Model başlangıçta modelin parametrelerini (örneğin, yapay sinir ağlarındaki nöronlar arasındaki bağlantıların ağırlıkları) uydurmak [için](https://en.wikipedia.org/wiki/Artificial_neural_networks) kullanılan bir dizi örnek olan bir **eğitim veri setine**[[3]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-James_2013_176-3) uyarlanır . [[4]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Ripley_1996_354-4) Model (örneğin [saf bir Bayes sınıflandırıcı ),](https://en.wikipedia.org/wiki/Naive_Bayes_classifier)[denetimli bir öğrenme](https://en.wikipedia.org/wiki/Supervised_learning) yöntemi kullanılarak, örneğin [gradyan iniş](https://en.wikipedia.org/wiki/Gradient_descent) veya [stokastik gradyan iniş](https://en.wikipedia.org/wiki/Stochastic_gradient_descent) gibi optimizasyon yöntemleri kullanılarak eğitim veri seti üzerinde eğitilir . [Uygulamada, eğitim veri seti genellikle bir giriş vektörü](https://en.wikipedia.org/wiki/Array_data_structure) (veya skaler) ve karşılık gelen çıkış vektörü (veya skaler) çiftlerinden oluşur ; burada cevap anahtarı genellikle *hedef* (veya *etiket* ) olarak gösterilir . Mevcut model, eğitim veri seti ile çalıştırılır ve eğitim veri setindeki her giriş vektörü için daha sonra *hedefle* karşılaştırılan bir sonuç üretir . Karşılaştırmanın sonucuna ve kullanılan spesifik öğrenme algoritmasına bağlı olarak modelin parametreleri ayarlanır. [Model uyumu hem değişken seçimini](https://en.wikipedia.org/wiki/Feature_selection) hem de parametre [tahminini](https://en.wikipedia.org/wiki/Estimation_theory) içerebilir .

Uygun model, **doğrulama veri seti** adı verilen ikinci bir veri setindeki gözlemlere yönelik yanıtları tahmin etmek için sırasıyla kullanılır . [[3]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-James_2013_176-3) Doğrulama veri seti, modelin [hiperparametrelerini](https://en.wikipedia.org/wiki/Hyperparameter_(machine_learning)" \o "Hiperparametre (makine öğrenimi))[[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5) (örneğin bir sinir ağındaki gizli birimlerin sayısı - katmanlar ve katman genişlikleri - [[4]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Ripley_1996_354-4) ) ayarlarken, eğitim veri setine uyan bir modelin tarafsız bir değerlendirmesini sağlar. [Doğrulama veri kümeleri, erken durdurma](https://en.wikipedia.org/wiki/Early_stopping) yoluyla [düzenleme](https://en.wikipedia.org/wiki/Regularization_(mathematics)) için kullanılabilir (doğrulama veri kümesindeki hata arttığında eğitimi durdurmak, çünkü bu, eğitim veri kümesine [aşırı uyumun bir işaretidir).](https://en.wikipedia.org/wiki/Overfitting)[[6]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-prechelt_early_2012-6) Bu basit prosedür, doğrulama veri kümesinin hatasının eğitim sırasında dalgalanabileceği ve birden fazla yerel minimum üretebileceği gerçeği nedeniyle pratikte karmaşık hale gelir. Bu karmaşıklık, aşırı uyumun gerçekten ne zaman başlayacağına karar vermek için birçok geçici kuralın oluşturulmasına yol açmıştır. [[6]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-prechelt_early_2012-6)

Son olarak **test veri seti** , eğitim veri setine uyan *nihai* modelin tarafsız bir değerlendirmesini sağlamak için kullanılan bir veri setidir . [[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5)[Test veri setindeki veriler eğitimde (örneğin çapraz doğrulamada](https://en.wikipedia.org/wiki/Cross-validation_(statistics)) ) hiç kullanılmamışsa , test veri setine aynı zamanda **bekletme veri seti** de denir . Bazı literatürde bazen "test seti" yerine "doğrulama seti" terimi kullanılır (örneğin, orijinal veri seti yalnızca iki alt gruba bölünmüşse, test seti doğrulama seti olarak adlandırılabilir). [[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5)

