



Yapay Zeka Destekli Araştırmalarda, Tahminin Yorumlamayı Gölgede Bıraktığı Durumlar: Jeolojide Veri Odaklı Yanılsama
When Accuracy Misleads Geological Interpretation: A Data-Driven Illusion

Oğuz Mülayim^{1,2*}

¹ *Türkiye Petrolleri A.O. Adıyaman Bölge Müdürlüğü, 02040, Adıyaman, Türkiye*

² *Türkiye Jeoloji Bülteni Editör Kurulu Üyesi/Member of the Geological Bulletin of Turkey Editorial Board*

• Geliş/Received: 16.02.2026 • Düzeltmiş Metin Geliş/Revised Manuscript Received: 30.03.2026 • Kabul/Accepted: 31.03.2026
• Çevrimiçi Yayın/Available online: 14.04.2026 • Baskı/Printed: 30.04.2026

Editorial/Editorial

Türkiye Jeol. Bül. / Geol. Bull. Turkey

Öz: Son yıllarda yapay zeka (YZ) modelleri; deprem erken uyarı sistemlerinden heyelan duyarlılık haritalarına, yeraltı kaynaklarının modellenmesinden stratigrafik sınıflandırmaya kadar yer bilimlerinin hemen her alanında rutin olarak kullanılmaktadır. Geleneksel yöntemleri tahmin doğruluğu açısından geride bırakan bu modeller, gözlem ile çıkarım arasında giderek daha fazla arabuluculuk rolü üstlenmektedir. Ancak bu teknik başarı, kritik bir soruyu gündeme getirmektedir: Yüksek tahmin doğruluğu, gerçek jeolojik anlayışı yansıtmakta mıdır? Bu yazı, veri odaklı modellerin artan tahmin kapasitesinin, yerbilimlerinin yorumlayıcı doğasını ve gerçek jeolojik koşulların ortaya konulmasını gölgeleme riskine dikkat çekmektedir. Gözlemlerin seyrek, belirsizliğin yapısal nitelikte ve doğruluğu kanıtlanmış bilginin sınırlı olduğu alanlarda —örneğin yeraltı yorumlaması ve afet değerlendirmesi— modeller, fiziksel süreç temelli olarak ilgisiz proksiler (vekil değişkenler) aracılığıyla görünürde yüksek sınıflandırma doğruluğuna ulaşabilmektedir. Bu durum, değişen çevresel veya tektonik koşullar altında modellerin taşınabilirliğini ve jeolojik tutarlılığını zayıflatmaktadır. Tahmine dayalı modellemeyi reddetmeksizin, bu perspektif çalışma, yeni bir algoritma önermekten ziyade yol gösterici bir yapı niteliğindeki kavramsal ve yarı-biçimsel bir “asistan” çerçevesi sunmayı amaçlamaktadır. Bu çerçeve, jeolojik kısıtlamaları, yorumlanabilir modellemeyi ve model sonrası jeolojik doğrulamayı bütünleştirerek tahmin performansının jeolojik akıl yürütmenin önüne geçmesini engellemeyi hedeflemektedir.

Abstract: In recent years, artificial intelligence (AI) models have become routine in geoscience applications ranging from earthquake early warning systems to landslide susceptibility mapping, subsurface resource modeling, and stratigraphic classification. Outperforming traditional methods for predictive accuracy, these models increasingly mediate between observation and inference. However, this technical success raises a critical question: does high predictive accuracy reflect true geological understanding? This paper draws attention to the risk that the growing predictive capacity of data-driven models may overshadow the interpretive nature of geoscience and the elucidation of actual geological conditions. In domains where observations are sparse, uncertainty is structural, and ground truth is limited—such as subsurface interpretation and hazard assessment—models can achieve seemingly high classification accuracy by relying on mechanistically irrelevant proxies. This undermines model transferability and geological consistency under changing environmental or tectonic conditions. Without rejecting predictive modeling, this perspective aims to propose a conceptual and semi-formalized “assistant” framework— not introducing a new algorithm but rather a guiding structure—that integrates geological constraints, interpretable modeling, and post-hoc geological validation to ensure that predictive performance does not override geological reasoning.

* Yazışma / Correspondence: omulayim@tpao.gov.tr

Anahtar Kelimeler: Açıklanabilir yapay zeka, fizik bilgili makine öğrenmesi, jeolojik yorum, tahmin modelleri, yapay zeka.

Keywords: *Artificial intelligence, explainable AI, geological interpretation, physics-informed machine learning, predictive models.*

Yapay zeka (YZ), yerbilimleri araştırmalarının ayrılmaz bir parçası haline gelmiş ve son yıllarda keşif amaçlı analizlerden karar verme süreçlerinin kritik olduğu durumlardaki operasyonel kullanıma doğru hızla ilerlemektedir. YZ odaklı modeller artık afet değerlendirme, erken uyarı sistemleri ve yeraltı kaynaklarına ilişkin karar alma süreçlerine giderek daha fazla entegre edilmekte; bu modellerin çıktıları politikaları, mühendislik uygulamalarını ve risk yönetimini doğrudan şekillendirmektedir. Bu değişim, giderek büyüyen, heterojen ve karmaşık veri kümelerinin analizini mümkün kılmıştır. Deprem erken uyarı sistemlerinden heyelan duyarlılık haritalamasına, stratigrafik sınıflandırmadan iklim verilerinin yeniden yapılandırılmasına kadar, makine öğrenmesi modelleri artık rutin olarak geleneksel yöntemleri tahmin performansı kıyaslamalarında geride bırakmaktadır (Lary vd. 2016; Reichstein vd. 2019; Rolnick vd. 2019; Bergen vd. 2019).

