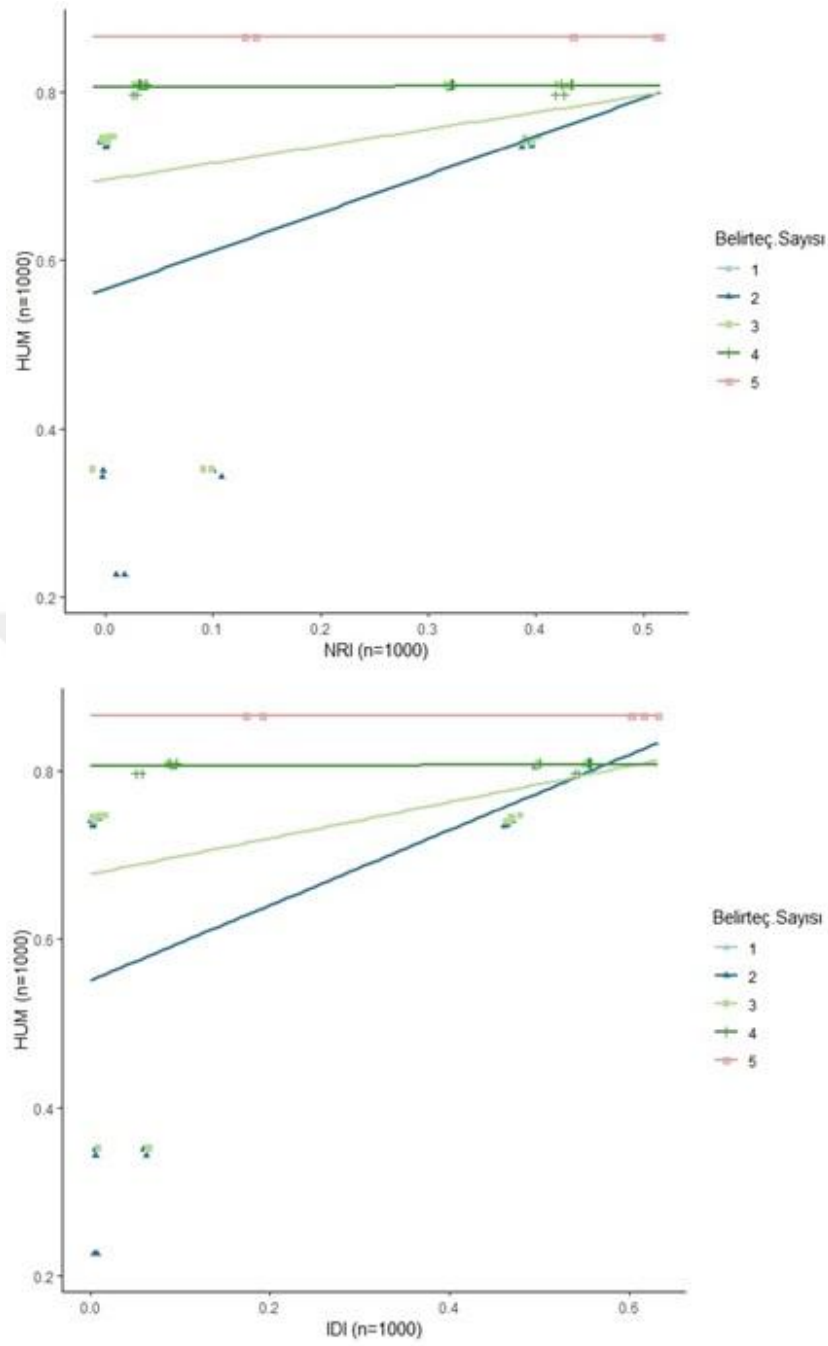
Şekil 4.2. Negatif ilişki yapısında benzetim yapılan verilerde, $n=100$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



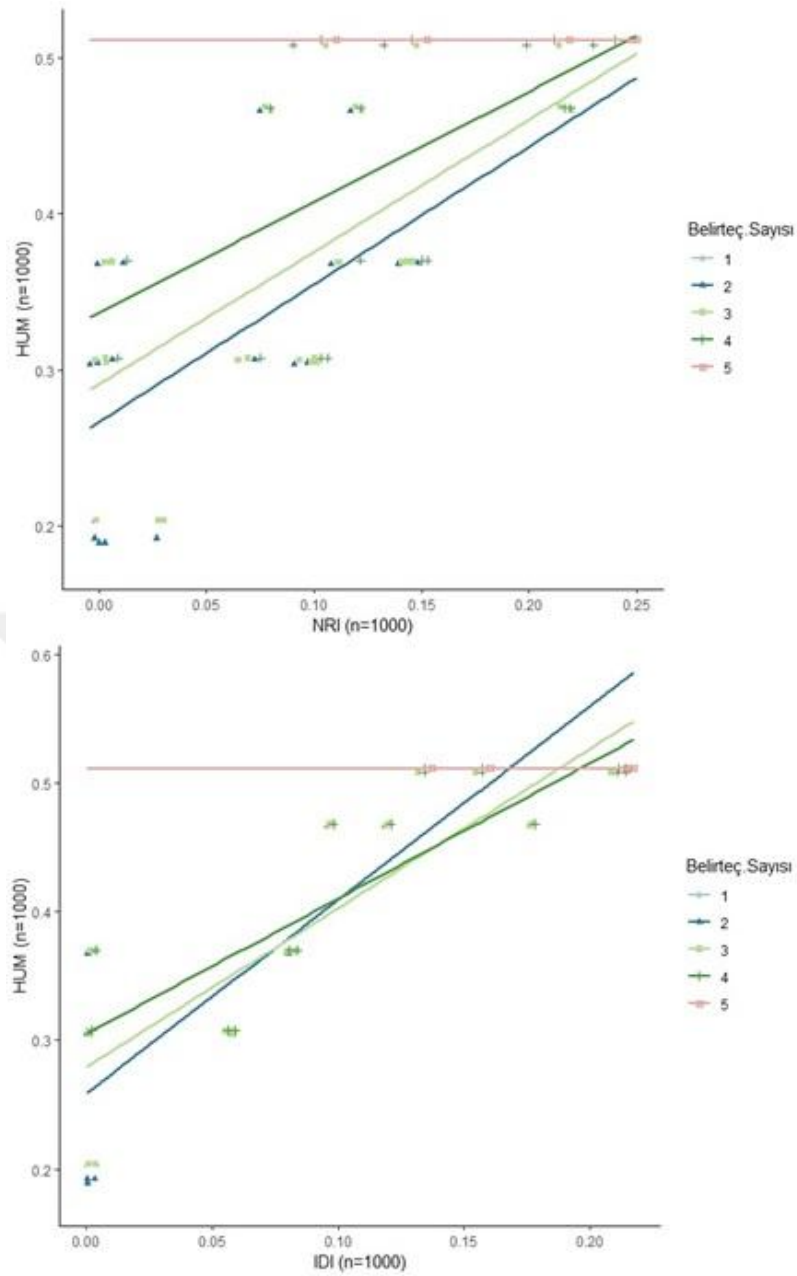
Şekil 4.3. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=500$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



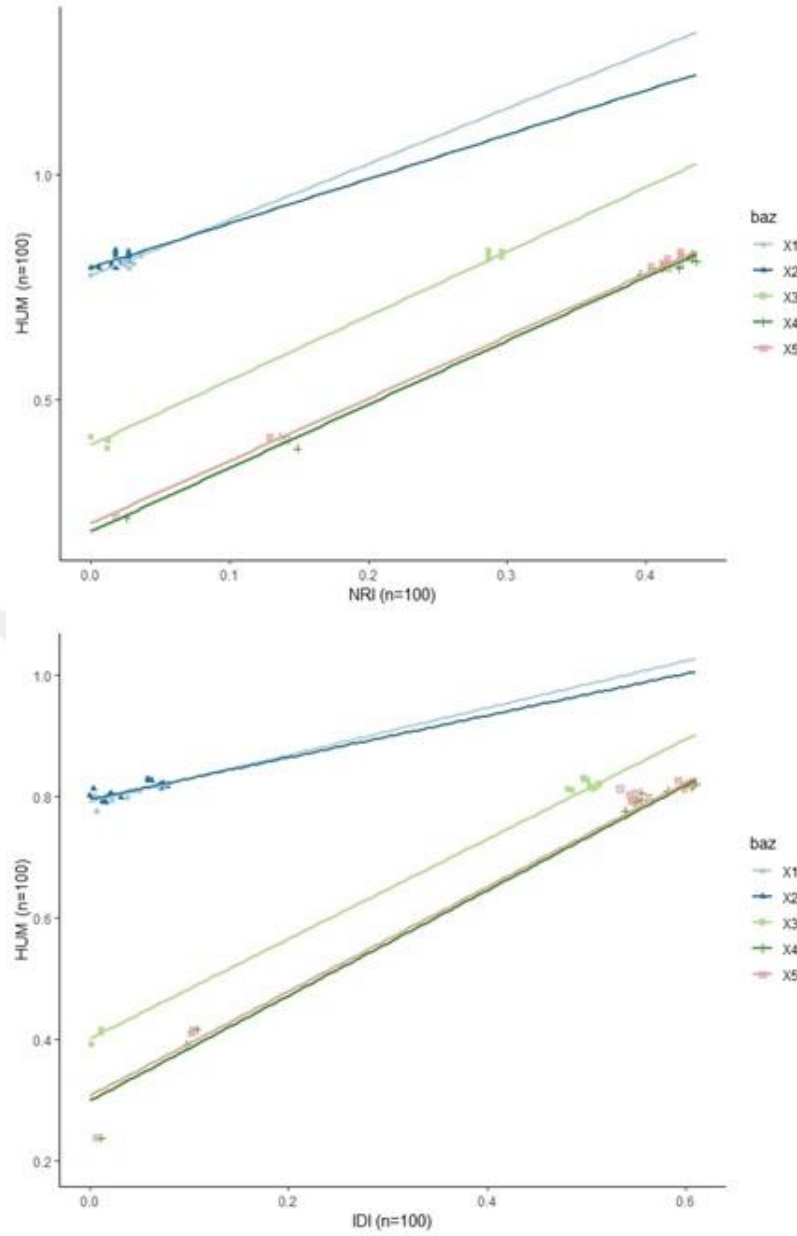
Şekil 4.4. Negatif ilişki yapısında benzetim yapılan verilerde, $n=500$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