Eğitim, test ve doğrulama setlerinde veri seti bölümü için boyutlara ve stratejilere karar vermek, büyük ölçüde soruna ve mevcut verilere bağlıdır

Eğitim veri seti [ [düzenle](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=1) ]

[daire, kalıp, desen, düzen, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu](https://en.wikipedia.org/wiki/File:Simplified_neural_network_training_example.svg)

Nesne tespitinde bir sinir ağının eğitiminin basitleştirilmiş örneği: Ağ, [denizyıldızlarını](https://en.wikipedia.org/wiki/Starfish) ve [deniz kestanelerini tasvir ettiği bilinen ve görsel](https://en.wikipedia.org/wiki/Sea_urchin)[özellikleri](https://en.wikipedia.org/wiki/Feature_(computer_vision)) temsil eden "düğümler" ile ilişkilendirilen birden fazla görüntü ile eğitilir . Denizyıldızı halkalı bir doku ve yıldız hatlarıyla eşleşirken deniz kestanelerinin çoğu çizgili doku ve oval şekille eşleşir. Ancak halka dokulu deniz kestanesi örneği aralarında zayıf ağırlıklı bir ilişki yaratıyor.

[daire, çizim, çocukların yaptığı resimler, çizgi film içeren bir resim

Açıklama otomatik olarak oluşturuldu](https://en.wikipedia.org/wiki/File:Simplified_neural_network_example.svg)

Ağın bir giriş görüntüsü üzerinde daha sonra çalıştırılması (solda): [[8]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-8) Ağ, denizyıldızını doğru bir şekilde algılar. Bununla birlikte, halkalı doku ile deniz kestanesi arasındaki zayıf ağırlıklı ilişki, ikincisine iki ara düğümden birinden gelen zayıf bir sinyal de vermektedir. Ayrıca eğitime dahil edilmeyen bir kabuk oval şekil için zayıf sinyal verirken, deniz kestanesi çıkışı için de zayıf sinyal oluşmasına neden olur. Bu zayıf sinyaller deniz kestanesi için [yanlış pozitif](https://en.wikipedia.org/wiki/False_positive) sonuca neden olabilir .  
Gerçekte, dokular ve ana hatlar tek düğümlerle değil, birden çok düğümün ilişkili ağırlık desenleriyle temsil edilir.

Bir eğitim veri seti, öğrenme süreci sırasında kullanılan örneklerden oluşan bir [veri setidir ve örneğin bir](https://en.wikipedia.org/wiki/Dataset)[sınıflandırıcının](https://en.wikipedia.org/wiki/Classifier_(machine_learning)) parametrelerine (örneğin ağırlıklara) uymak için kullanılır . [[9]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Ripley,_B.D._1996_p._354-9)[[10]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-cann-faq-10)

[Sınıflandırma görevleri için denetimli bir öğrenme algoritması, iyi bir tahmin modeli](https://en.wikipedia.org/wiki/Predictive_modelling) oluşturacak optimal değişken kombinasyonlarını belirlemek veya öğrenmek için eğitim veri setine bakar . [[11]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Larose2014-11) Amaç, yeni ve bilinmeyen verilere iyi bir şekilde genelleme yapan, eğitilmiş (uygun) bir model üretmektir. [[12]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Xu_Goodacre_2018-12) Uygun model, modelin yeni verileri sınıflandırmadaki doğruluğunu tahmin etmek için uzatılmış veri kümelerinden (doğrulama ve test veri kümeleri) "yeni" örnekler kullanılarak değerlendirilir. [[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5) Aşırı uyum gibi sorunların riskini azaltmak için doğrulama ve test veri kümelerindeki örnekler, modeli eğitmek için kullanılmamalıdır. [[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5)

Deneysel ilişkiler için eğitim verilerini araştıran yaklaşımların çoğu, verilere [gereğinden fazla uyum sağlama](https://en.wikipedia.org/wiki/Overfit) eğilimindedir ; bu, eğitim verilerinde genel olarak geçerli olmayan görünür ilişkileri tanımlayabilecekleri ve bunlardan yararlanabilecekleri anlamına gelir.

