



# MerQur'da Sınıflandırma: ROC Eğrisinden Gradient Boosting'e Modern Sınıflandırıcılar

## Classification in MerQur: From ROC Curves to Gradient Boosting — Modern Classifiers

Ömer K. Örucü<sup>1</sup>

<sup>1</sup> Süleyman Demirel Üniversitesi, Mimarlık Fakültesi, Peyzaj Mimarlığı Bölümü, Isparta/Türkiye ORCID:

0000-0002-2162-7553 E-posta: omerorucu@sdu.edu.tr · Resmi site:

<https://www.sekizgenacademy.com/journals/index.php/merqur/tr/index>

**Yazışmadan sorumlu yazar:** Ömer K. Örucü (omerorucu@sdu.edu.tr)

**Tür:** Davetli Editöryal Sunum / Invited Editorial Showcase **Geliş:** 2026-05-17 · **Kabul:** 2026-05-17 · **Yayın:** 2026-05-17 **DOI:** — (ISSN başvurusu sonrası eklenecek)

### Öz

Sınıflandırma, akademik araştırmanın ve uygulamalı veri biliminin en yaygın görevlerinden biridir: bir gözlemi önceden tanımlı kategorilerden birine atamak. Klasik istatistik (lojistik regresyon) ile modern makine öğrenmesi (random forest, gradient boosting, SVM) arasındaki köprünün üzerinde, sınıflandırma performansının doğru ölçülmesi ve raporlanması ayrı bir metodolojik özen gerektirir. Bu çalışmada **MerQur** masaüstü yazılımının Sınıflandırma kategorisinde sunulan **6 analiz** ayrıntılı olarak tanıtılmıştır: ROC Eğrisi, TSS (True Skill Statistic), Karmaşıklık (Confusion) Matrisi Metrikleri, Random Forest Sınıflandırma, Support Vector Machine (SVM) ve Gradient Boosting Sınıflandırma. Her analiz için (i) yöntemin temeli ve sınıflandırma görevinde nereye uyduğu, (ii) hiperparametreler ve seçim stratejileri, (iii) MerQur'daki form alanları ve seçenekler, (iv) raporlanan performans metrikleri (doğruluk, kesinlik, duyarlılık, F1, AUC, TSS, kappa, Matthews correlation), ve (v) tipik bir araştırma sorusu için yorumlama önerisi sunulmuştur. ROC ve AUC'nin eşik-bağımsız değerlendirmedeki rolü, dengesiz sınıflarda doğru metriğin seçimi (F1 / MCC / TSS), Random Forest'in değişken önemi yorumlanması, SVM'in çekirdek seçimi ve Gradient Boosting'in ardışık ağaçlarla aşırı uydurma kontrolü tartışılmıştır. MerQur'un Sınıflandırma kategorisi, klasik tanılayıcı eşik analizinden modern ensemble yöntemlerine uzanan kapsamı tek bir grafik arayüzde sunmakta ve k-katlı çapraz doğrulamayı standart olarak içermektedir.

**Anahtar Kelimeler:** sınıflandırma, ROC, AUC, TSS, confusion matrix, random forest, SVM, gradient boosting, makine öğrenmesi, MerQur

### Abstract

Classification is one of the most common tasks in academic research and applied data science: assigning an observation to one of predefined categories. On the bridge between classical statistics (logistic regression)

and modern machine learning (random forest, gradient boosting, SVM), the correct measurement and reporting of classification performance requires special methodological care. This study introduces in detail the **6 analyses** offered in MerQur's Classification category: ROC Curve, TSS (True Skill Statistic), Confusion Matrix Metrics, Random Forest Classification, Support Vector Machine (SVM), and Gradient Boosting Classification. For each: (i) the basis of the method and its place in the classification task, (ii) hyperparameters and selection strategies, (iii) form fields and options in MerQur, (iv) reported performance metrics (accuracy, precision, recall, F1, AUC, TSS, kappa, Matthews correlation), and (v) interpretation guidance for a typical research question. The role of ROC and AUC in threshold-independent evaluation, the selection of correct metrics in imbalanced classes (F1 / MCC / TSS), the interpretation of variable importance in Random Forest, kernel selection in SVM, and overfitting control in Gradient Boosting are discussed. MerQur's Classification category covers the spectrum from classical diagnostic threshold analysis to modern ensemble methods within a single graphical interface and includes k-fold cross-validation as standard.