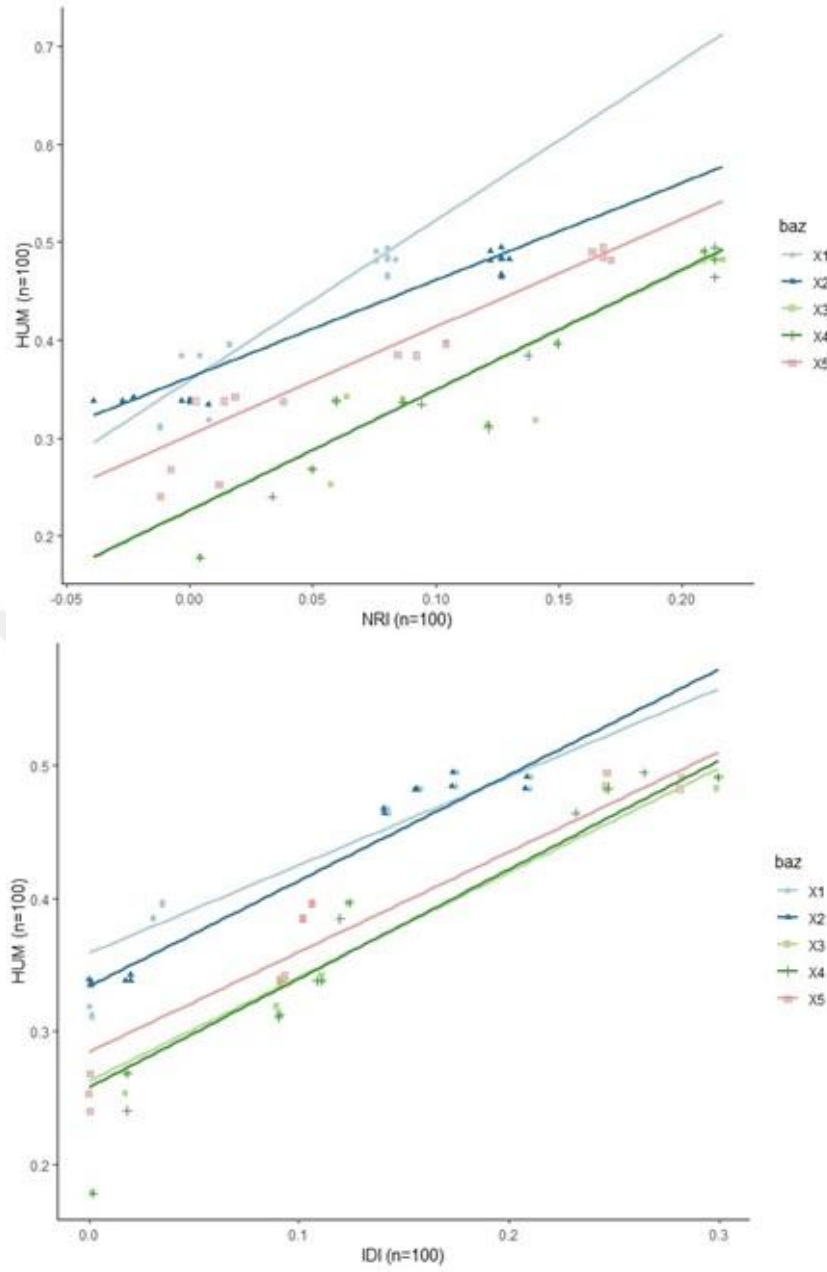
Şekil 4.5. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=1000$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



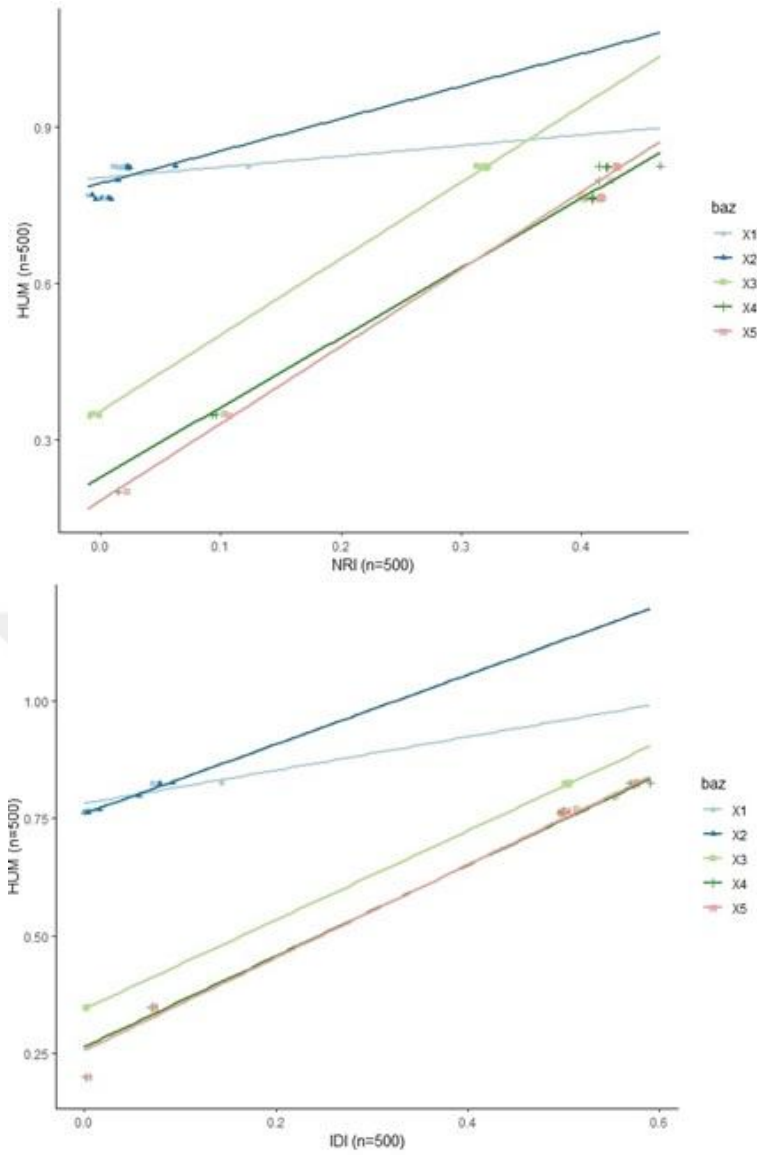
Şekil 4.6. Negatif ilişki yapısında benzetim yapılan verilerde, $n=1000$ örneklem için modelde bulunan belirteç sayısına göre modellerin performans sonuçları arasındaki ilişkiler



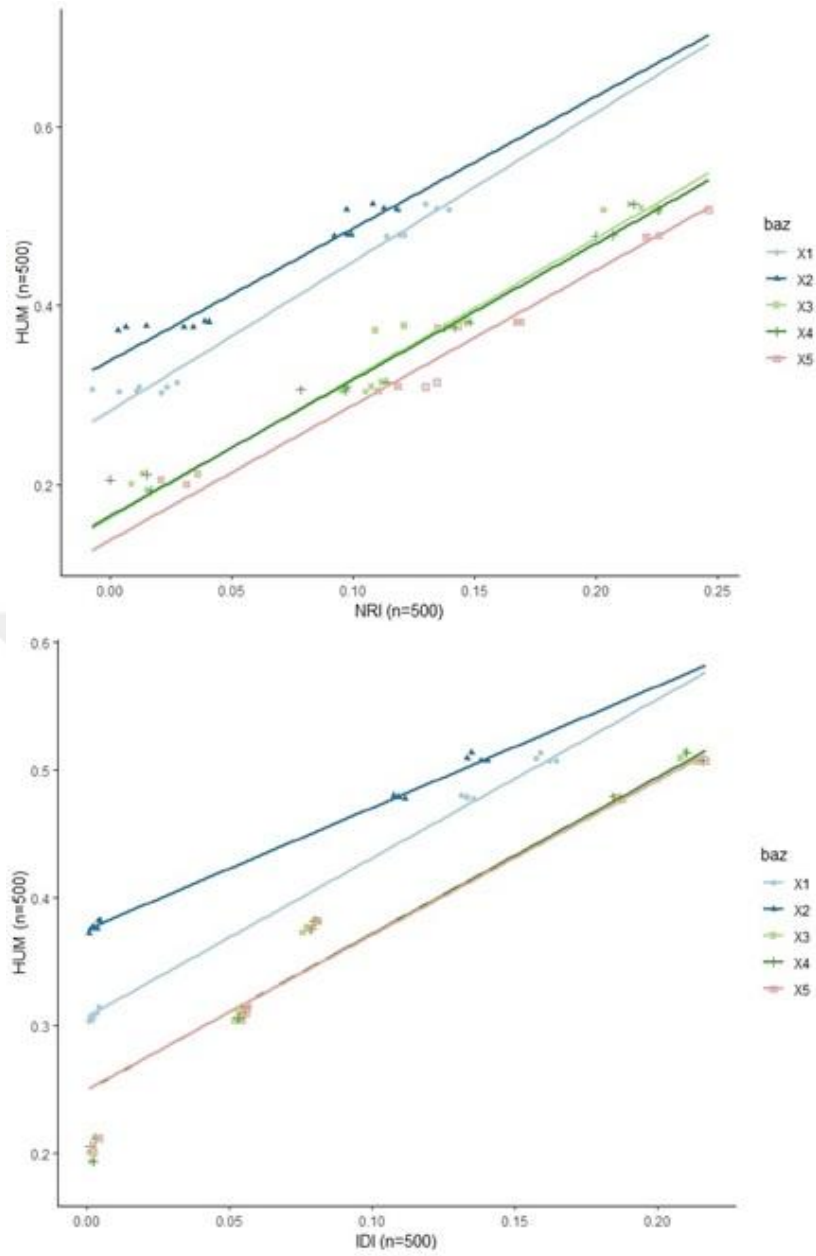
Şekil 4.7. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=100$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler



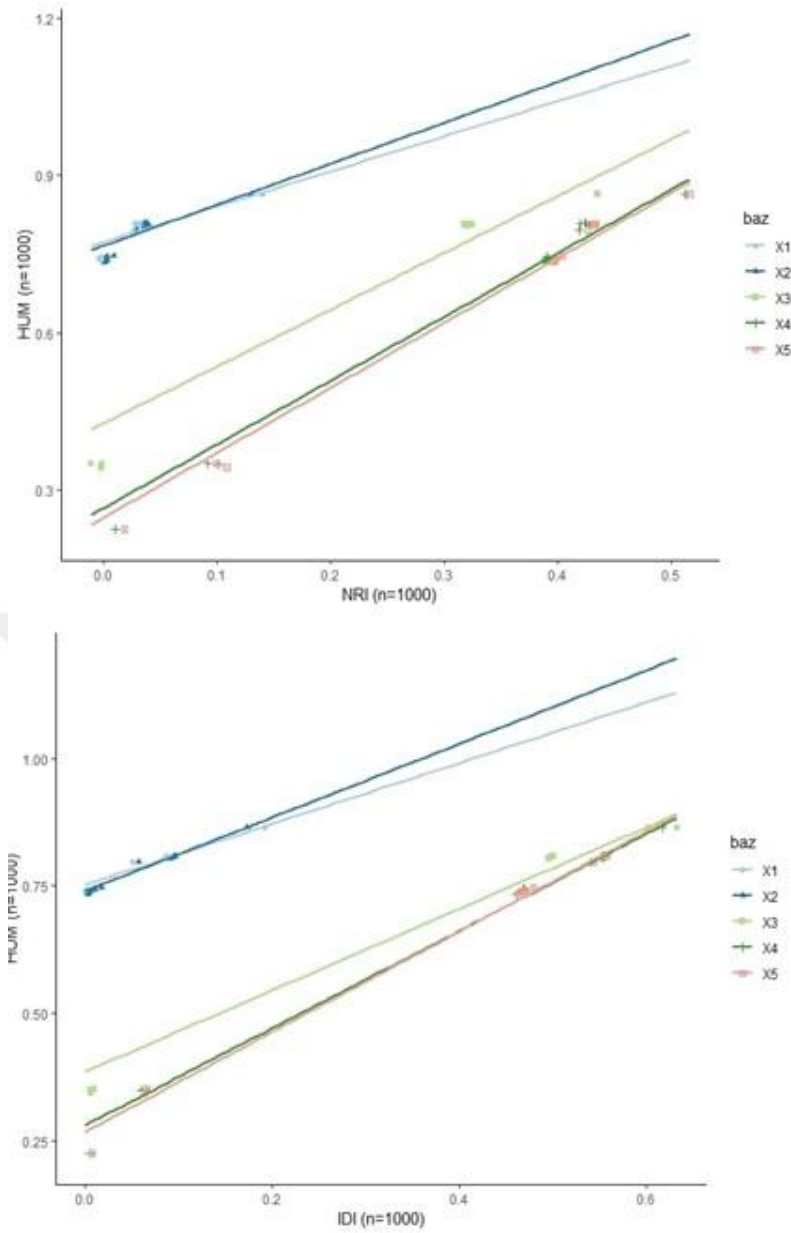
Şekil 4.8. Negatif ilişki yapısında benzetim yapılan verilerde, $n=100$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler



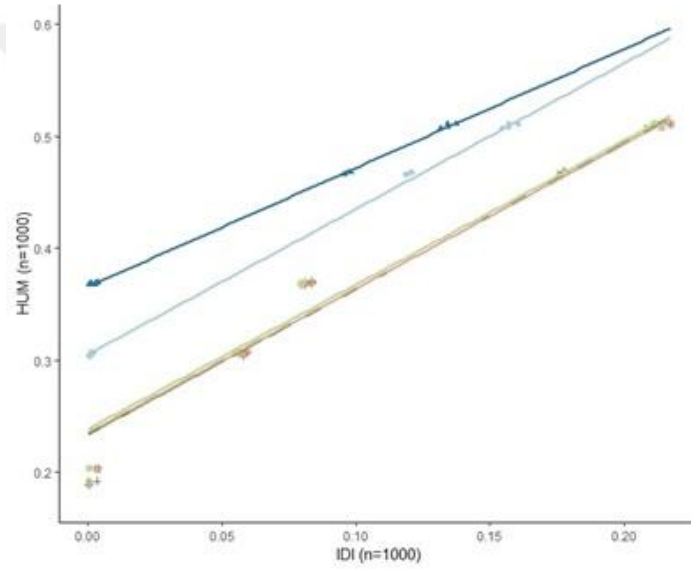
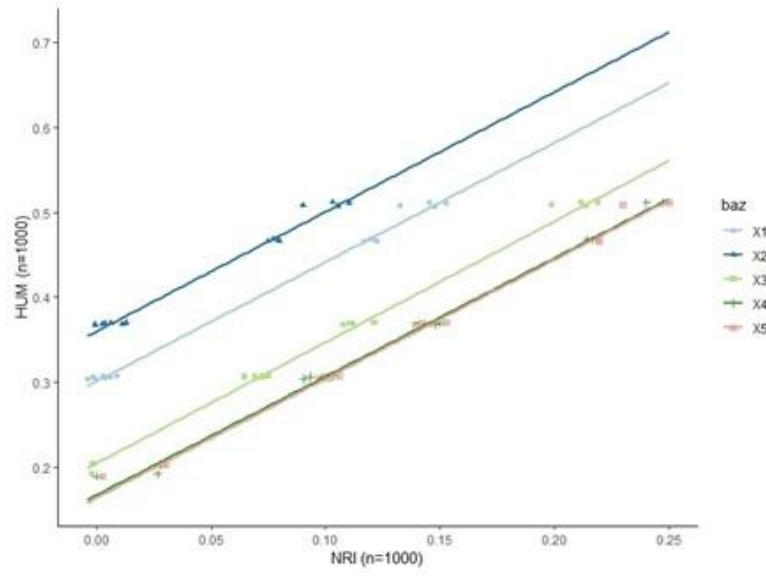
Şekil 4.9. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=500$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler



Şekil 4.10. Negatif ilişki yapısında benzetim yapılan verilerde, $n=500$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler



Şekil 4.11. Pozitif ilişki yapısında benzetim yapılan verilerde, $n=1000$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler



Şekil 4.12. Negatif ilişki yapısında benzetim yapılan verilerde, $n=1000$ örneklem için farklı belirteçlerin başlangıç modeli olarak incelendiği modellerin performans sonuçları arasındaki ilişkiler

4.2. Uygulama 2 - Bağımlı değişkenin sırasız yapıda olduğu durum için örnek uygulama (gerçek veri seti)

Bu çalışmada, orijinal veri seti üzerinden rastgele şekilde oluşturulan üç farklı veri seti ile çalışılmıştır. Bunun nedeni, incelenen performans ölçülerinin örneklem genişliği ile ilişkilerinin incelenmesidir.

Bu veri seti ile olan incelemelerde 325 adet HUM ve CCP ile hesaplanan model, 320 adet de NRI ve IDI ile hesaplanan model kullanılmıştır. Bu modeller, incelediğimiz belirteçler ile oluşturulabilecek olan model kombinasyonlarıdır. Her belirtecin modele ilk eklenen belirteç olduğu durum ve bu belirtecin arkasından modele eklenebilecek olan diğer belirteçlerin sırasının model performans sonuçlarına nasıl bir katkı yaptığı incelenmiştir.

İlk aşamada yapılan incelemelerde tek belirteç ile kurulan modeller kullanılmıştır. Tek belirteç ile kurulan modellerde NRI ve IDI değerleri hesaplanamadığı için performans sonuçları HUM ve CCP değerleri ile incelenmiştir. İncelenen tüm örneklem genişliklerinde, tiroid hastalığı sınıflaması için en yüksek performansı gösteren belirtecin TSH, en düşük sınıflama performansını gösteren belirtecin ise T4U olduğu görülmüştür (Tablo 4.12.).

Tablo 4.12. Tek belirteç ile kurulmuş olan modellerin performans sonuçları

Belirteç	n=100		n=500		n=1000	
	HUM	CCP	HUM	CCP	HUM	CCP
TSH	0,971	0,980	0,826	0,854	0,855	0,897
T3	0,430	0,590	0,513	0,620	0,523	0,615
TT4	0,656	0,700	0,612	0,666	0,632	0,670
T4U	0,212	0,570	0,285	0,546	0,262	0,547
FTI	0,776	0,740	0,713	0,712	0,685	0,712

İki belirteç ile kurulmuş olan modellerde tüm olası model kombinasyonları incelenmiştir. Sonuçlara bakıldığında; tek belirteç ile kurulmuş olan modellerde en yüksek sonucu veren TSH ve FTI belirteçleri ile kurulmuş olan model tüm örneklem genişliklerinde en yüksek sınıflama sonucunu vermiştir. HUM ve CCP değerlerinin, modelde bulunan belirtecin modele eklenme sırasından etkilenmedikleri görülmüştür.

TSH temel alınarak oluşturulan model ile FTI temel alınarak oluşturulan modellerin performans sonuçlarının aynı olduğu tablo 4.13.'de görülmektedir. Buna karşılık, NRI ve IDI değerlerinin modele eklenen yeni belirtecin modele olan katkısını incelemede daha hassas ölçüler olduğu bu tablodan görülebilir.

Daha ayrıntılı incelendiğinde; tek belirteç ile kurulmuş olan modellerde en yüksek sonucu veren TSH belirteci temel alınarak oluşturulan modellerde, NRI değerlerinin FTI belirtecine eklendiği durumda yüksek değerler aldığı görülmüştür. Ancak, tüm modellerin sonuçları göz önüne alındığında, NRI değerinin en yüksek değeri aldığı modelin “ $y = T4U + TSH$ ” olduğu görülmektedir. Bu model, tek belirteç ile kurulmuş olan modellerde en düşük sonucu veren T4U belirtecine temel alındığı ve bu belirtece TSH belirtecine eklendiği modeldir. Tek başına incelendiğinde zaten yüksek başarı ile sınıflama yapabilen bir belirtece herhangi bir belirtecin eklenmesinin model başarısını daha yüksek hale getirebileceği görülebildiği gibi, tek başına başarılı bir sınıflama yapamayan bir belirteç ile kurulan modellerde de hangi belirteçlerin model performansını yükselttiği görülmektedir. Bu durumda, modele eklenecek olan yeni ve modele en çok katkı sağlayabilecek olan belirtecin belirlenmesinde NRI yönteminin oldukça hassas ve yönlendirici olduğu görülebilmektedir.

Aynı HUM sonucunu veren diğer model olan “ $y = TSH + T4U$ ” modeline bakılacak olursa; TSH belirtecine T4U belirtecine eklenmesinin model performansında oldukça az bir katkı sağladığı benzer şekilde NRI değerleri ile görülebilmektedir.

Örneklem genişlikleri ile olan ilişkilerde ise; tüm örneklem genişliklerinde, incelenen modellerin sonuçları genel olarak aynı başarıyı göstermiştir. Sadece en küçük örneklem genişliğine sahip olan tek bir modelde IDI sonucunun beklenenin dışında bir değer aldığı görülmüştür. Bunun nedeninin ise IDI'nın formülasyonundan kaynaklandığı düşünülmektedir.

Tablo 4.13. İki belirteç ile kurulmuş olan modellerin performans sonuçları

Temel Alınan Belirteç	Kurulan Model	n=100				n=500				n=1000			
		HUM	CCP	IDI	NRI	HUM	CCP	IDI	NRI	HUM	CCP	IDI	NRI
TSH	y= TSH + T3	0,982	0,980	0,107	-0,010	0,830	0,862	0,034	0,008	0,884	0,914	0,027	0,015
TSH	y= TSH + TT4	0,986	0,880	0,088	0,000	0,912	0,888	0,093	0,040	0,954	0,934	0,099	0,041
TSH	y= TSH + T4U	0,984	0,960	0,084	-0,010	0,824	0,856	0,001	0,004	0,849	0,893	-0,002	0,001
TSH	y= TSH + FTI	0,988	0,990	0,026	0,010	0,968	0,896	0,124	0,044	0,989	0,944	0,119	0,047
T3	y= T3+ TSH	0,982	0,980	0,743	0,390	0,830	0,862	0,389	0,244	0,884	0,914	0,449	0,298
T3	y= T3 + TT4	0,653	0,690	0,300	0,100	0,621	0,654	0,188	0,036	0,650	0,670	0,213	0,057
T3	y= T3 + T4U	0,545	0,610	0,067	0,020	0,530	0,614	0,036	0,004	0,529	0,624	0,057	0,009
T3	y= T3 + FTI	0,771	0,740	0,344	0,150	0,735	0,714	0,254	0,094	0,726	0,725	0,277	0,108
TT4	y= TT4+ TSH	0,986	0,880	0,385	0,290	0,912	0,888	0,273	0,228	0,954	0,934	0,319	0,268
TT4	y= TT4 + T3	0,653	0,690	0,039	-0,010	0,621	0,654	0,009	-0,010	0,650	0,670	0,012	0,002
TT4	y= TT4 + T4U	0,790	0,760	0,105	0,050	0,709	0,704	0,058	0,038	0,682	0,714	0,049	0,045
TT4	y= TT4 + FTI	0,777	0,750	0,100	0,050	0,710	0,712	0,047	0,042	0,684	0,717	0,050	0,045
T4U	y= T4U + TSH	0,984	0,960	0,602	0,400	0,824	0,856	0,474	0,310	0,849	0,893	0,553	0,352
T4U	y= T4U + T3	0,545	0,610	0,085	0,030	0,530	0,614	0,177	0,078	0,529	0,624	0,180	0,078
T4U	y= T4U + TT4	0,790	0,760	0,437	0,180	0,709	0,704	0,363	0,158	0,682	0,714	0,382	0,167
T4U	y= T4U + FTI	0,781	0,750	0,430	0,180	0,707	0,712	0,364	0,166	0,684	0,713	0,383	0,166
FTI	y= FTI + TSH	0,988	0,990	0,346	0,250	0,968	0,896	0,237	0,184	0,989	0,944	0,293	0,232
FTI	y= FTI + T3	0,771	0,740	-0,002	0,000	0,735	0,714	0,024	0,000	0,726	0,725	0,026	0,008
FTI	y= FTI + TT4	0,777	0,750	0,012	0,010	0,710	0,712	0,003	-0,002	0,684	0,717	0,001	0,004
FTI	y= FTI + T4U	0,781	0,750	0,014	0,010	0,707	0,712	0,005	0,000	0,684	0,713	0,001	0,001

Üç belirteç ile kurulmuş olan modellerden elde edilen performans sonuçları incelendiğinde; tüm örneklem genişliklerinde, tek belirteç ile kurulmuş olan modellerde en yüksek performans sonucunu veren TSH belirteci ile kurulmuş olan model sonuçlarının en başarılı sınıflama sonuçlarını verdiği görülmüştür. TSH belirtecinin modele eklendiği sıranın önemi olmaksızın tüm incelemelerde en başarılı sonuçlar bu belirteçle kurulan modellerde görülmüştür (Tablo 4.14.).

