

T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

EVİRİSEL ALGORİTMALARA DAYALI YENİ DERİN AKTİF
ÖĞRENME YÖNTEMLERİN GELİŞTİRİLMESİ
(DEVELOPING NEW DEEP ACTIVE LEARNING TECHNIQUES BASED ON
EVOLUTIONARY ALGORITHMS)

Tezi Hazırlayan
Maliki MOUSTAPHA

Tezi Yöneten
Doç. Dr. Celal ÖZTÜRK

Bilgisayar Mühendisliği Anabilim Dalı
Doktora Tez Önerisi

Ağustos 2021

KAYSERİ

1. Genel bilgi ve araştırma amacı

Son yıllarda yapılan en önemli araştırmalardan biri yapay zekadır (YZ). YZ, makine öğrenme (MÖ) adı verilen bir alt alanı içeren ise geniş bir alan olarak bilinmektedir. Makine öğrenme ise Derin öğrenme dahil olmak üzere diğer alt alanlardan oluşmaktadır.

Herhangi bir makine öğrenmenin yaşam döngüsü sürecinde, genellikle makine öğrenme algoritmaları, özellikle derin öğrenme algoritmaları için model eğitim süreci için yeterli miktarda veri gerektirir.

Derin öğrenme (DÖ), özellik kesinleştirme yoluyla verimli bir Yapay zekâ modeli oluşturmak için verileri işlemek için insan beyninin işlevselliğini taklit eden günümüzde makine öğrenmenin önde gelen alanlarından biridir. Ancak, DÖ, modeli eğitmek için çok sayıda etiketli veri kümesine ihtiyaç duyulmaktadır ve bu süreç, mevcut verilerin çoğunlukla çok boyutlu ve etiketsiz olduğu alanlarda büyük veri işleme ile günümüzde çok zor hale gelmektedir. Bu sorunu çözmek için birçok araştırmacı şimdi Derin aktif öğrenmeye yöneliyor.

Çoğu durumda, iyi ve verimli bir eğitilmiş model elde etmek adına, makine öğrenme algoritmalarının, model tahmini ve sınıflandırma işlemi için daha iyi ve doğruluk sağlamak üzere verileri etiketlemesi için insanlara ihtiyacı vardır. Bu süreç, yeterli ve iyi etiketlenmiş veri elde etmek için çok fazla zaman, enerji ve maliyet gerektirir. Aktif öğrenme (AÖ), mevcut model performansına göre etiketlenecek veri noktalarını seçerek uygun örneklerin özellikleri aracılığıyla veri etiketleme sürecini daha verimli hale getirmek için kullanılan makine öğrenme tekniklerinden biridir.

Derin Aktif öğrenme (DAÖ) çerçeveleri (frameworks) artık, özellikle etiketlerin yüksek düzeyde uzmanlık gerektirdiği ve işlenmesi zor olduğu alanlarda veya büyük miktarda etiketlenmemiş verinin üretildiği ve veri modelini güncellemek için sürekli olarak eklenmesi gerektiği alanlarda veri etiketleme işlemleri için yaygın olarak kullanılmaktadır. DAÖ çerçevesinin uygulanmasında, sorgu stratejileri ile çok boyutlu verilerin işlenmesi gibi durumlarda farklı sorunlarla karşılaşır. Ve araştırmaların çoğu hala bu Aktif öğrenme stratejilerini geleneksel yöntemlerle optimize etmeye çalışmaktadır.

Evrimsel veya sürü zekâsı algoritmaları, çözüm optimizasyonu için karmaşık mühendislik problemlerini ele alma yeteneği ile bilinir ve bazı makine öğrenme problemlerini çözmek için kullanılır. Literatür incelemelerine dayanarak, bu algoritmalar henüz Derin aktif öğrenmeye

uygulanmadı. Bu nedenle, bu tez çalışması, uygulama sürecinde bazı Derin aktif öğrenme sorunlarını iyileştirmek ve çözmek için evrimsel algoritmaları (örneğin: genetik algoritma) ve sürü zeka algoritmalarını (örneğin: yapay arı kolonisi algoritması) entegre etme potansiyeli olan yeni bir DAL çerçevesi yani yeni yöntemleri geliştirmeyi amaçlamaktadır.

2. Derin öğrenme

Derin öğrenme, insan beyninin işlevsel yapısını simüle ederek Yapay Zekâ (YZ) veri modelleri oluşturmaya çalışan bir Makine öğrenmenin alt kümesidir. Modern derin öğrenmenin başlangıç noktası, 1943 yılında McCulloch-Pitts (MCP) tarafından önerilen model ile düşünülmüştür [7]. Son yıllarda derin öğrenme, makine öğrenme alanında devrim yarattı ve bilgisayarla görüntü işleme ve doğal dil işlemedeki (Natural Language Processing) diğer en son gelişmeler dahil olmak üzere çeşitli uygulamalarda son teknoloji performansa ulaştı.

Ayrıca klinik metinlerden sınıflandırma ve bilgi çıkarmayı otomatikleştirmede kayda değer ilerleme kaydetmiştir [2]. Mevcut gerçek dünya veri problemlerinde, genellikle makine öğrenmede kullanılan en son teknoloji teknikler Derin öğrenme mimarisidir. Derin öğrenme tabanlı mimari, Evrimsel Sinir Ağları (CNN'ler) olarak bilinmektedir ve normal makine öğrenme tekniklerinden daha iyi performans gösterdiğini gösterilmiştir.