**Keywords:** classification, ROC, AUC, TSS, confusion matrix, random forest, SVM, gradient boosting, machine learning, MerQur

## 1. Giriş

Sınıflandırma görevleri akademik araştırmanın çok geniş bir bölümünde merkezdedir: tıbbi teşhis (hasta vs sağlıklı), tür dağılım modelleri (uygun vs uygunsuz habitat), kredi skoru (geri ödenir vs geri ödenmez), spam tespiti, görüntü tanıma. Bu çalışmalarda **doğru sınıflandırmak** kadar **doğru ölçmek** ve **doğru raporlamak** kritiktir. Sınıflandırma performansının değerlendirilmesi çok boyutlu bir mesele olduğundan, tek bir "doğruluk" rakamı genellikle yetersiz, hatta yanıltıcıdır — özellikle dengesiz sınıflar bulunduğu (Saito & Rehmsmeier, 2015; Chicco & Jurman, 2020).

Yöntem ailesinde iki ana kuşak vardır: **klasik istatistiksel sınıflandırıcılar** (lojistik regresyon, ayırma analizi) ve **modern makine öğrenmesi sınıflandırıcıları** (random forest, gradient boosting, SVM, neural networks). MerQur, Sınıflandırma kategorisinde her iki kuşağa ait yöntemleri ve performans değerlendirme araçlarını bir arada sunar.

Bu çalışmanın amacı, MerQur'un Sınıflandırma kategorisinde yer alan 6 analizi sistemli olarak tanıtmak; performans metriklerinin doğru yorumlanmasına ve modern ML yöntemlerinin akademik araştırmada güvenli kullanımına ilişkin yönlendirici bilgi sunmaktır.

## 2. ROC Eğrisi ve AUC

### 2.1 Yöntem

Receiver Operating Characteristic (ROC) eğrisi, ikili bir sınıflandırıcının **olasılık tahminleri** üzerinden, tüm olası eşiklerde duyarlılık (sensitivity) — 1 - özgüllük (specificity) grafiğini çizer. Eğri altındaki alan **AUC** (area under curve), sınıflandırıcının rastgele seçilen bir pozitif örneği rastgele seçilen bir negatif örnekten daha yüksek olasılık vermesi şansıdır.

- AUC = 0.5: rastgele tahmin (eğri köşegen)
- AUC = 1.0: mükemmel sınıflandırma
- AUC > 0.7: kabul edilebilir, > 0.8: iyi, > 0.9: mükemmel (genelleme)

ROC eşik-bağımsızdır; bu nedenle modelin “kalitesini” tek bir rakamla özetler. Ancak dengesiz sınıflarda AUC yanıltıcı olabilir — bu durumda **PR (Precision-Recall) eğrisi** daha bilgilendiricidir.

## 2.2 MerQur’da uygulama

*İstatistik → Sınıflandırma → ROC Eğrisi*

. Form alanları:

- **Hedef Sütun:** ikili gerçek değer (0/1)
- **Skor Sütunu:** tahmin olasılığı veya skor
- **Pozitif sınıf:** hangi değer “pozitif” sayılacağı

Çıktıda ROC grafiği, AUC ve %95 GA (DeLong), optimum eşik (Youden’s J), bu eşikte duyarlılık/özgüllük, **PR eğrisi** (opsiyonel) ve farklı eşikte sınıflandırma matrisi tablosu sunulur.

## 2.3 Uygulama örneği

Bir tanı testinin AUC = 0.84, optimum eşik = 0.42 (Youden’s J = 0.58). Bu eşikte duyarlılık = 0.81, özgüllük = 0.77. Klinik kullanım için kabul edilebilir.

## 3. TSS (True Skill Statistic)

### 3.1 Yöntem

TSS, sınıflandırıcının “şans üstü” performansını ölçer:

$$TSS = \text{sensitivity} + \text{specificity} - 1$$

Aralık  $[-1, +1]$ ; 0 = şans, 1 = mükemmel. AUC’nin tek eşik versiyonu olarak düşünülebilir. Özellikle **tür dağılım modellerinde (SDM)** yaygındır çünkü dengesiz sınıflarda kappa’dan daha az duyarlıdır (Allouche ve ark., 2006).