Üç belirteç ile kurulmuş olan tüm modellerde en yüksek sınıflama performanslarının TSH, TT4 ve T4U belirteçleri ile kurulmuş olan modellerde olduğu görülmektedir. Sıra etkisini de göz önüne alarak inceleme yaptığımızda; modele temel olarak T4U belirteci alınarak kurulan modellerde NRI değerinin daha yüksek olduğu görülmüştür. T4U belirteci ile modelleme süreci başlatıldığında modele TSH ve TT4 belirteçlerinin eklenmesi durumunda diğer belirteçlerle kurulan modellerin performanslarına oranla daha yüksek bir başarı elde edilebileceği görülmektedir. Bu sonuç beklenen bir sonuçtur.

Tekrar iki belirteç ile kurulan modellerin sonuçlarına dönüp, bu sonuçları üç belirteç ile kurulan model sonuçları ile bağlayacak olursak; tek başına sınıflama performansına etkisi yüksek olan bir belirteç ile kurulan modellerde, modele yeni eklenen belirtecine de tek başına etkisinin oldukça yüksek olan bir belirteç olması nedeniyle, modele eklenecek olan üçüncü belirtecine de yüksek bir performans göstermesi gerektiği görülmüştür. Ancak tek başına sınıflama performansına etkisi düşük olan bir belirteç ile modelleme yapılması durumunda, modele eklenecek olan belirteçlerden hangisinin modele daha çok katkı sağlayacağı ve en az belirteç ile en yüksek performans sonucuna nasıl ulaşılacağı daha net görülebilmektedir. Üç belirteç ile kurulan modellerde aynı anda hem HUM hem de NRI değerinin en yüksek olduğu modelin “ $y = T4U + TSH + TT4$ ” modeli olduğu görülmüştür. Bu durum, her zaman modele en yüksek katkısı olan belirteç ile modellemeye başlanmasının kesinlikle gerekli olmadığını, modele katkısı düşük olan bir belirteç ile de modelleme sürecine başlanabileceğini göstermiştir. Bu durumda, elde edilen sınıflama performansları aynı olacaktır ama katkı sağlayan belirteç daha net bir şekilde ortaya çıkarılabilecektir.

Tablo 4.14. Üç belirteç ile kurulmuş olan modellerin performans sonuçları

	Temel Alınan Belirteç	Kurulan Model	HUM	CCP	IDI	NRI
n=100	TSH	y= TSH + (T3+ T4U)	0,984	0,970	0,108	0,000
	TSH	y= TSH + (TT4+ T4U)	0,988	0,960	-0,019	-0,010
	T3	y= T3+ (TSH + FTI)	0,985	0,970	0,693	0,400
	T3	y= T3 + (TT4+ FTI)	0,791	0,750	0,361	0,170
	TT4	y= TT4+ (T4U + FTI)	0,786	0,750	0,110	0,050
	TT4	y= TT4+ (TSH + T4U)	0,988	0,960	0,368	0,270
	T4U	y= T4U + (TT4+ FTI)	0,786	0,750	0,451	0,190
	T4U	y= T4U + (TSH + TT4)	0,988	0,960	0,656	0,400
	FTI	y= FTI + (TT4+ T4U)	0,786	0,750	0,024	0,020
	FTI	y= FTI + (TSH + TT4)	0,987	0,970	0,352	0,230
n=500			HUM	CCP	IDI	NRI
	TSH	y= TSH + (T3+ T4U)	0,853	0,874	0,033	0,026
	TSH	y= TSH + (TT4+ T4U)	0,968	0,892	0,123	0,044
	T3	y= T3+ (TSH + FTI)	0,964	0,898	0,481	0,276
	T3	y= T3 + (TT4+ FTI)	0,742	0,708	0,249	0,094
	TT4	y= TT4+ (T4U + FTI)	0,708	0,712	0,062	0,046
	TT4	y= TT4+ (TSH + T4U)	0,968	0,892	0,278	0,232
	T4U	y= T4U + (TT4+ FTI)	0,708	0,712	0,365	0,160
	T4U	y= T4U + (TSH + TT4)	0,968	0,892	0,600	0,348
	FTI	y= FTI + (TT4+ T4U)	0,708	0,712	0,006	-0,008
n=1000			HUM	CCP	IDI	NRI
	TSH	y= TSH + (T3+ T4U)	0,891	0,914	0,027	0,016
	TSH	y= TSH + (TT4+ T4U)	0,989	0,950	0,117	0,051
	T3	y= T3+ (TSH + FTI)	0,989	0,951	0,546	0,336
	T3	y= T3 + (TT4+ FTI)	0,732	0,713	0,276	0,102
	TT4	y= TT4+ (T4U + FTI)	0,684	0,715	0,050	0,046
	TT4	y= TT4+ (TSH + T4U)	0,989	0,950	0,318	0,270
	T4U	y= T4U + (TT4+ FTI)	0,684	0,715	0,384	0,168
	T4U	y= T4U + (TSH + TT4)	0,989	0,950	0,669	0,398
	FTI	y= FTI + (TT4+ T4U)	0,684	0,715	0,002	0,003
	FTI	y= FTI + (TSH + TT4)	0,988	0,943	0,280	0,230

Bu durum için daha net bir gösterge olarak, dört belirteç ile kurulmuş olan modeller incelendiğinde; en yüksek HUM değerini alan modelin TSH, T3, T4U ve FTI belirteçleri ile kurulan model olduğu görülmüştür. HUM değerlerinde sıra etkisi gözlenemediği için bu belirteçlerle kurulan tüm farklı model sonuçları aynıdır. Ancak NRI değerleri de incelendiğinde modele T4U belirteci ile başlamanın avantajlı olduğu görülmektedir (Tablo 4.15.).

Burada dikkat edilmesi gereken bir diğer nokta ise iki, üç ve dört belirteç ile kurulan model sonuçlarının HUM değerleridir. Tek belirteç ile kurulan modellerde en yüksek HUM değeri TSH belirteci ile kurulan modelden elde edilen 0.855'dir. İki belirteç ile kurulan modellerde en yüksek HUM değeri TSH ve FTI belirteçleri ile kurulan modelden elde edilen 0.944'tür. Üç belirteç ile kurulan modellerde en yüksek HUM değeri TSH, TT4 ve T4U belirteçleri ile kurulan modelden elde edilen 0.989'dur. Dört belirteç ile kurulan modellere gelindiğinde ise en yüksek elde edilen HUM değerinin 0.988 olduğu görülmektedir. Bu durumda dördüncü bir belirteci modele eklemenin model performansına getirisinin olmadığı görülmektedir. Son aşama olarak da beş belirteçle kurulan modelden edilen HUM değerinin de 0.989 olduğu görülmektedir (Tablo 4.16.). Modele iki belirteç daha eklenmiş olmasına rağmen model performansında bir yükselme olmadığı görülmektedir. Bu durumda, NRI değerleri dikkate alınarak en yüksek HUM değerine en kısa sürede ulaşılabileceği görülmektedir. NRI değerleri, en başarılı sınıflama modelinin belirlenebilmesi sürecinde, belirteç seçiminde oldukça önemli bir göstergedir. Ayrıca, örneklem büyüklüğünün artmasının NRI değerleri üzerinde bir etki göstermediği de görülmüştür. Örneklem genişliğinin küçük olduğu incelemelerde de belirteç seçimi için kullanılabilecek bir yöntem olabileceği görülmüştür.

Tablo 4.15. Dört belirteç ile kurulmuş olan modellerin performans sonuçları

	Temel Alınan Belirteç	Kurulan Model	HUM	CCP	IDI	NRI
n=100	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{FTI})$	0,979	0,960	0,075	-0,010
	TSH	$y = \text{TSH} + (\text{T3} + \text{T4U} + \text{FTI})$	0,987	0,970	0,104	0,000
	T3	$y = \text{T3} + (\text{TT4} + \text{T4U} + \text{FTI})$	0,797	0,770	0,372	0,170
	T3	$y = \text{T3} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,987	0,970	0,697	0,380
	TT4	$y = \text{TT4} + (\text{T3} + \text{T4U} + \text{FTI})$	0,797	0,770	0,115	0,060
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,985	0,960	0,445	0,270
	T4U	$y = \text{T4U} + (\text{T3} + \text{TT4} + \text{FTI})$	0,797	0,770	0,452	0,200
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{FTI})$	0,987	0,970	0,781	0,410
	FTI	$y = \text{FTI} + (\text{T3} + \text{TT4} + \text{T4U})$	0,797	0,770	0,025	0,020
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{T4U})$	0,987	0,970	0,323	0,230
			HUM	CCP	IDI	NRI
	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{FTI})$	0,965	0,890	0,137	0,050
	TSH	$y = \text{TSH} + (\text{T3} + \text{T4U} + \text{FTI})$	0,965	0,892	0,133	0,042
n=500	T3	$y = \text{T3} + (\text{TT4} + \text{T4U} + \text{FTI})$	0,747	0,716	0,259	0,102
	T3	$y = \text{T3} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,965	0,892	0,499	0,284
	TT4	$y = \text{TT4} + (\text{T3} + \text{T4U} + \text{FTI})$	0,747	0,716	0,084	0,056
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,968	0,896	0,291	0,230
	T4U	$y = \text{T4U} + (\text{T3} + \text{TT4} + \text{FTI})$	0,747	0,716	0,385	0,174
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{FTI})$	0,965	0,892	0,621	0,354
	FTI	$y = \text{FTI} + (\text{T3} + \text{TT4} + \text{T4U})$	0,747	0,716	0,032	0,008
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{T4U})$	0,965	0,892	0,237	0,186
			HUM	CCP	IDI	NRI
	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{FTI})$	0,988	0,946	0,123	0,055
	TSH	$y = \text{TSH} + (\text{T3} + \text{T4U} + \text{FTI})$	0,988	0,954	0,123	0,054
	T3	$y = \text{T3} + (\text{TT4} + \text{T4U} + \text{FTI})$	0,730	0,718	0,281	0,103
	T3	$y = \text{T3} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,988	0,954	0,550	0,337
n=1000	TT4	$y = \text{TT4} + (\text{T3} + \text{T4U} + \text{FTI})$	0,730	0,718	0,079	0,048
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T4U} + \text{FTI})$	0,988	0,943	0,325	0,274
	T4U	$y = \text{T4U} + (\text{T3} + \text{TT4} + \text{FTI})$	0,730	0,718	0,408	0,171
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{FTI})$	0,988	0,954	0,680	0,404
	FTI	$y = \text{FTI} + (\text{T3} + \text{TT4} + \text{T4U})$	0,730	0,718	0,034	0,006
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{T4U})$	0,988	0,954	0,298	0,239

Tablo 4.16. Beş belirteç ile kurulmuş olan tam modellerin performans sonuçları

	Temel Alınan Belirteç	Kurulan Model	HUM	CCP	IDI	NRI
n=100	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{T4U} + \text{FTI})$	0,975	0,940	0,124	-0,010
	T3	$y = \text{T3} + (\text{TSH} + \text{TT4} + \text{T4U} + \text{FTI})$	0,975	0,940	0,745	0,390
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T3} + \text{T4U} + \text{FTI})$	0,975	0,940	0,500	0,280
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{TT4} + \text{FTI})$	0,975	0,940	0,851	0,410
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{TT4} + \text{T4U})$	0,975	0,940	0,392	0,240
			HUM	CCP	IDI	NRI
n=500	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{T4U} + \text{FTI})$	0,966	0,906	0,136	0,046
	T3	$y = \text{T3} + (\text{TSH} + \text{TT4} + \text{T4U} + \text{FTI})$	0,966	0,906	0,490	0,280
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T3} + \text{T4U} + \text{FTI})$	0,966	0,906	0,319	0,234
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{TT4} + \text{FTI})$	0,966	0,906	0,619	0,352
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{TT4} + \text{T4U})$	0,966	0,906	0,256	0,186
			HUM	CCP	IDI	NRI
n=1000	TSH	$y = \text{TSH} + (\text{T3} + \text{TT4} + \text{T4U} + \text{FTI})$	0,989	0,951	0,123	0,055
	T3	$y = \text{T3} + (\text{TSH} + \text{TT4} + \text{T4U} + \text{FTI})$	0,989	0,951	0,549	0,338
	TT4	$y = \text{TT4} + (\text{TSH} + \text{T3} + \text{T4U} + \text{FTI})$	0,989	0,951	0,346	0,285
	T4U	$y = \text{T4U} + (\text{TSH} + \text{T3} + \text{TT4} + \text{FTI})$	0,989	0,951	0,683	0,406
	FTI	$y = \text{FTI} + (\text{TSH} + \text{T3} + \text{TT4} + \text{T4U})$	0,989	0,951	0,296	0,242

İncelenen performans ölçülerinin kendi aralarındaki ve modelde bulunan belirteç sayıları ile arasındaki ilişkileri incelenmiştir. HUM ve CCP değerleri arasındaki ilişkinin tüm örneklem genişliklerinde anlamlı olduğu görülmüştür. Burada göze çarpan nokta ise örneklem genişliği arttıkça HUM ve CCP değerleri arasındaki ilişki katsayısının düşüş göstermesidir (Tablo 4.17.).