Doğrulama veri seti [ [düzenle](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=2) ]

Doğrulama veri seti, bir sınıflandırıcının [hiperparametrelerini](https://en.wikipedia.org/wiki/Hyperparameter_(machine_learning)" \o "Hiperparametre (makine öğrenimi)) (yani mimarisini) ayarlamak için kullanılan örneklerden oluşan bir [veri setidir .](https://en.wikipedia.org/wiki/Dataset)Bazen geliştirme seti veya "geliştirme seti" olarak da adlandırılır. [[13]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-13)[Yapay sinir ağlarına](https://en.wikipedia.org/wiki/Artificial_neural_networks) yönelik bir hiperparametre örneği, her katmandaki gizli birimlerin sayısını içerir. [[9]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Ripley,_B.D._1996_p._354-9)[[10]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-cann-faq-10) Test seti (aşağıda belirtildiği gibi) gibi, eğitim veri seti ile aynı olasılık dağılımını takip etmelidir.

Aşırı uyumdan kaçınmak için herhangi bir [sınıflandırma](https://en.wikipedia.org/wiki/Statistical_classification) parametresinin ayarlanması gerektiğinde eğitim ve test veri setlerine ek olarak bir doğrulama veri setine de sahip olmak gerekir. Örneğin, problem için en uygun sınıflandırıcı aranıyorsa, farklı aday sınıflandırıcıları eğitmek için eğitim veri seti kullanılır, performanslarını karşılaştırmak ve hangisinin alınacağına karar vermek için doğrulama veri seti kullanılır ve son olarak test verileri kullanılır. set [doğruluk](https://en.wikipedia.org/wiki/Accuracy) , [duyarlılık](https://en.wikipedia.org/wiki/Sensitivity_and_specificity) , [özgüllük](https://en.wikipedia.org/wiki/Sensitivity_and_specificity) , [F-ölçüsü](https://en.wikipedia.org/wiki/Precision_and_recall#F-measure) vb. gibi performans özelliklerini elde etmek için kullanılır . Doğrulama veri seti bir hibrit olarak işlev görür: test için kullanılan eğitim verileridir, ancak ne düşük seviyeli eğitimin bir parçası ne de son testin bir parçası değildir.

[Model seçimi](https://en.wikipedia.org/wiki/Model_selection) için bir doğrulama veri seti kullanmanın temel süreci (eğitim veri seti, doğrulama veri seti ve test veri setinin bir parçası olarak): [[10]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-cann-faq-10)[[14]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-14)

Amacımız yeni veriler üzerinde en iyi performansa sahip ağı bulmak olduğundan, farklı ağların karşılaştırılmasında en basit yaklaşım, eğitim için kullanılandan bağımsız verileri kullanarak hata fonksiyonunu değerlendirmektir. Çeşitli ağlar, bir eğitim veri setine göre tanımlanan uygun bir hata fonksiyonunun en aza indirilmesiyle eğitilir. Daha sonra ağların performansı, bağımsız bir doğrulama seti kullanılarak hata fonksiyonunun değerlendirilmesiyle karşılaştırılır ve doğrulama setine göre en küçük hataya sahip ağ seçilir. *Bu yaklaşıma bekletme* yöntemi denir . Bu prosedürün kendisi doğrulama setine bir miktar aşırı uyum sağlamaya yol açabileceğinden, seçilen ağın performansı, test seti adı verilen üçüncü bir bağımsız veri seti üzerindeki performansı ölçülerek doğrulanmalıdır.

Bu sürecin bir uygulaması , aday modellerin aynı ağın ardışık yinelemeleri olduğu [erken durdurmadır](https://en.wikipedia.org/wiki/Early_stopping) ve doğrulama kümesindeki hata büyüdüğünde önceki modeli (minimum hataya sahip olanı) seçerek eğitim durur.

