



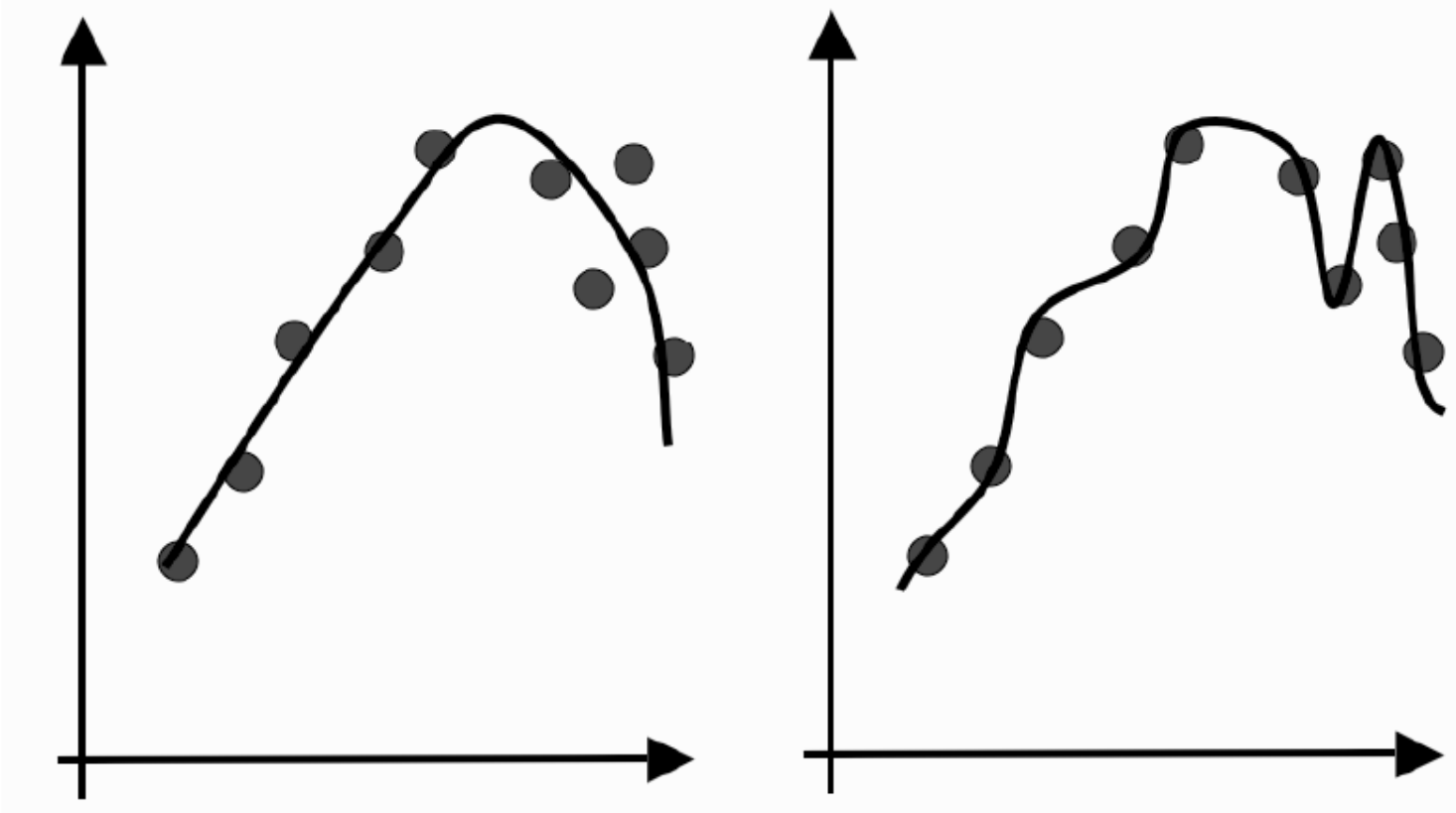
# **Model Performans Metrikleri**

**Doç. Dr. Mehmet Metin ÖZGÜVEN**

**Ankara Üniversitesi Ziraat Fakültesi  
Tarım Makinaları ve Teknolojileri Mühendisliği Bölümü**

**ZTM364 Tarımda Yapay Zeka Kullanımı**

# Ezber veya Aşırı Uyum (Overfitting)



**Şekil. Aşırı uydurmanın etkisi solda gösterildiği gibi üretme fonksiyonunu bulmak yerine, sinir ağının içlerindeki gürültü de dahil olmak üzere (sağda) girdilerle mükemmel bir şekilde eşleşmesidir.**

# Model Performans Metrikleri

## Karışıklık Matrisi (Confusion Matrix)

**Karışıklık matrisi bir sınıflandırma işlemleri için kurulan modellerin performans değerlendirmesinde kullanılabilen bir ölçüttür. Karşıklık matrisi, hem ikili sınıflandırmada hem de çok sınıflı sınıflandırma problemlerinde yaygın kullanılmaktadır. Tahmin edilen ve gerçek değerlerin 4 farklı kombinasyonundan oluşmaktadır. Şekilde karşıklık matrisi örneği verilmiştir.**

# Model Performans Metrikleri

Tablonun bir boyutunda, matris gerçek değerleri almaktadır. Matris daha sonra gerçek değerleri diğer boyuttaki tahmin edilen değerlerle eşlemektedir. Matristen Duyarlılık, Kesinlik, Özgüllük, Doğruluk ve en önemlisi AUC-ROC eğrileri gibi hesaplanabilen birçok performans metriği vardır.