Bu dönüşüm yalnızca küresel ölçekte değil, Türkiye’de de hız kazanmıştır. Özellikle deprem tehlikesinin yüksek olduğu ülkelerde erken uyarı ve sismik veri işleme sistemlerinde YZ temelli yaklaşımlar yaygınlaşmaktadır. Benzer şekilde heyelan duyarlılık haritalaması, maden arama ve hidrokarbon potansiyeli değerlendirmelerinde makine öğrenmesi yöntemleri hem akademik hem de kurumsal uygulamalarda kullanılmaktadır. MTA, AFAD ve TPAO gibi kurumların yürüttüğü veri yoğun projelerde algoritmik modelleme yaklaşımlarının artan rolü, bu tartışmayı Türkiye bağlamında da güncel kılmaktadır. Bu nedenle, tahmin doğruluğu ile jeolojik yorum arasındaki ilişkinin sağlıklı kurulması kritik öneme sahiptir. Bu gereklilik yalnızca bilimsel değil, aynı zamanda ekonomik ve toplumsal sonuçlar doğurmaktadır.

Artificial intelligence (AI) has become an integral part of geoscience research and has rapidly advanced in recent years, moving from exploratory analysis to operational use in contexts where decision-making is critical. AI-driven models are now increasingly integrated into disaster assessment, early warning systems, and decision-making processes related to subsurface resources; the outputs of these models directly shape policies, engineering practices, and risk management. This shift has enabled the analysis of ever-growing, heterogeneous, and complex datasets. From earthquake early warning systems to landslide susceptibility mapping, and from stratigraphic classification to the reconstruction of climate data, machine learning models now routinely outperform traditional methods for predictive performance benchmarks (Lary et al. 2016; Reichstein et al. 2019; Rolnick et al. 2019; Bergen et al. 2019).

This transformation has gained momentum not only globally but also in Turkey. AI-based approaches are becoming widespread in early warning and seismic data processing systems, particularly in countries with high earthquake risk. Similarly, machine learning methods are used in both academic and institutional practice for landslide susceptibility mapping, mineral exploration, and hydrocarbon potential assessment. The increasing role of algorithmic modeling approaches in data-intensive projects carried out by institutions such as MTA, AFAD, and TPAO makes this discussion highly relevant in the Turkish context. Therefore, establishing a sound relationship between predictive accuracy and geological interpretation has critical importance. This necessity carries not only scientific but also economic and societal consequences.

Ancak bu tahmin başarısının daha temel bir soruyu gölgeleme riski taşıdığı ileri sürülebilir. Gelişmiş tahmin yeteneği, gelişmiş jeolojik anlayışla karıştırılıyor mu? Bu ikilem yeni değildir. Yerbilimleri, devasa mekansal ve zamansal ölçeklere yayılan Dünya süreçlerini yeniden yapılandırmak için her zaman eksik kayıtlara, dolaylı kanıtlara ve uzman görüşüne dayanmıştır. Yeni olan, algoritmik sistemlerin gözlem ile çıkarım arasındaki arabuluculuk rolünü üstlendiği hız, ölçek ve bu süreçlerin şeffaf olmayışıdır. YZ bu arabuluculuk rolünü giderek daha fazla üstlendikçe, yerbilimlerinde nedensellik, anlam ve açıklama üzerine uzun süredir var olan bilginin doğasına ilişkin sorular hafiflemek yerine daha da keskinleşmiştir.

Günümüzde yerbilimlerindeki birçok YZ uygulaması, birincil başarı göstergesi olarak tahmin performansını önceliklendirmektedir. Modeller genellikle doğruluk metrikleri, çapraz doğrulama skorları veya sınıflandırma oranları kullanılarak değerlendirilmekte; tahminlerin nasıl üretildiği veya jeolojik olarak makul olup olmadığı çoğunlukla sorgulanmamaktadır. Oysa yüksek tahmin performansı, mutlaka jeolojik anlayış üstünlüğü anlamına gelmez (Breiman, 2001; Karpatne vd. 2017). Gözlemlerin seyrek olduğu, belirsizliğin istatistiksel olmaktan ziyade yapısal olduğu ve doğruluğu kanıtlanmış bilgiye genellikle erişilemeyen yeraltı yorumlaması veya afet değerlendirmesi gibi alanlarda bu ayırım özellikle önem arz eder.

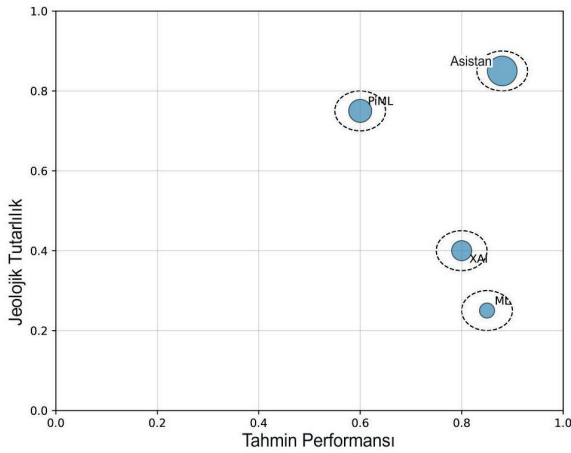
İç işleyişi, kullandığı veri kaynakları ve dayandığı varsayımlar yeterince şeffaf olmayan modellere olan artan bağımlılık, yerbilimlerini yorumlayıcı bir bilim dalı olmaktan çıkarıp, nedenselliğin yeterince sorgulanmadığı, korelasyonların kabul edildiği bir veri işleme faaliyetine dönüştürme riski taşımaktadır. Bu sorun, son yıllarda açıklanabilir yapay zeka (XAI) ve fizik bilgisine dayalı makine öğrenimi (PIML) gibi yaklaşımların geliştirilmesiyle daha görünür hale gelmiştir. XAI, özellikle yüksek riskli

It can be argued that this predictive success risks overshadowing a more fundamental question: Is advanced predictive capability being conflated with enhanced geological understanding? This dilemma is not new. The geosciences have always relied on incomplete records, indirect evidence, and expert opinion to reconstruct Earth processes that span vast spatial and temporal scales. What is new is the speed, scale, and opacity with which algorithmic systems now assume the role of mediating between observation and inference. As AI increasingly undertakes this mediating role, long-standing questions about the nature of knowledge in the geosciences—concerning causality, meaning, and explanation—have become sharper rather than diminished.