Test veri seti [ [düzenle](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=3) ]

Test veri seti, eğitim veri setinden [bağımsız](https://en.wikipedia.org/wiki/Independence_(probability_theory)) olan ancak eğitim veri seti ile aynı [olasılık dağılımını takip eden bir](https://en.wikipedia.org/wiki/Probability_distribution)[veri setidir .](https://en.wikipedia.org/wiki/Dataset)Eğitim veri setine uyan bir model aynı zamanda test veri setine de iyi uyuyorsa, minimal düzeyde [aşırı uyum](https://en.wikipedia.org/wiki/Overfitting) gerçekleşmiştir (aşağıdaki şekle bakın). Eğitim veri setinin test veri setine göre daha iyi uyumu genellikle aşırı uyum anlamına gelir.

Bu nedenle bir test seti, yalnızca tamamen belirlenmiş bir sınıflandırıcının performansını (yani genellemeyi) değerlendirmek için kullanılan bir örnekler setidir. [[9]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Ripley,_B.D._1996_p._354-9)[[10]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-cann-faq-10) Bunu yapmak için, test setindeki örneklerin sınıflandırmalarını tahmin etmek için son model kullanılır. Bu tahminler, modelin doğruluğunu değerlendirmek için örneklerin gerçek sınıflandırmalarıyla karşılaştırılır. [[11]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Larose2014-11)

Hem doğrulama hem de test veri setlerinin kullanıldığı bir senaryoda, test veri seti genellikle doğrulama süreci sırasında seçilen son modeli değerlendirmek için kullanılır. Orijinal veri kümesinin iki alt kümeye (eğitim ve test veri kümeleri) bölünmesi durumunda, test veri kümesi modeli yalnızca bir kez değerlendirebilir (örneğin, uzatma yönteminde [)](https://en.wikipedia.org/wiki/Holdout_method) . [[15]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Kohavi2001-15) Bazı kaynakların böyle bir yönteme karşı tavsiyede bulunduğunu unutmayın. [[12]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Xu_Goodacre_2018-12)[Bununla birlikte, çapraz doğrulama](https://en.wikipedia.org/wiki/Cross-validation_(statistics)) gibi bir yöntem kullanıldığında , yanlılığı ve değişkenliği azaltmaya yardımcı olmak için tekrarlanan model eğitimi ve test turlarından sonra sonuçların ortalaması alındığından iki bölüm yeterli ve etkili olabilir. [[5]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Brownlee-5)[[12]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-Xu_Goodacre_2018-12)

[çizgi, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu](https://en.wikipedia.org/wiki/File:Traintest.svg)Aynı istatistiksel popülasyondan bir eğitim seti (solda) ve bir test seti (sağda) mavi noktalarla gösterilmiştir. Eğitim verilerine iki tahmin modeli uygundur. Her iki takılan model de hem eğitim hem de test setleriyle çizilmiştir. Eğitim setinde turuncu ile gösterilen uyumun [MSE'si](https://en.wikipedia.org/wiki/Mean_squared_error" \o "Ortalama kare hatası) 4, yeşil ile gösterilen uyumun MSE'si 9'dur. Test setinde turuncu ile gösterilen uyumun MSE'si 15 ve ile gösterilen uyumun MSE'si 15'tir ve turuncu ile gösterilen uyumun MSE'si 9'dur. yeşil 13'tür. Turuncu eğri eğitim verilerine fazlasıyla uyuyor çünkü test seti ile eğitim seti karşılaştırıldığında MSE'si neredeyse dört kat artıyor. Yeşil eğri, MSE'si 2 kattan daha az arttığı için eğitim verilerine çok daha az uyuyor.