NRI ve IDI değerleri arasındaki ilişkiler incelendiğinde; tüm örneklem genişliklerinde oldukça kuvvetli katsayılar elde edildiği görülmüştür. NRI değerleri ve IDI değerleri bazı modellerde küçük farklılıklar göstermiş olsa da genel olarak bakıldığında oldukça kuvvetli bir ilişkiye sahip oldukları görülmektedir.

HUM ve NRI değerleri arasındaki ilişkilere bakılacak olursa; tüm örneklem genişliklerinde anlamlı ve pozitif yönde bir ilişki görülmektedir. Bu ilişkinin gücünün örneklem genişliği arttıkça daha da kuvvetlendiği görülmektedir. Tanı tahmin modeli incelemelerinde, örneklem genişliğinin yüksek olduğu durumlarda, tek başına sınıflama başarısı yüksek olan belirteçlerin tanı modeline katkısının da yüksek olacağı söylenebilir.

Modelde bulunan belirteç sayıları ile incelemiş olduğumuz performans ölçülerinin ilişkilerine bakıldığında; küçük örneklem genişliğinde sınıflama performansının modeldeki belirteç sayısı ile negatif yönde ilişkili olduğu görülmüştür. Örneklem genişliği yükseldikçe bu ilişkinin yönü pozitif olmuştur. Küçük örneklem genişliğinde negatif ilişki elde edilmiş olmasının ise bağımlı değişkenin sırasız olması durumundan kaynaklandığı düşünülmektedir. NRI ve IDI değerleri ile modeldeki belirteç sayıları arasında ise tüm örneklem genişliklerinde pozitif yönde ilişkiler olduğu görülmüştür.

Ayrıca, en iyi sınıflama performansına ulaşabilmek için klasik istatistiksel yöntemlerden birisi olan Kuadratik Diskriminant Analizi ile de inceleme yapılmıştır. İlk olarak verilerin dağılımı incelenmiştir. Değişkenlerin normal dağılım göstermemesi nedeniyle verilere “Box Cox dönüşümleri” yapılarak incelenmiştir.

Box Cox dönüşümü yapılan model sonuçları incelendiğinde; modele giren ilk belirtecin tüm örneklem genişliklerinde TSH olduğu görülmektedir. Örneklem genişliğinin düşük olduğu modelde, ikinci belirteç olarak FTI belirteci görülmektedir. Örneklem genişliğinin arttığı modellerde ise ikinci belirteç olarak FTI ve üçüncü belirteç olarak T3 belirtecinin modelde yer aldığı görülmüştür (Tablo 4.18.).

İncelediğimiz performans ölçüleri ile Kuadratik Diskriminant analizi sonuçları karşılaştırılmıştır. Her iki yöntemden elde edilen sonuçlarda da en iyi performans sonucu üç belirteç ile kurulan modeller ile elde edilmektedir. Kuadratik Diskriminant analizinde, modelde TSH, FTI ve T3 belirteçleri sıra ile modele girerken; yeni geliştirilen yöntemlerle elde edilen sonuçlarda da benzer şekilde TSH ilk olarak modele girdiğinde en iyi sınıflama performansını elde etmiş, arkasından FTI belirteci modele girerek en iyi sınıflama performansını vermiştir.

Tablo 4.17. Elde edilen performans sonuçlarının modelde bulunan belirteç sayısı ile ve kendi aralarındaki ilişkileri

	HUM		CCP		IDI		NRI	
n=100	r	p	r	p	r	p	r	p
M.B.S.	-0,152	0,010	-0,141	0,000	0,401	0,000	0,276	0,000
HUM			0,922	0,000	0,111	0,050	0,184	0,000
CCP					0,141	0,010	0,198	0,000
IDI							0,941	0,000
n=500	r	p	r	p	r	p	r	p
M.B.S.	0,542	0,000	0,861	0,000	0,343	0,000	0,356	0,000
HUM			0,761	0,000	0,244	0,000	0,309	0,000
CCP					0,360	0,000	0,390	0,000
IDI							0,946	0,000
n=1000	r	p	r	p	r	p	r	p
M.B.S.	0,699	0,000	0,587	0,000	0,333	0,000	0,403	0,000
HUM			0,610	0,000	0,328	0,000	0,427	0,000
CCP					0,420	0,000	0,488	0,000
IDI							0,965	0,000

Spearman Korelasyon Katsayısı; M.B.S. :Modelde bulunan belirteç sayısı

Tablo 4.18. Tiroid Verisi ile uygulanan Kuadratik Diskriminant Analizi Sonuçları

n		Wilks' Lambda	Kesin F	Doğru Sınıflama Olasılığı
100	TSH	0,293	116,934	0,980
	FTI	0,233	51,394	
500	TSH	0,357	447,325	0,934
	FTI	0,271	228,049	
	T3	0,250	164,848	
1000	TSH	0,331	1007,793	0,946
	FTI	0,257	484,265	
	T3	0,245	338,883	

4.3. Uygulama 3 - Mikrodizin verisi için örnek uygulama

Tek gen ekspresyon değerleri ile kurulan modellerde elde edilen performans sonuçları Tablo 4.19.'da gösterilmiştir. 7129 gen ekspresyon değerleri ayrı ayrı incelenerek bu tablo oluşturulmuştur.

Tablo 4.19. Tek gen ekspresyon değerlerinden elde edilen modellerin performans sonuçları

Veri Seti 1			Veri Seti 2		
En yüksek Sınıflama Performansına Sahip 5 Gen			Farklı Seviyelerden Sınıflama Performansına Sahip 5 Gen		
Gen Sırası	HUM	CCP	Gen Sırası	HUM	CCP
G1882 (x1)	0,812	0,842	G1882 (x1)	0,812	0,842
G248 (x2)	0,807	0,658	G248 (x2)	0,807	0,658
G5300 (x3)	0,803	0,789	G2794 (x3)	0,679	0,711
G1207 (x4)	0,791	0,737	G5153 (x4)	0,519	0,526
G4342 (x5)	0,774	0,737	G7006 (x5)	0,136	0,474

Tablo 4.20.'de, 2 gen ile kurulan modeller incelendiğinde; hem veri seti 1 hem de veri seti 2'de kurulan modellerin %100 doğru sınıflama yapabildiği görülmüştür. Veri seti 1'de; tek genle kurulan modellerde (Tablo 4.19.) en düşük başarıyı gösteren genin (x5), «2 gen ile kurulan modellerde» modele en yüksek katkıyı yaparak sınıflama sonucunu %100 değerine çıkaran gen olduğu görülmektedir. Bunun nedeni ise, genlerin kendi aralarında bulunan ilişkiler olabileceği düşünülmüştür.

Tablo 4.21.'den görülebileceği gibi, 1. ve 5. genlerinin ekspresyon düzeyleri arasındaki ilişki en düşük ilişkilerden birisidir. Ancak iki gen ile kurulan model sonuçları incelendiğinde en yüksek performans sonucunun elde edildiği iki modelin de bu iki gen ile kurulan modeller olduğu görülmektedir. Aralarındaki ilişkinin en yüksek olduğu görülen iki gen olan 4. ve 5. genler ile kurulan modellerde ise en düşük performansın elde edildiği görülmektedir. Bu durum, modellerde çoklu doğrusal bağlantı sorunu oluşabileceğini düşündürmektedir.

Veri seti 2’de ise, %100 doğru sınıflama yapan modelin x1 ve x3 genleri ile elde edildiği görülmektedir. Tek genle kurulan modellerde elde edilen performansları farklı seviyede seçilmiş olan bu genler incelendiğinde x1 geninin en yüksek performansı verdiği, x3 geninin ise orta düzeyde bir performans gösterdiği görülmektedir. Her iki gen ile elde edilen sonuç ise performansı %100’e çıkarmaktadır. Ayrıca bu genlerin kendi aralarında elde edilen ilişkilere bakılacak olursa orta düzeyde bir ilişkiye sahip oldukları görülmektedir (Tablo 4.21.). Tek genle kurulan modellerde x2 geninden elde edilen performansın x3 geninden daha iyi olmasına rağmen, 2 genle kurulan modellerde x3 geninin model sonucunu %100’e çıkarması ise genlerin kendi aralarında olan ilişkilerinden kaynaklandığı düşünülmektedir. x2 geni ile x1 geni arasındaki ilişki oldukça kuvvetlidir. Bu durumun, Veri seti 1’de olduğu gibi, çoklu doğrusal bağlantı sorunundan oluşabileceğini düşündürmektedir.

Tablo 4.20. İki gen ile kurulan modellerin performans sonuçları

Kurulan Model	Veri Seti 1				Veri Seti 2			
	HUM	CCP	NRI	IDI	HUM	CCP	NRI	IDI
$y = x_1 + x_2$	0,868	0,947	0,105	0,112	0,868	0,947	0,105	0,112
$y = x_1 + x_3$	0,972	0,921	0,079	0,252	1,000	1,000	0,158	0,380
$y = x_1 + x_4$	0,961	0,895	0,053	0,231	0,928	0,921	0,079	0,171
$y = x_1 + x_5$	1,000	1,000	0,158	0,380	0,868	0,921	0,079	0,107
$y = x_2 + x_1$	0,868	0,947	0,289	0,416	0,868	0,947	0,289	0,416
$y = x_2 + x_3$	0,909	0,842	0,184	0,371	0,931	0,842	0,184	0,395
$y = x_2 + x_4$	0,903	0,816	0,158	0,352	0,828	0,763	0,105	0,162
$y = x_2 + x_5$	0,981	0,947	0,289	0,579	0,800	0,658	0,000	0,003
$y = x_3 + x_1$	0,972	0,921	0,132	0,356	1,000	1,000	0,289	0,579
$y = x_3 + x_2$	0,909	0,842	0,053	0,173	0,931	0,842	0,132	0,291
$y = x_3 + x_4$	0,928	0,921	0,132	0,232	0,687	0,737	0,026	0,011
$y = x_3 + x_5$	0,923	0,921	0,132	0,218	0,722	0,789	0,079	0,027
$y = x_4 + x_1$	0,961	0,895	0,158	0,422	0,928	0,921	0,395	0,623
$y = x_4 + x_2$	0,903	0,816	0,079	0,240	0,828	0,763	0,237	0,310
$y = x_4 + x_3$	0,928	0,921	0,184	0,318	0,687	0,737	0,211	0,263
$y = x_4 + x_5$	0,833	0,816	0,079	0,203	0,504	0,553	0,026	0,047
$y = x_5 + x_1$	1,000	1,000	0,263	0,429	0,868	0,921	0,447	0,716
$y = x_5 + x_2$	0,981	0,947	0,211	0,325	0,800	0,658	0,184	0,309
$y = x_5 + x_3$	0,923	0,921	0,184	0,162	0,722	0,789	0,316	0,437
$y = x_5 + x_4$	0,833	0,816	0,079	0,061	0,504	0,553	0,079	0,205

Tablo 4.21. Veri Seti 1 ve Veri seti 2’de bulunan genlerin ekspresyon değerleri arasındaki ilişkiler

Veri Seti1								
	x2		x3		x4		x5	
	r	p	r	p	r	p	r	p
x1	0,705	0,000	-0,435	0,006	-0,244	0,140	-0,196	0,238
x2			-0,458	0,004	-0,308	0,060	-0,155	0,352
x3					-0,211	0,203	-0,244	0,140
x4							0,740	0,000
Veri Seti 2								
	x2		x3		x4		x5	
	r	p	r	p	r	p	r	p
x1	0,705	0,000	0,325	0,047	0,227	0,170	0,083	0,618
x2			0,167	0,316	0,087	0,602	0,186	0,262
x3					0,556	0,000	0,036	0,830
x4							0,243	0,142

Tüm incelemelerde Spearman Korelasyon Katsayısı kullanılmıştır.