Today, many AI applications in the geosciences prioritize predictive performance as the primary indicator of success. Models are often evaluated using accuracy metrics, cross-validation scores, or classification rates, while questions about how predictions are generated or whether they are geologically plausible frequently remain unanswered. However, high predictive performance does not necessarily imply superior geological understanding (Breiman, 2001; Karpatne et al., 2017). This distinction is particularly important in fields such as subsurface interpretation or disaster assessment, where observations are sparse, uncertainty is structural rather than merely statistical, and ground truth is often inaccessible.

The increasing reliance on models whose internal workings, data sources, and underlying assumptions are not sufficiently transparent risks transforming the geosciences from an interpretive science into a data-processing activity where correlations are accepted without adequate scrutiny of causality. This issue has become more visible in recent years with the development of approaches such as explainable AI (XAI) and physics-informed machine learning (PIML). While XAI aims to enhance model transparency

kararların alındığı alanlarda model şeffaflığını ve yorumlanabilirliğini artırmayı hedeflerken (Rudin, 2019; Molnar, 2022), PIML ise fizik yasalarını veya alana özgü teorik bilgiyi doğrudan öğrenme sürecine dâhil ederek yalnızca korelasyonel değil, nedensel olarak anlamlı modeller geliştirilmesine olanak tanır (Karpatne vd. 2017; Raissi vd. 2019). Bu yaklaşımlar, YZ'nın yer bilimlerinde salt bir tahmin aracı olmanın ötesine geçmesi için önemli metodolojik alternatifler sunmaktadır. Farklı YZ yaklaşımlarının şeffaflık, jeolojik tutarlılık ve tahmin performansı açısından konumlanışı Şekil 1 ve Çizelge 1'de özetlenmiştir.



Şekil 1. Yer bilimlerinde farklı yapay zeka yaklaşımlarının tahmin performansı ve jeolojik tutarlılık boyutlarındaki kavramsal konumlandırılması. Balon boyutu model şeffaflığını temsil ederken, kesikli elipsler kavramsal belirsizlik aralıklarını gösterir. Göreceli konumlar sezgiseldir ve nicel kıyaslama sonuçlarını temsil etmez.

Bununla birlikte, bu yaklaşımların da bazı sınırlılıkları bulunmaktadır. XAI yöntemleri çoğunlukla model sonrası (post-hoc) açıklamalar sunmakta olup, bu açıklamaların her zaman gerçek nedensel ilişkileri yansıtmayabileceği tartışılmaktadır. Benzer şekilde, PIML yaklaşımları fiziksel tutarlılığı artırma potansiyeline sahip olmakla birlikte, model karmaşıklığını artırabilir

and interpretability, especially in high-stakes decision-making domains (Rudin, 2019; Molnar, 2022), PIML enables the development of models that are not merely correlational but causally meaningful by incorporating physical laws or domain-specific theoretical knowledge directly into the learning process (Karpatne et al., 2017; Raissi et al., 2019). These approaches offer important methodological alternatives for AI to move beyond being a mere prediction tool in the geosciences. The positioning of different AI approaches in terms of transparency, geological consistency, and predictive performance is summarized in Figure 1 and Table 1.

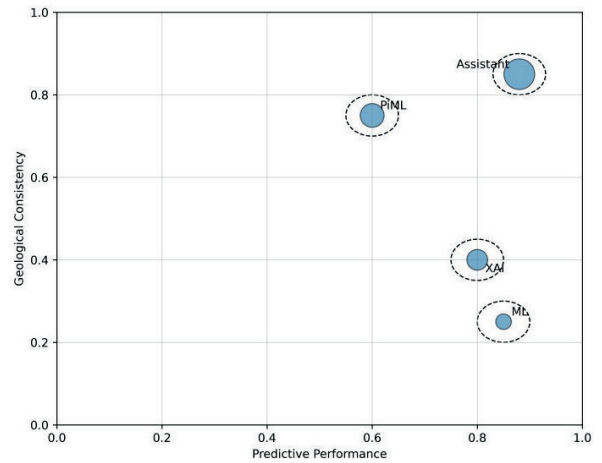


Figure 1. Conceptual positioning of different AI approaches in the geosciences in terms of prediction performance and geological consistency. Balloon size represents model transparency, while dashed ellipses indicate ranges of conceptual uncertainty. Relative positions are intuitive and do not represent quantitative comparison results.

However, these approaches also have certain limitations. XAI methods often provide post-hoc explanations, and it has been debated as to whether these explanations always reflect true causal relationships. Similarly, while PIML approaches have the potential to enhance physical consistency, they may increase model complexity and may not be easily applicable for

ve yüksek kaliteli veri ile iyi tanımlanmış fiziksel denklemlere olan bağımlılığı nedeniyle her uygulama için kolayca uygulanamayabilir. Bu sınırlılıklar, farklı yaklaşımların tamamlayıcı biçimde değerlendirilmesini gerekli kılmaktadır.

Çizelge 1. Farklı yapay zekâ yaklaşımlarının karşılaştırmalı değerlendirmesi.

Yaklaşım	Şeffaflık	Jeolojik tutarlılık	Performans	Risk
ML	düşük	düşük	yüksek	yüksek
XAI	orta	düşük-orta	yüksek	orta
PIML	orta	yüksek	orta	düşük
Asistan	yüksek	yüksek	yüksek (potansiyel)	düşük-orta

Son çalışmalar, geniş jeomorfolojik ve çevresel ipuçları temelinde geliştirilen modellerin, heyelan duyarlılık haritalamasında yüksek sınıflandırma doğruluğuna ulaşırken, eğim-süreç eşikleri veya malzeme sürekliliği gibi temel jeolojik kısıtlamaları ihmal edebildiğini göstermiştir. Örneğin, bir heyelan duyarlılık modelinde eğim yerine yalnızca yükselti değişkeninin güçlü bir belirleyici olarak model tarafından istatistiksel bir korelasyon şeklinde keşfedildiği bir durumda, model tarihsel veri seti üzerinde yüksek doğruluk gösterebilir. Ancak bu ilişki, jeolojik süreçler açısından nedensel değildir ve farklı topoğrafik koşullara sahip bir bölgede model başarısız olacaktır. Benzer şekilde, stratigrafik sınıflandırmada modelin litolojik birimleri doğru tahmin etmesine rağmen, bu birimlerin stratigrafik sıralamayı ihlal edecek şekilde konumlandırılması, yüksek doğruluğa rağmen düşük jeolojik tutarlılığa işaret eder. Bu tür durumlar, tahmin performansının tek başına yeterli bir değerlendirme ölçütü olmadığını açıkça

every application due to their reliance on high-quality data and well-defined physical equations. These limitations necessitate the complementary evaluation of different approaches.