Terminolojide karışıklık [\_](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=4) \_

Test etmek, bir şeyi bulmaya çalışmaktır (Collaborative International Dictionary of English'e göre "Kanıtlamak; doğruluğunu, orijinalliğini veya kalitesini deney yoluyla kanıtlamak") ve doğrulamak, bir şeyin geçerli olduğunu kanıtlamak demektir ( "Onaylamak; geçerli kılmak" İşbirlikçi Uluslararası İngilizce Sözlüğü). **Bu bakış açısıyla test seti** ve **doğrulama seti** terimlerinin en yaygın kullanımı burada açıklanandır. Bununla birlikte, hem endüstride hem de akademide, iç sürecin geliştirmek için farklı modelleri test ettiği (bir geliştirme seti olarak test seti) ve nihai modelin, gerçek kullanımdan önce doğrulanması gereken model olduğu dikkate alınarak, bazen birbirlerinin yerine kullanılırlar. görünmeyen bir veri (doğrulama seti). "Makine öğrenimi literatürü sıklıkla 'doğrulama' ve 'test' kümelerinin anlamını tersine çeviriyor. Bu, yapay zeka araştırmalarına hakim olan terminolojik kafa karışıklığının en bariz örneğidir." [[16]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-16) Bununla birlikte, korunması gereken önemli kavram, ister test ister doğrulama olarak adlandırılsın, son setin yalnızca son deneyde kullanılması gerektiğidir.

Çapraz doğrulama [ [düzenle](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=5) ]

Daha istikrarlı sonuçlar elde etmek ve tüm değerli verileri eğitim için kullanmak amacıyla, bir veri seti tekrar tekrar çeşitli eğitim ve doğrulama veri setlerine bölünebilir. [Bu, çapraz doğrulama](https://en.wikipedia.org/wiki/Cross-validation_(statistics)) olarak bilinir . Modelin performansını doğrulamak için normalde çapraz doğrulamadan elde edilen ek bir test veri seti kullanılır.

Hatanın nedenleri [\_](https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_data_sets&action=edit&section=6) \_

[metin, ekran görüntüsü, çizgi film, giyim içeren bir resim

Açıklama otomatik olarak oluşturuldu](https://en.wikipedia.org/wiki/File:Computer_definition_of_extra_hot.png)[Kurgusal hatalı bir bilgisayar çıktısını gösteren çizgi roman](https://en.wikipedia.org/wiki/Comic_strip) (önceki "ekstra sıcak" tanımından bir kahveyi 5 milyon derece [Santigrat veya Fahrenhayt belirtilmemiş] yapıyor). Bu hem mantık hatası hem de ilgili çeşitli çevresel koşulların dahil edilememesi olarak sınıflandırılabilir. [[17]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-ChandaBanerjee2022-17)

Algoritmaların eğitimindeki ihmaller hatalı çıktıların önemli bir nedenidir. [[17]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-ChandaBanerjee2022-17) Bu tür ihmallerin türleri şunları içerir: [[17]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-ChandaBanerjee2022-17)

* Belirli koşullar veya varyasyonlar dahil edilmemiştir.
* Eski veriler
* Belirsiz giriş bilgileri
* Yeni ortamlara geçememe
* Gerektiğinde bir insandan veya başka bir yapay zeka sisteminden yardım isteyememek

Belirli koşulların ihmal edilmesine bir örnek, bir çocuğun annesinin iç mekan, gece aydınlatması altında yüzünü kaydetmesi nedeniyle telefonun kilidini açabilmesidir; bu durum sistemin eğitimine uygun şekilde dahil edilmemiştir. [[17]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-ChandaBanerjee2022-17)[[18]](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-18)

[Göreceli olarak alakasız girdilerin kullanılması, algoritmaların nesne tespiti](https://en.wikipedia.org/wiki/Object_detection) için ilgilenilen nesne yerine arka planı kullandığı durumları içerebilir ; örneğin otlaklardaki koyun resimleriyle eğitilmek, farklı bir nesnenin bulunması durumunda koyun olarak yorumlanması riskine yol açar. bir çayırda. [[17](https://en.wikipedia.org/wiki/Training,_validation,_and_test_data_sets#cite_note-ChandaBanerjee2022-17)