Introduction

**Model kartı**, bir makine öğrenimi modeli hakkında temel bilgileri sağlayan kısa bir belgedir. Model kartları, eğitilmiş modeller hakkındaki bilgileri geniş kitlelere ileterek şeffaflığı artırır.

Bu eğitimde, model kartlarının hangi kitleler için yazılacağını ve hangi bölümleri içermesi gerektiğini öğreneceksiniz. Ardından, bir sonraki **alıştırmada**, öğrendiklerinizi gerçek dünyadan birkaç senaryoya uygulayacaksınız.

# Model cards[¶](https://www.kaggle.com/code/var0101/model-cards" \l "Model-cards" \t "_self)

Yapay zekâ sistemleri her sektörde giderek daha önemli roller üstlenmesine rağmen, çok az kişi bu sistemlerin nasıl çalıştığını anlıyor. Yapay zekâ araştırmacıları, yapay zekâ sistemlerini kullanan, yapay zekâ sistemlerinden etkilenen ve diğer ilgili kişileri bilgilendirmek için modeller hakkındaki temel bilgileri iletmenin birçok yolunu araştırıyor.

**Model kartları** - 2019 tarihli bir makalede tanıtılan - ekiplerin yapay zekâ sistemleri hakkındaki temel bilgileri geniş bir kitleye iletmesinin bir yoludur. Bu bilgiler genellikle modelin amaçlanan kullanımlarını, modelin nasıl çalıştığını ve farklı durumlarda nasıl performans gösterdiğini içerir.

Model kartlarını, paketli gıdalarda bulduğunuz besin değerleri etiketlerine benzetebilirsiniz.

# Examples of model cards[¶](https://www.kaggle.com/code/var0101/model-cards" \l "Examples-of-model-cards" \t "_self)

Devam etmeden önce, model kartlara ait bazı örneklere kısaca göz atmak faydalı olabilir.

* [Salesforce's model cards](https://blog.einstein.ai/model-cards-for-ai-model-transparency/)
* [Open AI’s model card for GPT-3](https://github.com/openai/gpt-3/blob/master/model-card.md)
* [Google Cloud's example model cards](https://modelcards.withgoogle.com/face-detection)

# Who is the audience of your model card?[¶](https://www.kaggle.com/code/var0101/model-cards#Who-is-the-audience-of-your-model-card?)

Bir model kartı, kolay anlaşılır olmak ile önemli teknik bilgileri iletmek arasında bir denge kurmalıdır. Bir model kartı yazarken, hedef kitlenizi, yani model kartınızı okuma olasılığı en yüksek olan insan gruplarını göz önünde bulundurmalısınız. Bu gruplar, yapay zekâ sisteminin amacına göre değişecektir.

Örneğin, kas-iskelet sistemi yaralanmalarını daha iyi teşhis etmek için tıp uzmanlarının röntgenleri yorumlamasına yardımcı olan bir yapay zekâ sistemi için hazırlanan model kartı, muhtemelen tıp uzmanları, bilim insanları, hastalar, araştırmacılar, politika yapıcılar ve benzer yapay zekâ sistemlerinin geliştiricileri tarafından okunacaktır. Bu nedenle, model kartı sağlık hizmetleri ve yapay zekâ sistemleri hakkında bazı bilgilere sahip olunduğunu varsayabilir.

# What sections should a model card contain?[¶](https://www.kaggle.com/code/var0101/model-cards#What-sections-should-a-model-card-contain?)

Orijinal makaleye göre, bir model kart aşağıdaki dokuz bölümden oluşmalıdır. Farklı kuruluşların ihtiyaçlarına göre model kart bölümlerini ekleyebileceğini, çıkarabileceğini veya yeniden düzenleyebileceğini unutmayın (ve bunu yukarıdaki örneklerin bazılarında fark etmiş olabilirsiniz).

Farklı bölümler hakkında okurken, orijinal makaleden iki örnek model kartı incelemeniz önerilir. Devam etmeden önce, bu model kart örneklerinin her birini yeni bir pencerede açın:

* [Model Card - Smiling Detection in Images](https://github.com/Kaggle/learntools/blob/master/notebooks/ethics/pdfs/smiling_in_images_model_card.pdf)
* [Model Card - Toxicity in Text](https://github.com/Kaggle/learntools/blob/master/notebooks/ethics/pdfs/toxicity_in_text_model_card.pdf)

## **1. Model Details[¶](https://www.kaggle.com/code/var0101/model-cards" \l "1.-Model-Details" \t "_self)**

Geliştirici ve model sürümü gibi arka plan bilgilerini ekleyin.