Table 1. Comparative evaluation of different artificial intelligence approaches.

Approach	Transparency	Geological consistency	Performance	Risk
ML	low	low	high	high
XAI	medium	low-medium	high	medium
PIML	medium	high	medium	low
Assistant	high	high	high (potential)	low-medium

Recent studies show that models developed on the basis of broad geomorphological and environmental cues can achieve high classification accuracy in landslide susceptibility mapping, while neglecting fundamental geological constraints such as slope-process thresholds or material continuity. For example, consider a landslide susceptibility model where, instead of slope, only the elevation variable is discovered to be a strong predictor in the form of a statistical correlation by the model. The model may have high accuracy for the historical dataset. However, this relationship is not causal in terms of geological processes, and the model will fail in a region with different topographic conditions. Similarly, in stratigraphic classification, even if the model correctly predicts lithological units, the placement of these units in a way that violates stratigraphic ordering indicates low geological consistency despite high accuracy. Such cases clearly demonstrate that predictive performance alone is not a sufficient evaluation criterion. For instance, in a study by Karpatne et al. (2017), a machine learning model trained

ortaya koymaktadır. Örneğin, Karpatne vd. (2017) tarafından yapılan bir çalışmada, yalnızca geçmiş heyelan envanteri ile eğitilen bir makine öğrenmesi modeli, fiziksel süreç temelli olarak ilgisiz ancak geçmişte heyelanlarla ilişkili proksilere (örneğin belirli bir yükselti aralığı) dayanarak yüksek doğruluk göstermiş; ancak değişen iklim koşullarında bu korelasyonlar geçersiz hale geldiğinde modelin tahmin başarısı ciddi ölçüde düşmüştür. Benzer şekilde, Rudin (2019), yüksek doğruluklu “açıklanamayan karmaşık modellerin” (black-box) adli veri analizinde yanıltıcı sonuçlar üretebildiğini, oysa yorumlanabilir modellerin daha güvenilir olduğunu vurgulamıştır.

Stratigrafik sınıflandırma alanında da benzer sorunlar gözlenmektedir. Derin öğrenme modelleri, litolojik birimleri yüksek doğrulukla ayırt edebilmekte; ancak süperpozisyon ilkesi veya bilinen yapısal ilişkiler model çıktılarıyla çeliştiğinde, bu çelişki yalnızca performans metriği üzerinden değil, jeolojik tutarlılık açısından da değerlendirilmelidir (Bergen vd. 2019). Bu tür örnekler, model geliştirme sürecinde jeolojik ön bilgilerin dışlanması durumunda, tahmin başarısının jeolojik anlayışı temsil etmekten uzaklaşabileceğini açıkça göstermektedir.

Bu durumu daha somutlaştırmak için yarı gerçekçi bir örnek düşünülebilir. Bir heyelan duyarlılık modelinin sayısal yükseklik modeli (DEM), yağış verisi ve litolojik haritalar kullanılarak eğitildiği varsayalım. Model, eğitim ve doğrulama veri setlerinde %90'ın üzerinde doğruluk (örneğin %92) elde edebilir. Ancak model girdileri arasında eğim açısının (slope) yer almaması durumunda, modelin yüksek doğruluğu jeolojik süreçleri temsil etmekten ziyade, tarihsel veri setinde heyelanlarla korelasyon gösteren ancak fiziksel süreç temelli olarak ikincil olan değişkenlere dayanabilir. Bu durumda model, farklı topoğrafik özelliklere sahip bir bölgede uygulandığında başarısız olacaktır. Çünkü temel süreç kontrolü olan eğim parametresi modele dahil edilmemiştir. Bu tür örnekler, yüksek tahmin doğruluğunun jeolojik anlamlılıkla her zaman

solely on a past landslide inventory achieved high accuracy based on proxies that were physically unrelated but historically associated with landslides (e.g., a specific elevation range). However, when these correlations became invalid under changing climatic conditions, the model's predictive performance dropped significantly. Similarly, Rudin (2019) emphasized that high-accuracy “black-box” models can produce misleading results in forensic data analysis, whereas interpretable models are more reliable.

Similar problems are observed in the field of stratigraphic classification. Deep learning models can distinguish lithological units with high accuracy; but when the principle of superposition or known structural relationships contradict the model outputs, this contradiction should be evaluated not only in terms of performance metrics but also in terms of geological consistency (Bergen et al. 2019). Such examples clearly show that when prior geological knowledge is excluded from the model development process, predictive success may cease to represent geological understanding

To make this situation more concrete, consider a semi-realistic example. Suppose a landslide susceptibility model is trained using a digital elevation model (DEM), rainfall data, and lithological maps. The model might achieve over 90% accuracy (e.g., 92%) on the training and validation datasets. However, if slope angle is not included among the model inputs, the model's high accuracy may rely not on representing geological processes but on variables that correlate with landslides in the historical dataset but are secondary from a physical process perspective. In this case, the model will fail when applied to a region with different topographic characteristics, because the slope parameter—the fundamental process control—was not included in the model. Such examples clearly demonstrate that high predictive accuracy does not always coincide with geological meaningfulness.

örtüşmediğini açıkça göstermektedir.

Özerk bir karar alıcı olarak konumlandırmak yerine, YZ'nin veri ile jeolojik akıl yürütme arasında yorumlayıcı bir araç olarak sınırlandırılmalıdır ve bu rol kavramsal olarak "asistan/yardımcı" olarak değerlendirilmelidir. Bu çerçeveye, üç metodolojik ilkeye dayanmaktadır.