		Gerçek Değer	
		Pozitif (1)	Negatif (0)
Tahmin Edilen Değer	Pozitif (1)	TP	FP
	Negatif (0)	FN	TN

Şekil. Karışıklık matrisi (Altaş, 2023).

# Model Performans Metrikleri

- **Doğru pozitif (True Positive-TP);** sınıflandırıcı tarafından pozitif sınıfa ait verilerden kaç tanesinin doğru şekilde sınıflandırıldığı temsil edilmektedir.
- **Doğru negatif (True Negative-TN);** sınıflandırıcı tarafından negatif sınıfa ait verilerden kaç tanesinin doğru şekilde sınıflandırıldığı temsil edilmektedir.
- **Yanlış pozitif (False Positive (Tip 1 Hata)-FP);** gerçekte negatif sınıfa ait olan bir verinin sınıflandırma sonucunda pozitif sınıf olarak etiketlenmesidir.
- **Yanlış negatif (False Negative (Tip 2 Hata)-FN);** gerçekte pozitif sınıfa ait olan bir verinin sınıflandırma sonucunda negatif sınıf olarak etiketlenmesidir.

# Model Performans Metrikleri

## Kesinlik (Precision)

**Kesinlik, modelin yaptığı pozitif tahminlerin doğruluğunu ölçmek için önemli bir metriktir. Özellikle, modelin gerçek nesnelere yanlış pozitiflerden ne kadar iyi ayırdığını değerlendirmektedir. Temelde kesinlik, modelin doğru pozitif tahminler yapma yeteneğine odaklanmakta ve gerçekten doğru pozitif olan tahminlerin oranını belirtmektedir. Yüksek bir kesinlik skoru, modelin yanlış pozitifleri en aza indirmeye ve güvenilir pozitif tahminler sunma yeteneğine sahip olduğunu göstermektedir. Kesinlik (Precision) Eşitlikteki gibi hesaplanmaktadır.**

$$\text{Kesinlik (Precision)} = \frac{TP}{(TP+FP)}$$

# Model Performans Metrikleri

## Duyarlılık (Recall)

**Duyarlılık, aynı zamanda kesinlik veya doğru pozitif oranı olarak da bilinen, model performansını değerlendirmede, özellikle de nesne tespiti görevlerinde kullanılan önemli bir metriktir. Duyarlılık, modelin resimdeki tüm ilgili nesnelere yakalama yeteneğini ölçmektedir. Yüksek bir duyarlılık skoru, modelin verideki çoğu ilgili nesneyi etkili bir şekilde tanımladığını göstermektedir. Duyarlılık (Recall), Eşitlikteki gibi hesaplanmaktadır.**

$$\text{Duyarlılık (Recall)} = \frac{TP}{(TP+FN)}$$

# Model Performans Metrikleri

## F1 Skoru (F1 Score)

**F1 Skoru (F1 Score), modelin kesinlik (precision) ve duyarlılık (recall) değerlerinin harmonik ortalamasıdır. Bu metrik, modelin performansını değerlendirirken hem yanlış pozitifleri hem de yanlış negatifleri dikkate almaktadır F1 Skoru, modelin doğruluğunu değerlendirmek ve performansını daha kapsamlı bir şekilde anlamak için kullanılmaktadır. F1 Skor Eşitlikteki gibi hesaplanmaktadır.**

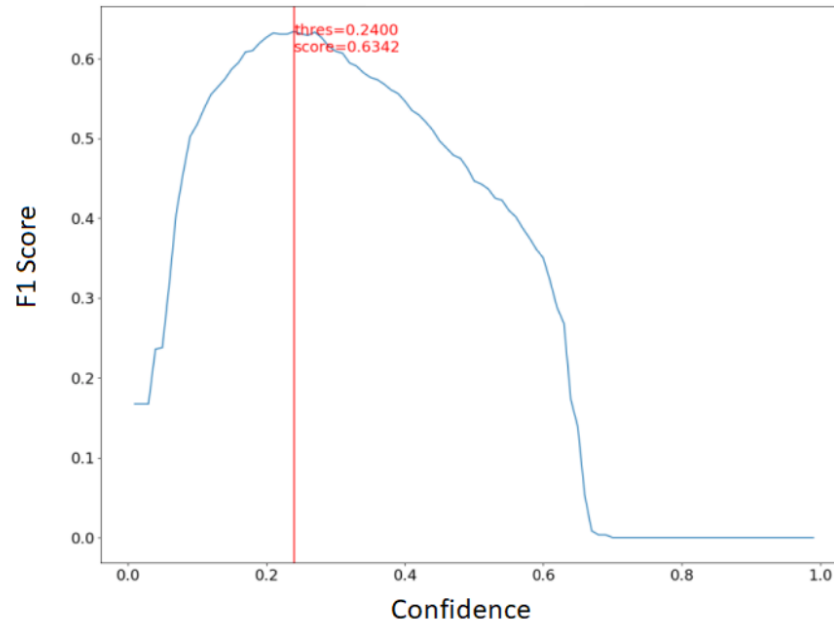
$$\text{F1 Skor} = \frac{2 * \text{Duyarlılık} * \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}}$$