### 3.2 MerQur’da uygulama

*İstatistik → Sınıflandırma → TSS.* Form: **Hedef + Tahmin + Eşik** (varsayılan: optimum TSS).

Çıktıda TSS, optimum eşik, bu eşikte duyarlılık/özgüllük, eşik fonksiyonu grafiği.

### 3.3 Uygulama örneği

Bir ağaç türünün varlık/yokluk tahmininde TSS = 0.62 (eşik 0.37) — şans üstü kabul edilebilir performans.

## 4. Karmaşıklık (Confusion) Matrisi Metrikleri

### 4.1 Yöntem

Confusion matrisi  $2 \times 2$  (ikili) veya  $K \times K$  (çok sınıflı) tablodur; satırlar gerçek, sütunlar tahmin sınıflarıdır. Bu matristen pek çok performans metriği türetilir:

- **Doğruluk (Accuracy):**  $(TP + TN)/N$
- **Kesinlik (Precision):**  $TP/(TP + FP)$
- **Duyarlılık (Recall/Sensitivity):**  $TP/(TP + FN)$

- **Özgüllük (Specificity):**  $TN/(TN + FP)$
- **F1 skoru:**  $2 \cdot P \cdot R/(P + R)$
- **Cohen's Kappa:** şans üstü uyumun düzeltilmiş ölçüsü
- **Matthews Correlation Coefficient (MCC):** dengeli sınıflarda doğruluk için en sağlam metrik

## 4.2 MerQur'da uygulama

*İstatistik* → *Sınıflandırma* → *Karmaşıklık Matrisi Metrikleri*. Form: **Gerçek + Tahmin**.

Çıktıda 2x2 veya KxK confusion matrisi (renkli ısı haritasıyla), tüm metrikler tablo halinde, sınıf-spesifik kesinlik/duyarlılık (multi-class için), **macro / micro / weighted ortalama F1**.

## 4.3 Uygulama örneği

Dengesiz veri seti (95% negative, 5% positive). Doğruluk = 0.93 yanıltıcı görünüyor; ama MCC = 0.18, F1 = 0.22 — gerçek performans zayıf.

# 5. Random Forest Sınıflandırma

## 5.1 Yöntem

Random Forest, çok sayıda karar ağacının "ensemble" (topluluk) yöntemidir (Breiman, 2001). Her ağaç:

1. Bootstrap örneklemeden eğitilir (bagging)
2. Her bölünmede tüm değişkenler değil, **rastgele alt küme** değişkenden en iyisi seçilir

Sonuç: yüksek varyanslı bireysel ağaçlar, ortalama ile düşük varyanslı topluluk haline gelir. Aşırı uydurmaya görece dayanıklıdır; değişken önemi (variable importance) doğal olarak çıkarılır.

## 5.2 MerQur'da uygulama

*İstatistik* → *Sınıflandırma* → *Random Forest Sınıflandırma*

. Form alanları:

- **Hedef:** kategorik (sınıflandırma için)
- **Bağımsızlar:** sayısal/kategorik özellikler
- **Ağaç sayısı (n\_estimators):** 100 (varsayılan), 200, 500, 1000
- **Maksimum derinlik:** None / sayı
- **Sınıf ağırlıkları:** balanced / custom / none
- **k-katlı CV:** k = 5 (varsayılan)

Çıktıda CV-doğruluk + diğer metrikler, **değişken önem grafiği**, OOB (Out-of-Bag) hata tahmini, sınıflandırma matrisi.

## 5.3 Uygulama örneği

20 özellikten 8 önemli çıkar; CV-AUC = 0.89, OOB hatası = 0.11. Önemli değişken: yaş ve gelir.

## 6. Support Vector Machine (SVM)

### 6.1 Yöntem

SVM, sınıflar arasındaki en geniş “marjlı” hiperdüzlemi bulur. Doğrusal ayrılabilir olmayan veriler için **kernel hilesi** (kernel trick) kullanılır: veri yüksek-boyutlu uzaya implicit olarak çıkarılır (RBF, polynomial, sigmoid).

Önemli hiperparametreler:  $C$  (yanlış sınıflandırma cezası — büyük  $C$  = sıkı uyum),  $\gamma$  (RBF kernel için bant genişliği).