Önerilen "Asistan/Yardımcı" Çerçevesi

Bu çalışma kapsamında önerilen "asistan/yardımcı" yaklaşımı yalnızca kavramsal bir çerçeve değil, aynı zamanda yarı-formal bir modelleme iş akışı (workflow) olarak tanımlanabilir:

Aşama 1 – Girdi:

- Gözlemsel veri (jeofizik, jeokimyas, uzaktan algılama vb.)
- Temel jeolojik kısıtlar (stratigrafik ilişkiler, fiziksel eşikler, yapısal süreklilik)

Aşama 2 – Modelleme:

- Makine öğrenmesi / derin öğrenme modelleri
- Gerekliğinde fizik bilgili yaklaşımlar (PIML)

Aşama 3 – Çıktıların Dönüştürülmesi:

- XAI teknikleri (ör. SHAP, LIME)
- Model çıktılarının jeolojik kavramlara (fasiyes, süreç, deformasyon vb.) çevrilmesi

Aşama 4 – Doğrulama:

- Klasik performans metrikleri (accuracy, F1, ROC vb.)
- Jeolojik tutarlılık testleri

Aşama 5 – Geri Besleme:

- Jeolojik olarak tutarsız sonuçların modele geri beslenmesi
- Modelin yeniden eğitilmesi / kısıtların güncellenmesi

Instead of positioning AI as an autonomous decision-maker, AI should be constrained as an interpretive tool between data and geological reasoning, and this role should be conceptually considered as that of an "assistant." This framework is based on three methodological principles.

Proposed "Assistant" Framework

The "assistant" approach proposed within the scope of this study can be defined not only as a conceptual framework but also as a semi-formal modeling workflow:

Stage 1 – Input:

- *Observational data (geophysical, geochemical, remote sensing, etc.)*
- *Basic geological constraints (stratigraphic relationships, physical thresholds, structural continuity)*

Stage 2 – Modeling:

- *Machine learning/deep learning models*
- *Physics-informed approaches (PIML) where necessary*

Stage 3 – Transformation of Outputs:

- *XAI techniques (e.g., SHAP, LIME)*
- *Translation of model outputs into geological concepts (facies, process, deformation, etc.)*

Stage 4 – Validation:

- *Classical performance metrics (accuracy, F1, ROC, etc.)*
- *Geological consistency tests*

Stage 5 – Feedback:

- *Feeding geologically inconsistent results back into the model*
- *Retraining the model/updating constraints*

Bu yapı, YZ'yi özerk bir karar verici olmaktan çıkararak, veri ile jeolojik akıl yürütme arasında insan faktörünün dahil olduğu (human-in-the-loop) bir yorumlayıcı sistem olarak konumlandırır. Bu ilkeler ve aralarındaki ilişkiler Şekil 2'de akış şeması olarak gösterilmiştir. Şekil 1'de ise önerilen yaklaşımın diğer YZ paradigmalarıyla karşılaştırması sunulmaktadır. Çizelge 1'de, geleneksel makine öğrenmesi (ML), açıklanabilir yapay zekâ (XAI), fizik bilgili makine öğrenmesi (PIML) ve önerilen “asistan/yardımcı” çerçevesi; şeffaflık, jeolojik tutarlılık, tahmin performansı ve risk açısından nitel (kalitatif) olarak karşılaştırılmaktadır. Bu karşılaştırmada PIML'in jeolojik tutarlılığı, modelleme sürecine doğrudan gömülü fizik kısıtlarından (örneğin diferansiyel denklemler, korunum yasaları) kaynaklanırken; önerilen “asistan/yardımcı” çerçevesi jeolojik tutarlılığı hem PIML benzeri fizik kısıtlarını hem de model çıktılarının insan faktörünün dahil olduğu (human-in-the-loop) ile sonradan doğrulanmasını birleştirerek sağlamaktadır. Bu nedenle her iki yaklaşım da jeolojik tutarlılık açısından “yüksek” olarak değerlendirilmiş, ancak asistan çerçevesi ek bir doğrulama mekanizması sunmaktadır. Değerlendirme, her bir yaklaşımın jeolojik yorumlama süreçlerine katkısı ve potansiyel sınırlılıkları dikkate alınarak kavramsal bir çerçevede sunulmuştur.

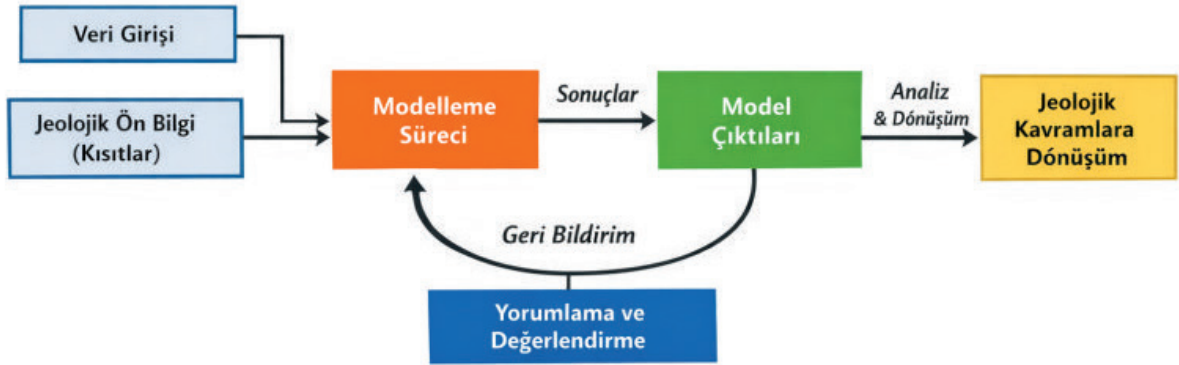
Jeolojik önbilginin modellenmesi: Model girdilerine stratigrafik süreklilik, süreç eşikleri (örneğin eğim açısı, malzeme dayanımı) veya yapısal ilişkiler gibi yerleşik jeolojik kavramların temel jeolojik kısıt olarak eklenmesi. Bu, PIML yaklaşımlarıyla doğrudan ilişkilidir (Karpatne vd. 2017; Raissi vd. 2019). Bu süreç adımları Şekil 1'de görselleştirilmiştir.