Tablo 4.22.’de, 3 gen ile kurulan modeller incelendiğinde; hem veri seti 1 hem de veri seti 2’de kurulan modellerin çoğunun %100 doğru sınıflama yapabildiği görülmüştür. Veri seti 1’de, x1 ve x3 genleri ile kurulmuş olan modeller birlikte %100 sınıflamaya ulaşabildiği için yanlarında bulunan herhangi bir gen daha fazla katkı sağlamamıştır. Ancak, modele ilk alınan genin bu iki gen dışında bir gen olduğu durumda, x1 ve x3 geninin modele katkısını gösteren NRI ve IDI değerlerinin diğer modellerden elde edilen değerlere göre daha yüksek olduğu görülmüştür. Veri seti 2’de de aynı durum söz konusudur.

Tablo 4.23.’de, 4 gen ile kurulan modeller incelendiğinde; hem veri seti 1 hem de veri seti 2’de kurulan modellerin biri hariç %100 doğru sınıflama yapabildiği görülmüştür. Her iki incelemede de %100 başarıya ulaşmayan tek model, tek gen ile yapılan incelemelerde en yüksek performansın elde edildiği genin bulunmadığı modeldir.

Tablo 4.22. Üç gen ile kurulan modellerin performans sonuçları

Kurulan Model	Veri Seti 1				Veri Seti 2			
	HUM	CCP	NRI	IDI	HUM	CCP	NRI	IDI
$y = x_1 + (x_2 + x_3)$	1,000	1,000	0,158	0,358	1,000	1,000	0,158	0,380
$y = x_1 + (x_2 + x_4)$	0,974	0,895	0,053	0,243	0,954	0,921	0,079	0,199
$y = x_1 + (x_2 + x_5)$	1,000	1,000	0,158	0,380	0,868	0,921	0,079	0,119
$y = x_1 + (x_3 + x_4)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_1 + (x_3 + x_5)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_1 + (x_4 + x_5)$	1,000	1,000	0,158	0,380	0,947	0,947	0,105	0,218
$y = x_2 + (x_1 + x_3)$	1,000	1,000	0,342	0,682	1,000	1,000	0,342	0,683
$y = x_2 + (x_1 + x_4)$	0,974	0,895	0,237	0,547	0,954	0,921	0,263	0,502
$y = x_2 + (x_1 + x_5)$	1,000	1,000	0,342	0,683	0,868	0,921	0,263	0,422
$y = x_2 + (x_3 + x_4)$	0,952	0,895	0,237	0,508	0,941	0,868	0,211	0,413
$y = x_2 + (x_3 + x_5)$	0,995	0,974	0,342	0,641	0,935	0,868	0,211	0,411
$y = x_2 + (x_4 + x_5)$	0,99	0,947	0,289	0,614	0,883	0,763	0,105	0,205
$y = x_3 + (x_1 + x_2)$	1,000	1,000	0,211	0,484	1,000	1,000	0,289	0,579
$y = x_3 + (x_1 + x_4)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_3 + (x_1 + x_5)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_3 + (x_2 + x_4)$	0,952	0,895	0,105	0,310	0,941	0,868	0,158	0,308
$y = x_3 + (x_2 + x_5)$	1,000	1,000	0,184	0,423	0,935	0,868	0,158	0,307
$y = x_3 + (x_4 + x_5)$	0,923	0,947	0,158	0,250	0,719	0,737	0,026	0,047
$y = x_4 + (x_1 + x_2)$	0,974	0,895	0,158	0,434	0,954	0,921	0,395	0,650
$y = x_4 + (x_1 + x_3)$	1,000	1,000	0,263	0,571	1,000	1,000	0,474	0,831
$y = x_4 + (x_1 + x_5)$	1,000	1,000	0,263	0,571	0,947	0,947	0,421	0,669
$y = x_4 + (x_2 + x_3)$	0,952	0,895	0,158	0,396	0,941	0,868	0,342	0,561
$y = x_4 + (x_2 + x_5)$	0,990	0,947	0,211	0,488	0,883	0,763	0,237	0,353
$y = x_4 + (x_3 + x_5)$	0,923	0,947	0,211	0,337	0,719	0,737	0,211	0,299
$y = x_5 + (x_1 + x_2)$	1,000	1,000	0,263	0,429	0,868	0,921	0,447	0,728
$y = x_5 + (x_1 + x_3)$	1,000	1,000	0,263	0,429	1,000	1,000	0,526	0,989
$y = x_5 + (x_1 + x_4)$	1,000	1,000	0,263	0,429	0,947	0,947	0,474	0,826
$y = x_5 + (x_2 + x_3)$	1,000	0,974	0,263	0,389	0,935	0,868	0,395	0,717
$y = x_5 + (x_2 + x_4)$	0,990	0,947	0,211	0,357	0,883	0,763	0,289	0,511
$y = x_5 + (x_3 + x_4)$	0,923	0,947	0,211	0,194	0,719	0,737	0,263	0,457

Tablo 4.23. Dört gen ile kurulan modellerin performans sonuçları

Kurulan Model	Veri Seti 1				Veri Seti 2			
	HUM	CCP	NRI	IDI	HUM	CCP	NRI	IDI
$y = x_1 + (x_2 + x_3 + x_4)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_1 + (x_2 + x_3 + x_5)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_1 + (x_2 + x_4 + x_5)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_1 + (x_3 + x_4 + x_5)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_2 + (x_1 + x_3 + x_4)$	1,000	1,000	0,342	0,683	1,000	1,000	0,342	0,683
$y = x_2 + (x_1 + x_3 + x_5)$	1,000	1,000	0,342	0,683	1,000	1,000	0,342	0,683
$y = x_2 + (x_1 + x_4 + x_5)$	1,000	1,000	0,342	0,683	1,000	1,000	0,342	0,683
$y = x_2 + (x_3 + x_4 + x_5)$	1,000	1,000	0,342	0,683	0,931	0,868	0,211	0,435
$y = x_3 + (x_1 + x_2 + x_4)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_3 + (x_1 + x_2 + x_5)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_3 + (x_1 + x_4 + x_5)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_3 + (x_2 + x_4 + x_5)$	1,000	1,000	0,211	0,466	0,931	0,868	0,158	0,331
$y = x_4 + (x_1 + x_2 + x_3)$	1,000	1,000	0,263	0,571	1,000	1,000	0,474	0,831
$y = x_4 + (x_1 + x_2 + x_5)$	1,000	1,000	0,263	0,571	1,000	1,000	0,474	0,831
$y = x_4 + (x_1 + x_3 + x_5)$	1,000	1,000	0,263	0,571	1,000	1,000	0,474	0,831
$y = x_4 + (x_2 + x_3 + x_5)$	1,000	1,000	0,263	0,568	0,931	0,868	0,342	0,583
$y = x_5 + (x_1 + x_2 + x_3)$	1,000	1,000	0,263	0,429	1,000	1,000	0,526	0,989
$y = x_5 + (x_1 + x_2 + x_4)$	1,000	1,000	0,263	0,429	1,000	1,000	0,526	0,989
$y = x_5 + (x_1 + x_3 + x_4)$	1,000	1,000	0,263	0,429	1,000	1,000	0,526	0,989
$y = x_5 + (x_2 + x_3 + x_4)$	1,000	0,974	0,263	0,429	0,931	0,868	0,395	0,741

Tablo 4.24.'de, 5 gen ile kurulan tam modeller incelendiğinde; hem veri seti 1 hem de veri seti 2'de kurulan modellerin tümünün %100 doğru sınıflama yapabildiği görülmüştür.

Tablo 4.24. Beş gen ile kurulan modellerin performans sonuçları

Kurulan Model	Veri Seti 1				Veri Seti 2			
	HUM	CCP	NRI	IDI	HUM	CCP	NRI	IDI
$y = x_1 + (x_2 + x_3 + x_4 + x_5)$	1,000	1,000	0,158	0,380	1,000	1,000	0,158	0,380
$y = x_2 + (x_1 + x_3 + x_4 + x_5)$	1,000	1,000	0,342	0,683	1,000	1,000	0,342	0,683
$y = x_3 + (x_1 + x_2 + x_4 + x_5)$	1,000	1,000	0,211	0,485	1,000	1,000	0,289	0,579
$y = x_4 + (x_1 + x_2 + x_3 + x_5)$	1,000	1,000	0,263	0,571	1,000	1,000	0,474	0,831
$y = x_5 + (x_1 + x_2 + x_3 + x_4)$	1,000	1,000	0,263	0,429	1,000	1,000	0,526	0,989

Her iki veri setinin de öncelikle Box Cox dönüşümü ile normal dağılım göstermesi sağlanmıştır. Sonrasında Kuadratik diskriminant analizi yapılmıştır. Veri Seti 1’de elde edilen sonuçlarda, 4 gen değeri ile %100 sınıflama başarısı sağlanmıştır. Model dışında kalan tek gen ise ikinci sırada yüksek performans göstermiş olan x_2 geni olmuştur. Modele öncelikli giren gen ise en düşük performansa sahip olan 5.gen olmuştur.

Veri Seti 2’de de öncelikle Box Cox dönüşümü ile normal dağılım varsayımı sağlanmıştır. Sonrasında Kuadratik diskriminant analizi yapılmıştır. Elde edilen sonuçlarda, 3 gen değeri ile %97.4 oranında sınıflama başarısı sağlanmıştır. Bu modelde %100 başarıya ulaşılamamıştır. Model dışında kalan genler, tek gen ile kurulan modellerde en düşük performans sonucunu göstermiş olan x_4 ve x_5 genleri olmuştur (Tablo 4.25.).