### 6.2 MerQur’da uygulama

*İstatistik → Sınıflandırma → Support Vector Machine (SVM)*

. Form alanları:

- **Hedef:** kategorik
- **Bağımsızlar:** sayısal (kategorik için one-hot encoding)
- **Kernel:** linear / RBF / polynomial / sigmoid
- **C:** 1.0 (varsayılan), grid CV ile arama opsiyonu
- **$\gamma$ :** scale / auto / sayı
- **k-katlı CV:**  $k = 5$

Çıktıda CV-performans metrikleri, optimum  $C$  ve  $\gamma$  (grid search seçildiyse), confusion matris, decision boundary görselleştirme (2D için).

### 6.3 Uygulama örneği

Kanser tanı veri seti, 30 özellik. SVM-RBF,  $C=1$ ,  $\gamma=auto$ : CV-doğruluk = 0.92, AUC = 0.96. Grid CV ile  $C=10$  daha iyi sonuç verir.

## 7. Gradient Boosting Sınıflandırma

### 7.1 Yöntem

Gradient Boosting, ardışık olarak öncekinin hatalarını düzeltmeye odaklanan ağaçların ensemble’idir (Friedman, 2001). Her yeni ağaç, mevcut ensemble’ın kalıntılarına (gradient) uydurulur. **LightGBM** veya **XGBoost** benzeri implementasyonlar yüksek hız ve performans sağlar.

Önemli hiperparametreler: öğrenme oranı (learning rate), ağaç derinliği, ağaç sayısı, alt-örneklem oranı, erken durdurma (early stopping).

### 7.2 MerQur’da uygulama

*İstatistik → Sınıflandırma → Gradient Boosting Sınıflandırma*

. Form alanları:

- **Hedef:** kategorik

- **Bağımsızlar:** sayısal/kategorik
- **Öğrenme oranı ( $\eta$ ):** 0.1 (varsayılan)
- **n\_estimators:** 100 (varsayılan)
- **max\_depth:** 3-7 (varsayılan 3)
- **subsample:** 0.8
- **Erken durdurma:** validation loss durdurma turu
- **k-katlı CV:** k = 5

Çıktıda CV-metrikler, **değişken önemi** (gain-based), öğrenme eğrisi (train-val loss), confusion matrix, ROC ve PR eğrileri.

### 7.3 Uygulama örneği

Bir e-ticaret terk tahmini: AUC = 0.91, F1 = 0.74. Erken durdurma 150 ağaçta gerçekleşti; en önemli özellik son satın alma günü.

## 8. Hangi Sınıflandırıcı Ne Zaman? Bir Karar Matrisi

Tablo 1 ana karar eksenlerini özetler.

**Tablo 1.** Sınıflandırma yöntemi seçim rehberi.

Senaryo	Önerilen yöntem
İkili yanıt + az değişken + yorumlama önemli	Lojistik Regresyon (Regresyon kategorisinde)
Çok değişken + doğrusal olmayan ilişki + yorumlama orta	Random Forest
Yüksek tahmin performansı + biraz fazla hesaplama tolere edilir	Gradient Boosting
Boyut yüksek ( $p > n$ ) + ikili yanıt + lineer ayırım	Lineer SVM
Karmaşık geometri + orta boyutlu veri	RBF-SVM
Eşik-bağımsız performans değerlendirme	ROC + AUC
Tür dağılım modellemesi (dengesiz sınıflar)	TSS
Çoklu sınıf + dengesiz	F1 (macro) + MCC + confusion matrix

## 9. Karşılaştırmalı Değerlendirme

Tablo 2, MerQur'un Sınıflandırma kategorisindeki 6 analizin diğer GUI tabanlı açık erişimli alternatiflerle karşılaştırmasını özetler.

**Tablo 2.** Sınıflandırma ailesinin MerQur, JASP, jamovi ve PSPP'deki desteği.

Analiz	MerQur	JASP	jamovi	PSPP
ROC Eğrisi (AUC)	✓	✓	+	-
TSS	✓	-	-	-
Confusion Matris Metrikleri	✓	+	+	-
Random Forest	✓	+ (modül)	+ (modül)	-
SVM	✓	+ (modül)	+ (modül)	-
Gradient Boosting	✓	-	-	-

MerQur'un belirgin avantajları: (i) Gradient Boosting'in doğrudan panel olarak yer alması (çoğu alternatifte yok), (ii) TSS gibi ekoloji/SDM odaklı metriklerin built-in olması, (iii) ROC, TSS, confusion matrix metriklerinin tek arayüzde toplanması.