Jeolojik olarak makul olmayan yüksek doğrulukların tespiti: Modelin yüksek performansına rağmen, tahminlerin dayandığı korelasyonların jeolojik olarak anlamlandırılmadığı durumların sistematik biçimde işaretlenmesi ve bu tür modellerin karar destek süreçlerinde kullanılmadan önce yeniden değerlendirilmesi.

This structure positions AI not as an autonomous decision-maker but as an interpretive system with humans-in-the-loop between data and geological reasoning. These principles and their relationships are shown as a flowchart in Figure 2. Figure 1 presents a comparison of the proposed approach with other AI paradigms. Table 1 provides a qualitative comparison of traditional machine learning (ML), explainable AI (XAI), physics-informed machine learning (PIML), and the proposed “assistant” framework in terms of transparency, geological consistency, predictive performance, and risk. In this comparison, while PIML’s geological consistency arises from physical constraints (e.g., differential equations, conservation laws) embedded directly into the modeling process, the proposed “assistant” framework achieves geological consistency by combining both PIML-like physical constraints and post-hoc validation with human-in-the-loop. Therefore, both approaches are rated as “high” in terms of geological consistency, but the assistant framework offers an additional validation mechanism. The evaluation is presented within a conceptual framework, considering each approach’s contribution to geological interpretation processes and its potential limitations.

Modeling of prior geological knowledge: Adding inherent geological concepts such as stratigraphic continuity, process thresholds (e.g., slope angle, material strength), or structural relationships as basic geological constraints to model inputs. This is directly related to PIML approaches (Karpatne et al., 2017; Raissi et al., 2019). These process steps are visualized in Figure 2.

Translation of model outputs into geological concepts: Reporting and interpreting predictions not merely as class labels or numerical values, but together with their geological meaning (e.g., facies relationships, deformation evolution, production potential). This stage can be supported



Şekil 2. “Asistan/Yardımcı” çerçevesinin kavramsal iş akışı. Bu şemada veri girişi, jeolojik ön bilgi (kısıtlar), modelleme süreci, model çıktıların jeolojik kavramlara dönüştürülmesi, yorumlama aşaması ve geri bildirim döngüsü arasındaki ilişkiler insan faktörünün dahil olduğu (human-in-the-loop) yaklaşımı kapsamında gösterilmektedir.

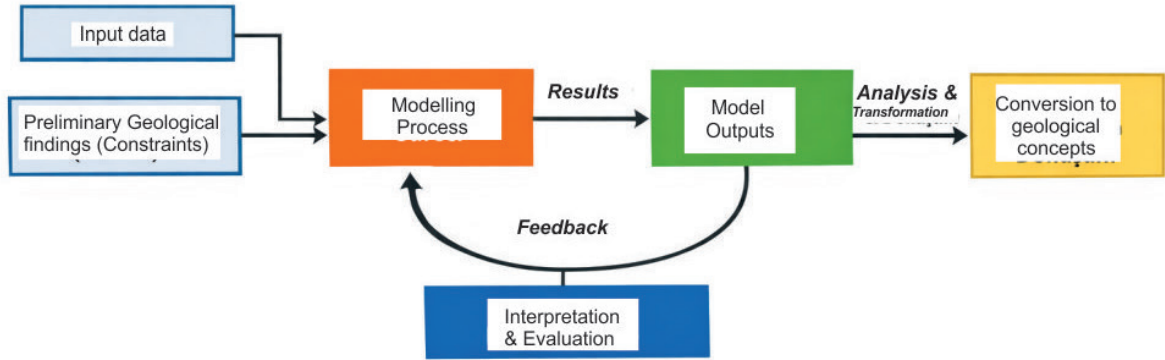


Figure 2. Conceptual workflow for the “Assistant” framework. In this diagram, the relationships between data input, geological prior knowledge (constraints), the modeling process, the conversion of model outputs into geological concepts, the interpretation phase, and the feedback loop are depicted within the context of a human-in-the-loop approach.

Bu çalışmada “jeolojik tutarlılık”, yalnızca nitel bir kavram olarak ele alınmamakta, aynı zamanda çeşitli ölçütler üzerinden yarı nicel biçimde değerlendirilmektedir. Bu kapsamda, öncelikle stratigrafik uyum dikkate alınarak süperpozisyon ilkesi ve bilinen stratigrafik sıralamaların ihlal edilip edilmediği kontrol edilmektedir. Ayrıca fiziksel eşiklerin korunumu değerlendirilmekte; eğitim açısı, kaya dayanımı ve akışkan basıncı gibi süreç temelli parametrelerin fiziksel sınırlarının aşılıp aşılmadığı incelenmektedir. Bunun yanı sıra uzaysal süreklilik ölçütü ile litolojik birimlerin ve yapısal özelliklerin jeolojik olarak beklenen

by XAI techniques (e.g., SHAP, LIME) (Lundberg & Lee, 2017; Ribeiro et al., 2016).

Detection of geologically implausible high accuracy: Systematically flagging cases where, despite a model’s high performance, the correlations underlying the predictions cannot be geologically interpreted, and re-evaluating these models before using them in decision-support processes.

In this study, “geological consistency” is not treated merely as a qualitative concept but is also evaluated in a semi-quantitative manner

mekânsal devamlılığı gösterip göstermediği analiz edilmektedir. Son olarak, model çıktıları alan uzmanlar tarafından yarı nicel ölçekler kullanılarak (örneğin 1–5 arası tutarlılık puanı) değerlendirilmektedir. Bu çok boyutlu yaklaşım, yüksek tahmin doğruluğuna sahip olsa dahi jeolojik açıdan anlamlı olmayan modellerin tespit edilmesini mümkün kılmaktadır.