# Model Performans Metrikleri

## F1 Güven Skoru Eşığı Eğrisi

**F1 güven skoru eşığı eğrisi, bir sınıflandırma modelinin çeşitli güven eşığı değerlerine göre kesinlik ve duyarlılık değerlerinin nasıl değiştiğini gösteren bir eğridir. Bir modelin çıkışı genellikle bir güven skoru (confidence score) içermekte ve bu skor, modelin tahmin ettiği bir örneğin ne kadar doğru olma olasılığına işaret etmektedir.**



**Şekil. F1 güven skoru eşığı eğrisi (Uygun, 2023)**

# Model Performans Metrikleri

## Ortalama Kesinliklerin Ortalaması (mAP- mean Average Precision)

**Ortalama Kesinliklerin Ortalaması (mAP), tespit modelinin farklı sınıflar için hesaplanan Ortalama Kesinlik (AP)'lerin ortalamasıdır. mAP metriği aşağıda verilen Eşitlikteki gibi hesaplanmaktadır. Burada N, toplam sınıf sayısını ifade etmektedir. AP<sub>i</sub>, her bir sınıfın hesaplanmış Ortalama Kesinlik (AP) değerini temsil etmektedir. Bu değer nesne tespiti modellerinin farklı kesinlik seviyelerinde, ne kadar etkili olduğunu değerlendirmek için kullanılmaktadır.**

$$\text{Ortalama Kesinliklerin Ortalaması (mAP)} = \frac{1}{N} \sum_{i=1}^N AP_i$$

# Model Performans Metrikleri

## **mAP@0.5 ile mAP@0.5:0.95 Metrikleri**

**mAP@0.5 ile mAP@0.5:0.95 metrikleri ise Eşitlikteki gibi hesaplanmaktadır. Burada  $\text{ort}(mAP_i)$ , 0.5 ila 0.95 aralığındaki farklı IoU değerlerinde elde edilen mAP değerlerinin ortalamasını ifade etmektedir. Bu, genellikle nesne tespiti modellerinin farklı kesinlik eşiklerindeki performansını anlamak için kullanılmaktadır.**

$$mAP@0.5:0.95 = \text{ort}(mAP_i), \quad i = 0.5:0.05:0.95$$

# Performans Metriklerine Örnekler

Table 2 Confusion matrix of the Faster RCNN model

		Predict				Accuracy	Precision	Recall	F1 Score
		0 (Healthy)	1 (Low)	2 (Severe)	3 (Low & Severe)				
Actual	0 Healthy	70	5	1	0	96.15	0.92	0.92	0.92
	1 Low	5	66	3	0	93.91	0.86	0.89	0.87
	2 Severe	1	5	77	3	94.23	0.90	0.90	0.90
	3 Low & Severe	0	1	5	70	97.12	0.96	0.92	0.94
Overall Accuracy						90.71			

# Performans Metrikleri Örnekler

**Table 4** Performance comparisons of YOLOv8(n/s/m/l/x)-Seg models

Model	Parameter	Layer	Train time	GFLOP <sub>s</sub>	Tesla A100 max. GPU usage (G)	Mask			
						<u>mAP</u> 0.5	<u>mAP</u> 0.5-0.95	P	R
YOLOv8n-Seg	3M	195	27m, 37s	12.0	7.5	0.934	0.795	0.903	0.899
YOLOv8s-Seg	11M	195	29m, 50s	42.4	8.7	0.929	0.797	0.926	0.885
YOLOv8m-Seg	27M	245	35m, 19s	110.0	12	0.930	0.8	0.896	0.895
YOLOv8l-Seg	46M	295	42m, 32s	220.1	16.1	0.935	0.806	0.956	0.859
YOLOv8x-Seg	71M	295	54m, 40s	343.7	20.9	0.931	0.797	0.872	0.912



Doç. Dr. Mehmet Metin ÖZGÜVEN

[mmozguven@ankara.edu.tr](mailto:mmozguven@ankara.edu.tr)

ZTM364 Tarımda Yapay Zeka Kullanımı