**Sınırlılıklar:** (i) Neural network/deep learning panelleri henüz yoktur (yol haritasında), (ii) Naive Bayes panel olarak yer almaz (Lojistik regresyon kısmen alternatif), (iii) k-Nearest Neighbors (kNN) panel olarak yer almaz.

## 10. Sonuç

Bu çalışma, MerQur masaüstü yazılımının Sınıflandırma kategorisinde sunulan 6 analizi sistemli olarak tanıtmıştır. ROC ve TSS gibi eşik-temelli performans araçlarından modern ensemble yöntemlerine (Random Forest, Gradient Boosting) ve SVM'e uzanan kapsam, akademik araştırmalardaki ikili ve çok sınıflı sınıflandırma görevlerini geniş bir yelpazede karşılar. k-katlı çapraz doğrulama tüm ML panellerinde standart olarak sunulur; performans metrikleri APA 7'ye yakın formatta otomatik raporlanır. MerQur'un Sınıflandırma kategorisi özellikle tür dağılım modellemesi, klinik tanılama, kredi/finans modelleme ve genel sınıflandırma uygulamaları için bütünlük bir araç ortamıdır. Sonraki davetli editöryal sunumda **Kümeleme** kategorisi ayrıntılı olarak incelenecektir.

## Beyanlar

**Etik Kurul Onayı:** Bu çalışma insan ya da hayvan denek içermediğinden etik kurul onayı gerektirmemiştir.

**Çıkar Çatışması:** Yazar, MerQur yazılımının geliştiricisidir.

**Finansman:** Spesifik bir dış fon alınmamıştır.

**Veri ve Kod Erişim Beyanı:** Bu derleme orijinal araştırma verisi içermez. MerQur yazılımı <https://merqur.sdu.edu.tr> adresinden ücretsiz indirilebilir.

**Yapay Zekâ Kullanımı:** Bu makalenin yazımı sırasında üretken yapay zekâ (Claude, Anthropic) dil ve yapı düzeltmesi amacıyla destekleyici olarak kullanılmıştır.

**Yazar Katkı Beyanı (CRediT):** Ömer K. Örüçü — Kavramsallaştırma, Yöntem, Yazılım, Doğrulama, Yazma (orijinal taslak), Yazma (gözden geçirme & düzenleme).

## Kaynakça

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS).

---

*Journal of Applied Ecology*

, 43(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>

Breiman, L. (2001). Random forests.

---

*Machine Learning*

, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In

---

*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*

(pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation.

---

*BMC Genomics*

, 21, 6. <https://doi.org/10.1186/s12864-019-6413-7>

Cortes, C., & Vapnik, V. (1995). Support-vector networks.

---

*Machine Learning*

, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach.

---

*Biometrics*

, 44(3), 837–845. <https://doi.org/10.2307/2531595>

Fawcett, T. (2006). An introduction to ROC analysis.

---

*Pattern Recognition Letters*

, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine.

---

*Annals of Statistics*

, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>

Hastie, T., Tibshirani, R., & Friedman, J. (2009).

---

*The elements of statistical learning*

(2nd ed.). Springer.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013).

---

*Applied logistic regression*

(3rd ed.). Wiley.

---

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree.

---

*Advances in Neural Information Processing Systems*  
, 30, 3146–3154.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest.

---

*R News*  
, 2(3), 18–22.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme.

---

*Biochimica et Biophysica Acta*  
, 405(2), 442–451.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python.

---

*Journal of Machine Learning Research*  
, 12, 2825–2830.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets.

---

*PLOS ONE*  
, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>

Youden, W. J. (1950). Index for rating diagnostic tests.

---

*Cancer*  
, 3(1), 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)

---

*Bu makale "Davetli Editöryal Sunum" bölümü kapsamında yayımlanmıştır. Bölüm politikası gereği harici hakem değerlendirmesinden geçmemiş, MerQur Veri Bilimi ve Yöntemleri Dergisi Yayın Kurulu tarafından editöryal incelemeye tabi tutulmuştur. Bu makale Creative Commons Atıf 4.0 Uluslararası (CC-BY 4.0) lisansı altında yayımlanmıştır.*