Bu değerlendirmeyi daha sistematik hale getirmek amacıyla, jeolojik tutarlılık yarı nicel bir skor ile ifade edilebilir. Bu kapsamda, jeolojik tutarlılık üç temel bileşen üzerinden tanımlanabilir: (i) stratigrafik uyum (S_1), (ii) fiziksel eşiklerin korunumu (S_2) ve (iii) uzaysal süreklilik (S_3). Her bir bileşen 0 ile 1 arasında normalize edilerek değerlendirilir. Buna göre basit bir Jeolojik Tutarlılık Skoru (JTS) aşağıdaki şekilde tanımlanabilir:

$$JTS = (S_1 + S_2 + S_3) / 3$$

Bu tür bir yaklaşım, model performansının yalnızca doğruluk metrikleri ile değil, aynı zamanda jeolojik anlamlılık açısından da değerlendirilmesine olanak tanır. Amaç kesin bir ölçüm üretmekten ziyade, jeolojik muhakemenin model değerlendirme sürecine sistematik olarak dahil edilmesini sağlamaktır.

Bu çerçeve yalnızca terminolojik bir öneri değil, aynı zamanda uygulamaya dönük bir metodolojik yönlendirmedir. Örneğin, heyelan duyarlılık haritalamasında geliştirilen bir makine öğrenmesi modeli, geçmiş heyelan envanteri ile eğitildiğinde yüksek doğruluk elde edebilir; ancak model, fiziksel olarak anlamlı eğitim-süreç eşiklerini dikkate almaksızın yalnızca korelasyonel örüntülere dayanıyorsa, değişen iklimsel koşullar altında yanıltıcı sonuçlar üretebilir. “Asistan/yardımcı” çerçevesi, bu tür modellerin geliştirilmesi sırasında jeolojik kısıtların işe koşulmasını ve çıktıların jeolojik akıl yürütmeyle sınanmasını zorunlu kılar.

Böyle bir yaklaşım, tahmine dayalı modellemeyi reddetmediği gibi geleneksel yorumlamayı da yüceltmez. Bunun yerine,

based on various criteria. In this context, first, stratigraphic conformity is considered, checking whether the principle of superposition and known stratigraphic sequences are violated. Additionally, the preservation of physical thresholds is assessed, examining whether process-based parameters such as slope angle, rock strength, and fluid pressure exceed physical limits. Furthermore, using the spatial continuity criterion, whether lithological units and structural features exhibit the spatially expected geological continuity or not is analyzed. Finally, model outputs are evaluated by domain experts using semi-quantitative scales (e.g., a consistency score from 1 to 5). This multidimensional approach makes it possible to identify models that are geologically meaningless even if they possess high predictive accuracy.

To make this evaluation more systematic, geological consistency can be expressed as a semi-quantitative score. In this context, geological consistency can be defined based on three fundamental components: (i) stratigraphic conformity (S_1), (ii) preservation of physical thresholds (S_2), and (iii) spatial continuity (S_3). Each component is evaluated by normalizing between 0 and 1. Accordingly, a simple Geological Consistency Score (GCS) can be defined as follows:

$$JTS = (S_1 + S_2 + S_3) / 3$$

Such an approach allows model performance to be evaluated not only in terms of accuracy metrics but also in terms of geological meaningfulness. The aim is not to produce a precise measurement but to ensure the systematic incorporation of geological reasoning into the model evaluation process.

This framework is not merely a terminological proposal but also a practical methodological guideline. For example, a machine learning model developed for landslide susceptibility mapping may achieve high accuracy when trained on a past landslide inventory; however, if the model relies solely on correlational patterns without considering physically meaningful slope-

yerbilimleri camiasındaki farklı aktörler için belirli çıkarımları beraberinde getirir:

Araştırmacılar: Jeolojik varsayımları gizlemek yerine açığa çıkaran modeller tasarlamaya teşvik edilir. Model geliştirme sürecine jeolojik ön bilgileri ve kısıtlamaları açıkça dâhil etmeleri, yalnızca tahmin performansına değil, model çıktılarının jeolojik tutarlılığına da öncelik vermeleri beklenir. Yorumlanabilir ya da kuram rehberliğinde öğrenme yaklaşımlarının tercih edilmesi, sorumlu bir araştırma pratiğinin parçasıdır.

Hakemler ve Editörler: Değerlendirme süreçlerinde tahmin doğruluğu metriklerinin yanında jeolojik tutarlılık ve yorumlanabilirliği de teşvik edici ölçütler olarak benimsemelidir. Çalışmalarda model çıktılarının yerleşik jeolojik kavramlarla ilişkilendirilmesini talep etmelidir; yüksek doğruluğa jeolojik olarak makul olmayan yollardan ulaşan modelleri eleştirel biçimde değerlendirmelidir.

Fon sağlayan kurumlar ve karar alıcılar: Kuram rehberliğinde ve yorumlama bilincine sahip YZ çerçevelerini öncelikli destek alanları arasında konumlandırılmalı; disiplinlerarası projelerde jeoloji ve bilgisayar bilimlerinin dengeli biçimde temsil edilmesini gözetmelidir. Yalnızca tahmin başarısını değil, jeolojik anlayışa katkıyı da ölçen değerlendirme kriterlerinin geliştirilmesi, sürdürülebilir ve nitelikli bir bilimsel ortamın tesis edilmesine katkı sağlayacaktır.

Yapay zeka, yerbilimlerinde gözlem ile sonuç arasındaki asistan/yardımcı rolünü benzeri görülmemiş bir hız ve ölçekte yerine getirebilir. Deprem erken uyarı sistemlerinden heyelan duyarlılık haritalarına, yeraltı kaynaklarının modellenmesinden stratigrafik sınıflandırmaya kadar pek çok alanda YZ modelleri, geleneksel yöntemleri tahmin performansı açısından geride bırakmaktadır. Ancak bu teknik başarı, jeolojik anlayışla eşdeğer görülmemelidir.

process thresholds, it may produce misleading results under changing climatic conditions. The “assistant” framework necessitates the deployment of geological constraints during the development of such models and requires that outputs be tested against geological reasoning.

Such an approach neither rejects predictive modeling nor glorifies traditional interpretation. Instead, it carries specific implications for different actors within the geoscience community:

Researchers: It is recommended that researchers be encouraged to design models that reveal rather than conceal geological assumptions. They are expected to explicitly incorporate prior geological knowledge and constraints into the model development process, prioritizing not only predictive performance but also the geological consistency of model outputs. Preferring interpretable or theory-guided learning approaches is part of responsible research practice.