## **2. Intended Use[¶](https://www.kaggle.com/code/var0101/model-cards" \l "2.-Intended-Use" \t "_self)**

Hangi kullanım durumları kapsam dahilinde?

Hedef kullanıcılarınız kimler?

Hangi kullanım durumları kapsam dışında?

## **3. Factors[¶](https://www.kaggle.com/code/var0101/model-cards" \l "3.-Factors" \t "_self)**

Modelin etkisini hangi faktörler etkiler? Örneğin, gülümseme algılama modelinin sonuçları yaş, cinsiyet veya etnik köken gibi demografik faktörlere, aydınlatma veya yağmur gibi çevresel faktörlere ve kamera türü gibi cihazlara göre değişiklik gösterir.

## **4. Metrics[¶](https://www.kaggle.com/code/var0101/model-cards" \l "4.-Metrics" \t "_self)**

"Modelin performansını ölçmek için hangi metrikleri kullanıyorsunuz? Neden bu metrikleri seçtiniz?"

**Sınıflandırma sistemleri** için – çıktının bir sınıf etiketi olduğu sistemler – potansiyel hata türleri arasında yanlış pozitif oranı, yanlış negatif oranı, yanlış buluş oranı ve yanlış atlama oranı bulunur. Bunların her birinin göreli önemi, kullanım durumuna bağlıdır.

**Puana dayalı analizler** için – çıktının bir puan veya fiyat olduğu analizler – gruplar arasındaki model performansını raporlamayı düşünün.

## **5. Evaluation Data**[**¶**](https://www.kaggle.com/code/var0101/model-cards#5.-Evaluation-Data)

Model performansını değerlendirmek için hangi veri kümelerini kullandınız? Mümkünse veri kümelerini belirtin.

Değerlendirme için neden bu veri kümelerini seçtiniz?

Veri kümeleri tipik kullanım durumlarını, beklenen test durumlarını ve/veya zorlu durumları temsil ediyor mu?

## **6. Training Data**[**¶**](https://www.kaggle.com/code/var0101/model-cards#6.-Training-Data)

Model hangi veriler üzerinde eğitildi?

## **7. Quantitative Analyses[¶](https://www.kaggle.com/code/var0101/model-cards" \l "7.-Quantitative-Analyses" \t "_self)**

Model, seçtiğiniz metriklerde nasıl performans gösterdi? Performansı önemli faktörlere ve bunların kesişim noktalarına göre ayırın. Örneğin, gülümseme tespiti örneğinde performans yaşa (örneğin genç, yaşlı), cinsiyete (örneğin kadın, erkek) ve ardından her ikisine (örneğin yaşlı-kadın, yaşlı-erkek, genç-kadın, genç-erkek) göre ayrıştırılır.

## **8. Ethical Considerations[¶](https://www.kaggle.com/code/var0101/model-cards" \l "8.-Ethical-Considerations" \t "_self)**

Modelin eğitilmesinde kullanılan hassas veriler, modelin insan yaşamı, sağlığı veya güvenliği üzerinde etkileri olup olmadığı, riskin nasıl azaltıldığı ve model kullanımında hangi zararların bulunabileceği gibi modelle ilgili etik hususları açıklayın.

## **9. Caveats and Recommendations[¶](https://www.kaggle.com/code/var0101/model-cards" \l "9.-Caveats-and-Recommendations" \t "_self)**

Model kartında başka bir yerde değinmediğiniz önemli bir şeyi ekleyin.

# How can you use model cards in your organization?[¶](https://www.kaggle.com/code/var0101/model-cards#How-can-you-use-model-cards-in-your-organization?)

Ayrıntılı model kartlarının kullanımı, bir kuruluşun süreçlerini, tescilli verilerini veya ticari sırlarını ifşa etmek istememesi nedeniyle genellikle zorlayıcı olabilir. Bu gibi durumlarda, geliştirici ekibi, hassas bilgiler içermeden model kartlarının nasıl faydalı ve güçlendirici olabileceğini düşünmelidir.

Bazı ekipler, ML model bilgilerini toplamak ve kaydetmek için Bilgi Sayfaları gibi farklı formatlar kullanır.