Tablo 4.25. Veri Seti 1 ve Veri Seti 2 ile uygulanan Kuadratik Diskriminant Analizi Sonuçları

Veri Seti 1	Wilks' Lambda	Kesin F	Doğru Sınıflama Olasılığı
x5	0,309	39,091	1,000
x3	0,126	30,869	
x1	0,074	29,481	
x4	0,057	25,515	
Veri Seti 2	Wilks' Lambda	Kesin F	Doğru Sınıflama Olasılığı
x1	0,392	27,197	0,974
x3	0,171	24,094	
x2	0,135	18,919	

5.TARTIŞMA

Sağlık alanında, testlerin ya da kestirim modellerinden elde edilen risklerin doğru değerlendirilmesi, hasta ve sağlıklı kişileri belirlemede oldukça önemlidir. Hastalık durumunun iki kategorili olduğu durumlarda, hastalığı sınıflandırma ya da riskleri belirleyebilmek için geliştirilmiş ve yaygın olarak kullanılan birçok yöntem bulunmasına rağmen, hastalık durumu çok kategorili olduğu durumlarda bu yöntemlerin çoğu uygulanamamaktadır. Özellikle çok kategorili hastalık durumlarında sınıflandırma işlemleri, tanısal modelin performansını ve bu modele eklenen yeni bir belirtecin model performansı üzerinde yarattığı değişimi incelemek oldukça karmaşık işlemlerdir.

Belirteçlerin performansını değerlendirmede en yaygın kullanılan yöntem ROC eğrisi yöntemidir. HUM'dan önce geliştirilmiş olan 3 Yönlü ROC analizi (VUS) yöntemi de sadece 3 sınıflı hastalık durumlarında belirteçlerin performansını inceleyebildiği için HUM gibi 3 ve üzeri kategori için inceleme yapabilen bir yöntem geliştirilme ihtiyacı ortaya çıkmıştır. Bu yöntemden yola çıkılarak geliştirilmiş olan HUM yöntemi 2004 yılından beri kullanılmaktadır.

Yapmış olduğumuz çalışmada, ikinci ve üçüncü uygulamalarda kullanmış olduğumuz veri setlerinin farklı sınıflandırma yöntemleri ile incelendiği bir çok çalışma vardır. Bu çalışmalarda, klasik yöntemler, makine öğrenmesi yöntemleri ya da veri madenciliği yöntemleri gibi birçok farklı yaklaşıma sahip sınıflandırma yöntemleri kullanılmıştır. Özellikle Golub ve arkadaşlarının verisi ile mikrodizin verilerinin incelendiği, kanser hastalığının sınıflandırılması açısından en önemli genleri belirlemeye yönelik çalışma sayısı oldukça fazladır. Bu çalışmaların genel amacı sınıflandırma doğruluğu açısından en önemli genleri belirlemek olduğu gibi, bu genlerin bir kısmı ile kurulan çok belirteçli modellerin incelenmesini de içermektedir. Ancak bu çalışmaların çok azı çok kategorili hastalık sınıflandırması açısından yaklaşmıştır. Genel olarak 2 sınıflı hastalık durumlarını tahmin etme üzerine uygulamalar yapılmıştır.

Golub ve arkadaşları (24) yapmış oldukları çalışmada, 10 gen ile %100 sınıflama başarısını ulaştıklarını belirtmişlerdir. Golub ve arkadaşlarının veri seti ile

yapılmış çalışmalar incelendiğinde; Sewak ve arkadaşları (37), 7129 gen olan bu veri setini 250 gene kadar indirebilen bir model geliştirmiş ve bu model ile 3 sınıfı %100 doğrulukla ayırabildiklerini belirtmişlerdir. Xu ve arkadaşları (38), Lineer Diskriminant analizi ile yapmış oldukları çıkarsamalarla farklı yöntemleri sınıflama performansları açısından incelemişlerdir. En başarılı buldukları yöntem olan modifiye edilmiş Lineer Diskriminant Analizi ile elde edilen sınıflama ile %100 başarıya ulaştıkları gen sayısının 70 olduğunu göstermişlerdir. Li ve arkadaşları (39) yapmış oldukları çalışmada, VUS ve CCP ile inceleme yapmış ve 2 gen ile %100 başarıya ulaşabilmişlerdir ayrıca HUM yöntemi ile yapmış oldukları çalışmada da (1) 2 gen ile %100 başarıya ulaşmışlardır.

Her ne kadar HUM yöntemi geliştirilmiş olsa da birden çok belirteç ile kurulmuş ya da kurulması planlanan tanı modellerinde, modele yeni bir belirteç eklenmesinin modele olan katkısı tam olarak incelenememektedir. Yani modele eklenen belirtecin yeterli performans artışı sağlayıp sağlayamayacağı sadece HUM yöntemi kullanılarak belirlenememektedir. Bu çalışmalardan farklı olarak; hastalık durumunun 2 sınıflı olduğu ve incelenen modellere yeni bir belirteç eklenmesi ile elde edilen performans artışını inceleyen çalışmalar da oldukça fazladır. 2’den fazla sınıf sayısına sahip hastalık durumları için yapılmış çalışma sayısı ise oldukça azdır.

Bu ihtiyaçtan dolayı da 2008 yılında NRI yöntemi geliştirilmiştir. Pencina ve arkadaşları 2008 yılında modele yeni bir belirteç eklendiğinde oluşturulan yeni sınıflama sonucunun performansını gösteren iki yeni performans ölçüsü sunmuşlardır. Bunlar; NRI ve IDI yöntemleridir (4).

Özellikle çok kategorili hastalık durumlarında hastalık sınıfları arasındaki geçişlerin tablolar ile incelenebilmesi oldukça zor ve yanıltıcı olabileceği için bu konuda yapılmış olan çalışma bulunmamaktadır. Çok kategorili hastalık durumlarında yapılmış olan 2 adet çalışma vardır ve bu çalışmaların her ikisi de Li ve arkadaşları tarafından “Sürekli NRI” yöntemi ile incelenmiştir.

Li ve arkadaşlarının 2013 ve 2017 yıllarında yapmış oldukları bu çalışmalarda; HUM ve NRI yöntemleri birlikte incelenmiş ve 2 gen ile %100 başarıya ulaştıklarını göstermişlerdir. Bu genleri belirlerken en yüksek HUM ve CCP değerine sahip olan 2

geni incelemişlerdir. İkinci gen modele eklendiğinde %100 doğru sınıflandırma başarısına ulaşabildiği için diğer genleri modelde incelememişlerdir.

Bu çalışmada, en yüksek HUM değerine sahip olan 5 gen, farklı kombinasyonlarla kurulmuş modellerle incelendiğinden, alanyazında bu çalışmaya benzer bir çalışma bulunmamaktadır. Daha önce yapılmış benzer çalışmalarda olduğu gibi bu çalışmada da en yüksek HUM değerine sahip olan 2 gen ile modelleme yapıldığında %100 doğru sınıflandırma başarısına ulaşılmıştır. Diğer çalışmalardan farklı olarak, farklı modeller de incelenmiş olduğundan, en yüksek performansa ulaşabilecek alternatif modeller de incelenmiştir. Örneğin, ekspresyon düzeyleri hatalı elde edilmiş bir gen olduğu durumda farklı bir gene yönelme durumu olabilecek ya da ulaşılması daha kolay olan üçüncü bir gen ile modelleme yapılabilir.

6. SONUÇ VE ÖNERİLER

Farklı ilişki yapılarından ve örnek genişliklerinden elde edilen modellerin performans sonuçları arasındaki ilişkilerde; modelde bulunan belirteç sayısı 2 iken pozitif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin sonuçlarında, klasik yöntemler ile NRI ve IDI arasında bir ilişki görülmemiştir. Ancak negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin performans sonuçlarında, klasik yöntemlerle NRI ve IDI arasında pozitif yönde ilişkiler olduğu görülmektedir. Tüm modellerde, klasik yöntemler olan HUM ve CCP yöntemlerinden elde edilen sonuçlarda, CCP değerleri HUM değerlerine göre NRI ve IDI ile daha fazla ilişki göstermiştir.

CCP yöntemi örneklem genişliğinden etkilenmektedir. HUM yöntemi ise örneklem genişliğinden etkilenmemektedir. NRI ve IDI yöntemlerinin sonuçları ise örneklem genişliğinden etkilenmemekle birlikte, bağımlı değişken ve bağımsız belirteçlerin aralarındaki ilişkilerden etkilenmektedir.

Genel olarak, bağımlı değişken ile negatif ilişki yapısında benzetim yapılarak elde edilen verilerle oluşturulan modellerin sonucunda elde edilen NRI ve IDI değerlerinin klasik yöntemler olan HUM ve CCP ile daha yüksek derecede ilişkili olduğu görülmektedir.

Klasik yöntemler, modele eklenen belirtecin sırasından etkilenmemektedir. Ancak NRI ve IDI modele eklenen belirtecin sırasından etkilenmektedir. Doğru bir klinik modelleme yapılması durumunda birçok belirtecin birlikte modellendiğinde elde edilebilecek doğru sınıflama performansının, doğru sırada eklenen daha az sayıda belirteç ile yapabileceği düşünülmektedir.

Ayrıca bu yöntemler, tanı modeli belirlemede ve özellikle modele katkı sağlayacak olan belirteçlerin seçiminde oldukça yol göstericidir. Tanı kestirim modeli kurulması aşamalarında büyük sıkıntı yaratabilen sorunların aşılmasında bu yöntemlerin fayda sağlayabileceği düşünülmektedir. Özellikle büyük veri yapılarında modelleme yapılırken, boyut indirgeme yapılabilmesi açısından da başarılı yöntemler oldukları düşünülmektedir.

Modele eklenen bir belirtecin, model performans sonucunu yükseltmesi tanı tahmin modelinin kurulabilmesi için tek başına yeterli bir kriter değildir. Mutlaka sınıflandırma performansının da incelendiği bir performans ölçüsü ile birlikte, modele katkı sağlayan en başarılı belirtecin bulunması için NRI ve IDI yöntemleri de kullanılarak çalışma amacına uygun olan modelin belirlenmesi gereklidir. Tek belirteçli modelleme sonucunda, oldukça yüksek bir sınıflandırma başarısına sahip bir belirtecin modelde bulunması ne kadar önemli ise, bu belirteçten sonra modele eklenen belirteçlerin performanslarının adımsal olarak incelenmesi de o kadar önemlidir. Ayrıca bazı durumlarda tek başına oldukça yüksek bir sınıflandırma başarısına sahip olmasına rağmen farklı sebeplerle (maliyet, risk, ölçülemez vb.gibi) modele eklenemeyen ya da eklenmesi oldukça zor olan bazı belirteçler olabilir. Örneğin, tek başına sınıflandırma başarısı yüksek olan ama maliyeti de oldukça yüksek olan ya da girişimsel olarak oldukça risk taşıyan bir tanı testi olabilir. Bu gibi durumlarda, adımsal olarak NRI ve IDI yöntemleri ile modele eklenebilecek diğer belirteçler belirlenerek alternatif bir model oluşturulabilir. Ya da performansı yüksek olmayan bir çok belirtecin bulunduğu bir durumda, model oluşturulma aşamasında adımsal olarak incelemeler yapılabileceği için en az belirteç ile en yüksek sınıflandırma performansına ulaşılabilir. Her durumda, HUM ve CCP yöntemleri ile NRI ve IDI yöntemlerinin birlikte kullanılması oldukça önemlidir.

Genel olarak sonuçlara bakıldığında ise, Kuadratik Diskriminant analizi ile elde edilen doğru sınıflama başarısı sonuçlarının HUM ve CCP ile elde edilen model sınıflandırma performansı sonuçlarından daha düşük olduğu görülmüştür. Kuadratik Diskriminant Analizi olmasa da literatürde kullanılan birçok farklı yöntemde parametrik test varsayımlarına ihtiyaç duyulmasından dolayı; HUM, CCP, NRI ve IDI yöntemleri parametrik test varsayımları gerektirmediğinden oldukça avantajlıdır. Ayrıca HUM, CCP, NRI ve IDI sonuçlarının incelenebilmesi için veriye ve amaca uygun olarak her değişken türü ile her türlü regresyon modeli kullanılabilir. Bu yöntemler, gerek bağımlı değişkenin gerekse modelleme için incelenecek olan belirteçlerin, her değişken türünde kullanılabilmesi ve tüm regresyon modelleri ile modelleme yapılabilmesi açısından oldukça tercih edilen performans ölçüleri olmuşlardır. Ayrıca daha az sayıda belirteç ile daha doğru sınıflandırma performansı gösterebilmektedir.