Reviewers and Editors: In evaluation processes, reviewers and editors should adopt geological consistency and interpretability as encouraging criteria alongside predictive accuracy metrics. They should demand that studies relate model outputs to established geological concepts; they should critically evaluate models that achieve high accuracy through geologically implausible means.

Funding agencies and decision-makers: It is recommended that funding agencies and decision-makers should position theory-guided and interpretation-aware AI frameworks among priority funding areas and ensure balanced representation of geology and computer science in interdisciplinary projects. Developing evaluation criteria that measure not only predictive success but also contributions to geological understanding will contribute to establishing a sustainable and high-quality scientific environment.

Artificial intelligence can fulfill the assistant role between observation and outcome in the geosciences at an unprecedented speed and

Yüksek tahmin doğruluğu, jeolojik tutarlılık ve nedensel çıkarım anlamına gelmez; aksine, yorumlayıcı muhakemenin marjinalleşmesi halinde modellerin fiziksel süreç temelli olarak ilgisiz proksilere dayanan yanıltıcı sonuçlar üretebileceği dikkatle değerlendirilmelidir. Özellikle gözlemlerin seyrek olduğu, belirsizliğin yapısal nitelikte olduğu ve doğruluğu kanıtlanmış bilginin sınırlı olduğu yeraltı yorumlaması ve afet değerlendirmesi gibi alanlarda bu risk daha da belirgindir.

Bu sorunu aşmak amacıyla önerilen YZ'nın asistan/yardımcı vasfı, YZ'yi özerk bir karar alıcı olarak değil, veri ile jeolojik akıl yürütme arasında yorumlayıcı bir aracı olarak konumlandırmaktadır. Bu yaklaşım, tahmine dayalı modellemeyi reddetmeksizin, yorumlamayı aktif ve eleştirel bir bilimsel pratik olarak korumayı hedefler.

Yapay zeka yerbilimlerini yeniden şekillendirmeye ve dönüştürmeye devam edecektir. Bu dönüşümün, kavrayışı derinleştirip derinleştirmeyeceği ya da yalnızca örüntü tespitini hızlandırıp hızlandırmayacağı, araştırmacıların, editörlerin ve kurumların alacağı kararlara bağlıdır. Yorumlamayı temel bir bilimsel pratik olarak korumak, veri odaklı bilimsel gelişmenin uzun ömürlü ve uygulanabilir jeolojik öngörüye dönüşmesine yardımcı olacaktır.

KATKI BELİRTME

Bu yazı, Prof. Dr. Nizamettin Kazancı'nın teşviki üzerine hazırlanmış olup kendisine katkıları için teşekkür ederim. Ayrıca, Dr. Fatih Köroğlu'na değerli yorumları için minnettarım. Makalenin ilk taslağı Prof. Dr. Erdinç Yiğitbaş tarafından incelenmiş olup, kendisinin detaylı incelemesi teknik notun kalitesini önemli ölçüde artırmıştır. Ayrıca üç isimsiz hakemden gelen yapıcı yorumlara teşekkür ederim.

scale. From earthquake early warning systems to landslide susceptibility maps, from subsurface resource modeling to stratigraphic classification, AI models outperform traditional methods in terms of predictive performance in many areas. However, this technical success should not be equated with geological understanding.

High predictive accuracy does not imply geological consistency and causal inference; on the contrary, it should be carefully considered that if interpretive reasoning becomes marginalized, models may produce misleading results based on physically irrelevant proxies. This risk is even more pronounced in areas such as subsurface interpretation and disaster assessment, where observations are sparse, uncertainty is structural in nature, and ground truth is limited.

To overcome this problem, the proposed "assistant" role of AI positions AI not as an autonomous decision-maker but as an interpretive intermediary between data and geological reasoning. Without rejecting predictive modeling, this approach aims to preserve interpretation as an active and critical scientific practice.

Artificial intelligence will continue to reshape and transform the geosciences. Whether this transformation deepens understanding or merely accelerates pattern detection depends on the decisions made by researchers, editors, and institutions. Preserving interpretation as a fundamental scientific practice will help transform data-driven scientific advances into long-lasting and applicable geological foresight.

ACKNOWLEDGMENTS

This paper was prepared with the encouragement of Prof. Dr. Nizamettin Kazancı, to whom I extend my thanks for his contributions. I am also grateful to Dr. Fatih Köroğlu for his valuable comments. The first draft of the manuscript was reviewed by Prof. Dr. Erdinç Yiğitbaş, whose detailed examination significantly improved the quality of the technical note. I also thank three anonymous reviewers for their constructive comments.

ORCID

Oğuz Mülayim  <https://orcid.org/0000-0003-3283-8070>

KAYNAKÇA / REFERENCES

- Bergen, K. J., Johnson, P. A., de Hoop, M. V., Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363, Article eaau0323.
<https://doi.org/10.1126/science.aau0323>
- Breiman, L. (2001). Statistical modeling: the two cultures. *Statistical Science*, 16(3), 199-231.
- Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V. (2017). Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10), 2318-2331.
<https://doi.org/10.1109/TKDE.2017.2720168>
- Lary, D. J., Alavi, A. H., Gandomi, A. H., Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1), 3-10.
<https://doi.org/10.1016/j.gsf.2015.07.003>
- Lundberg, S.M., Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.
- Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. <https://christophm.github.io/interpretable-ml-book/>
- Raissi, M., Perdikaris, P., Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707.
<https://doi.org/10.1016/j.jcp.2018.10.045>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195-204. <https://doi.org/10.1038/s41586-019-0912-1>
- Ribeiro, M. T., Singh, S., Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.
<https://doi.org/10.1145/2939672.2939778>
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C., Ng, A. Y., Hassabis, D., Platt, J. C., Bengio, Y. (2019). *Tackling Climate Change with Machine Learning*. arXiv:1906.05433.
<https://doi.org/10.48550/arXiv.1906.05433>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1, 206-215.
<https://doi.org/10.1038/s42256-019-0048-x>