Sonuçlardan da görülebileceği gibi klasik sınıflandırma yöntemleri ile kurulan modellerde, incelenen belirteçlerin tek belirteç ile kurulan modellerle elde edilen sonuçlarının göz önüne alınması gerektiği görülmüştür. Tek belirteç ile kurulan modellerde sınıflandırma performansı düşük olan belirteçlerin, klasik yöntemlerle kurulan modellerde modele girme sırası önemsizdir ve modele katkı sağlayamayarak model dışında kalmaktadır.

Bu sonuçlardan yola çıkılarak, çok kategorili hastalık durumlarında, adımsal olarak NRI ve IDI yöntemlerinin kullanılabileceği bir yöntem geliştirilerek tüm belirteçlerin performansları göz önüne alınabilir ve özellikle belirteç sayısı yüksek olan veri setlerinde (örneğin herhangi bir hastalığın mikrodizin veri seti) birkaç belirteç ile incelenen hastalık sınıflarını doğru sınıflandırma başarısı yüksek tanı modelleri elde edilebilir. Ayrıca kategori bazlı incelemeler de bu modellere eklenerek hangi belirtecin hangi kategori üzerinde etkisinin yüksek olduğu belirlenebilir. Bu durumda sınıflandırma başarısı düşük olan kategoriye etkisi olan belirteç ya da belirteçler belirlenebilir ve daha az sayıda belirteç ile tüm kategoriler doğru sınıflandırılabilir.

Ayrıca çok kategorili hastalık durumlarında risk sınıflandırmaları için iki kategorili NRI yönteminde olduğu gibi hesaplamalar geliştirebilir ve çapraz tablolar ile Net Yeniden Sınıflandırma İyileştirmesi görselleştirilebilir.

KAYNAKLAR

1. Li J, Jiang B, Fine J.P. Multicategory reclassification statistics for assessing improvements in diagnostic accuracy. *Biostatistics*. 2013; 14, 2, pp. 382–394.
2. Mossman D. Three-way ROCs. *Med Decis Making*. 1999; vol 19: 78.
3. Nakas C.T, Yiannoutsos C.T. Ordered multiple-class ROC analysis with continuous measurements. *Statist. Med*. 2004; 23: 3437–3449.
4. Li J, Fine J.P. ROC analysis with multiple classes and multiple tests: methodology and its application in microarray studies. *Biostatistics*. 2008; 9, 3, pp. 566–576.
5. Zhou X, Obuchowski N.A, McClish D.K. *Statistical methods in diagnostic medicine*. New Jersey: John Wiley & Sons, Inc; 2011.
6. Pencina M, D’Agostino Sr R.B, D’Agostino Jr R.B, Vasan R.S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statist. Med*. 2008; 27: 157–172.
7. Karaismailoğlu E. İzlem Süresi İçerisinde Belirteçlerin Performanslarını Değerlendirmede Zamana Bağlı ROC Eğrisinin Kullanımı [Bütünleşik Doktora tezi]. Ankara: Hacettepe Üniversitesi; 2015.
8. Karakaya J. Üç Yönlü ROC Analizi ve Ortak Değişken Düzeltmesi [Doktora tezi]. Ankara: Hacettepe Üniversitesi; 2012.
9. Scurfield B.K. Multiple-event forced-choice tasks in the theory of signal detectability. *Journal of Mathematical Psychology*. 1996; 40, 253_269.
10. Karaağaoğlu E, Karakaya J, Kılıçkap M. Tanı testlerinin değerlendirilmesinde istatistiksel yöntemler. Ankara: Detay Yayıncılık; 2016.
11. Xiong C, van Belle G, Miller J.P, Morris J.C. Measuring and estimating diagnostic accuracy when there are three ordinal diagnostic groups. *Statist. Med*. 2006; 25: 1251–1273.
12. Li J, Zhou X. Nonparametric and semiparametric estimation of the three way receiver operating characteristic surface. *Journal of Statistical Planning and Inference*. 2009; 139: 4133- 4142.
13. Dreiseitl S, Ohno-Machado L, Binder M. Comparing Three-class diagnostic tests by three-way ROC analysis. *Med Decis Making*. 2000; 20: 323.
14. Li J, Fine J.P, Pencina M. Multi-category diagnostic accuracy based on logistic regression. *Statistical Theory and Related Fields*. 2017; 1: 2, 143-158.
15. Steyerberg E.W, Vickers A.J, Cook N.R, Gerds T, Gonen M, Obuchowski N ve arkadaşları. Assessing the performance of prediction models A framework for traditional and novel measures. *Epidemiology*. 2010; 21 – 1.
16. Steyerberg E.W, Calster B.V, Pencina M. Performance measures for prediction models and markers: evaluation of predictions and classifications. *Rev Esp Cardiol*. 2011;64(9):788–794.

17. Pencina M.J, D'Agostino R.B, Pencina K.M, Janssens A.C.J.W, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol.* 2012;176(6): 473–481.
18. Muchlenbruch K, Heraclides A, Steyerberg E.W, Joost H, Boeing H, Schulze M.B. Assessing improvement in disease prediction using net reclassification improvement: impact of risk cut-offs and number of risk categories. *Eur J Epidemiol.* 2013; 28: 25–33.
19. Pencina M.J, Steyerberg E.W, D'Agostino R.B. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med.* 2011; January 15; 30(1): 11–21.
20. Leening M.J.G, Vedder M.M, Witteman J.C.M, Pencina M.J, Steyerberg E.W. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014;160:122-131.
21. Kerr K.F, Wang Z, Janes H, McClelland R.L, Psaty B.M, Pepe M.S. Net Reclassification indices for evaluating risk-prediction instruments: A critical review. *Epidemiology.* 2014; January; 25(1): 114–121.
22. Li J, Gao M. mcca: Multi-category classification accuracy. R package version 0.5.0. <https://CRAN.R-project.org/package=mcca>. 2019.
23. Quinlan, J.R. Thyroid Disease Data Set [Internet]. 1987 [01 Ocak 2018]. Erişim adresi: <http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease>.
24. Golub T. R, Slonim D. K, Tamayo P, Huard C, Gaasenbeek M, Mesirov J. P ve arkadaşları. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science.*1999; 286 (5439), 531-537.
25. Greenland S. The need for reorientation toward cost-effective prediction comments on 'Evaluating the added predictive ability of a new marker from area under the ROC curve to reclassification and beyond. *Statist. Med.* 2008; 27: 199–206.
26. Steyerberg E.W, Pencina M.J, Lingsma H.F, Kattan M.W, Vickers A.J, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers a review and illustration. *Eur J Clin Invest.* 2012; 42 (2): 216–228.
27. Demler O.V, Pencina M.J, Cook N.R, D'Agostino Sr R.B. Asymptotic distribution of Δ AUC, NRIs, and IDI based on theory of U-statistics. *Statist. Med.* 2017; 36, 3334–3360.
28. Pickering J.W, Endre Z.H. New metrics for assessing diagnostic potential of candidate biomarkers. *Clin J Am Soc Nephrol.* 2012; 7, 1- 10.
29. Alpar R. Uygulamalı çok değişkenli istatistiksel yöntemler. Ankara: Detay Yayıncılık; 2017.
30. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. 2014
31. Demirtas H, Amatya A. OrdNor: An R Package for concurrent generation of correlated ordinal and normal data. *Journal of Statistical Software.* 2015; 68(2), 1-14.

32. Venables W. N, Ripley B. D. Modern applied statistics with S. Fourth Edition. New York: Springer; 2002.
33. To Duc K, Chiogna M, Adimari G. bcROCsurface: Bias-Corrected methods for Estimating the ROC surface of continuous diagnostic tests. R package version 1.0-3. <https://CRAN.R-project.org/package=bcROCsurface>. 2018.
34. Kundu S, Aulchenko Y.S, Janssens A.C.J.W. PredictABEL: Assessment of risk prediction models. R package version. 1.2-2. <https://CRAN.R-project.org/package=PredictABEL>. 2014.
35. Warnes G. R, Bolker B, Lumley T. gtools: Various R Programming Tools. R package version 3.8.1. <https://CRAN.R-project.org/package=gtools>. 2018.
36. IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.
37. Sewak M.S, Reddy N.P, Duan Z. Gene expression based leukemia sub-classification using committee neural networks. *Bioinformatics and Biology Insights*. 2009; 3 89–98.
38. Xu P, Brock G.N, Parrish R.S. Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. *Computational Statistics and Data Analysis*. 2009; 1674–1687.
39. Li J, Chow Y, Wong W.K, Wong T.Y. Sorting multiple classes in multi-dimensional ROC analysis: parametric and nonparametric approaches. *Biomarkers*. 2014; 19: 1, 1-8.

8. EKLER

EK-1: Tez Çalışması Orijinallik Raporu

ÇOK KATEGORİLİ NRI_1

ORIJINALLIK RAPORU

% 4	% 4	% 2	% 3
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	www.math.niu.edu İnternet Kaynağı	% 1
2	www.soften.ktu.lt İnternet Kaynağı	% 1
3	www.docstoc.com İnternet Kaynağı	% 1
4	interm.gtk.gau.hu İnternet Kaynağı	<% 1
5	Submitted to TechKnowledge Turkey Öğrenci Ödevi	<% 1
6	www.etkinlik.adu.edu.tr İnternet Kaynağı	<% 1
7	docslide.us İnternet Kaynağı	<% 1
8	polen.itu.edu.tr İnternet Kaynağı	<% 1
9	dizc.ru İnternet Kaynağı	<% 1

EK-2: Dijital Makbuz

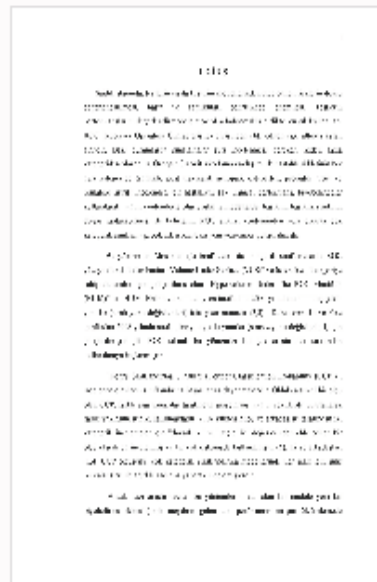


Dijital Makbuz

Bu makbuz ödevinizin **TumutIn**'e ulaştığını bildirmektedir. Gönderiminize dair bilgiler şöyledir:

Gönderinizin ilk sayfası aşağıda gönderilmektedir.

Gönderen: **Hande Şenol**
 Ödev başlığı: **ÇOK KATEGORİLİ NRI_1**
 Gönderi Başlığı: **ÇOK KATEGORİLİ NRI_1**
 Dosya adı: **tumitin.docx**
 Dosya boyutu: **592.57K**
 Sayfa sayısı: **85**
 Kelime sayısı: **17,281**
 Karakter sayısı: **101,484**
 Gönderim Tarihi: **23-Tem-2019 03:31PM (UTC+0300)**
 Gönderim Numarası: **1154321480**



9. ÖZGEÇMİŞ

1983 yılında Kocaeli ili İzmit ilçesinde doğdu. İlk, orta ve lise öğrenimini İzmit'te 24 Kasım Anadolu Lisesi'nde tamamladı. 2007 yılında Mimar Sinan Üniversitesi İstatistik bölümünden mezun oldu. 2011 yılında Pamukkale Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Yüksek Lisans programında eğitime başladı ve 2014 yılında tamamladı. Aynı yıl Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Doktora programında eğitime başladı. Halen Pamukkale Üniversitesi Biyoistatistik Anabilim Dalı'nda Öğretim Görevlisi olarak çalışmaktadır.