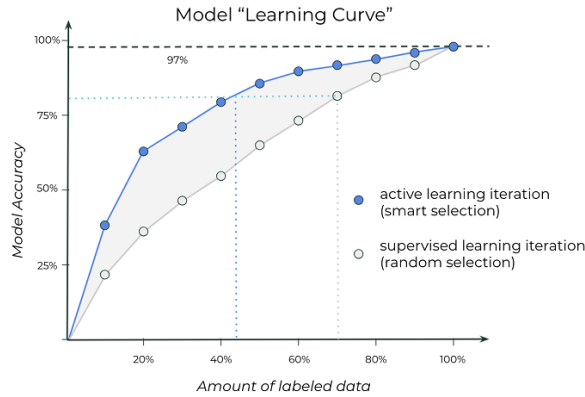
Örnek olarak, klinik metin uygulamalarında CNN ve kural tabanlı yaklaşımların kombinasyonu iyi performans göstermiştir [8-10]. Bununla birlikte, karşılaşılan genel dezavantaj, sağlam bir derin sinir ağı eğitimini, büyük miktarda elle etiketlenmiş eğitim verisi gerekmektedir; ancak bu işlem, yararlı etiketli verileri elde etmek için zaman alıcı ve emek yoğundur.

3. Aktif öğrenme

Aktif öğrenmenin temel amacı, belirli bir makine öğrenme modeli için ek etiketli eğitim verilerinin elde edilmesinin etkinliğini en üst düzeye çıkarmaktır. Bu süreç, az eğitilmiş modeli kendisi kullanarak model için en bilgilendirici olacak etiketlenmemiş veri setini aktif olarak seçerek gerçekleştirilir [2].

Bu kavramın arkasındaki temel fikir, bir makine öğrenme algoritmasının, öğrenmek istediği verileri seçebilmesi durumunda daha az eğitim etiketi kullanarak potansiyel olarak daha iyi doğruluk elde edebilmesidir. Bu tür algoritma, aktif öğrenen olarak bilinmektedir. Şekil 3.1'de, aktif öğrenmenin, mevcut toplam veri hacminin yalnızca %45'i ile %80 doğruluğa ulaşılmasına

nasıl yardımcı olduğunu ve normal düzenli denetimli öğrenme kullanılarak %70'lik doğruluk elde edildiğini göstermektedir.

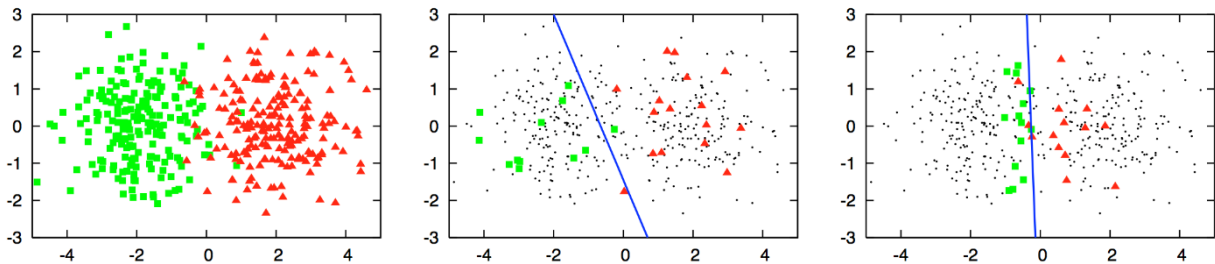


Şekil 3.1 Aktif öğrenmeye dayalı model doğruluğu

Aktif öğrenmenin arkasındaki ana temeli, bir makine öğrenme algoritmasının, öğrenmek istediği verileri seçmesine izin verilirse daha az eğitimle daha iyi performans gösterebilmesidir [6]. Daha sezgisel bir şekilde, Aktif öğrenme, veri işleme görevinin daha az etiketli veri kullandığı ve veri modelinin performansını iyileştirmek için etkileşimli olarak yeni veri noktalarını etiketlediği bir makine öğrenme tekniği olarak tanımlanır.

3.1 Aktif Öğrenmenin Motive Edici Örneği

Şekil 3.2'deki görüntü, basit örneklerle aktif öğrenmenin temel anlayışını açıklamaktadır. Bu görüntüde geleneksel öğrenme süreci ile aktif öğrenme arasındaki farkı kısaca açıklar. Yeşil ve kırmızı renkler iki kümenin bir sınıflandırmasını temsil eder. Tüm örnekler için etiket bulmaya çalışmak çok maliyetli olabilir. Buradaki amaç, küçük bir veri nokta alt kümesinin seçilmesi, ardından modeli eğitmek için birkaç etiketli veriyi kullanmak ve daha sonra kalan veri noktalarını model aracılığıyla sınıflandırılabilir.



Şekil 3.2 Aktif Öğrenmenin Motive

Soldaki ilk görüntü, iki boyutlu özellik alanında temsil edilen iki Gauss sınıfından 400 numunenin bir örneğini göstermektedir. Ortadaki görüntü, problem alanından rastgele seçilen 30 etiketli örnekle eğitilmiş bir lojistik regresyon modelini göstermektedir. Çizgi ise, elde edilen %70 doğrulukla sınıflandırıcının karar sınırını temsil eder. En sağdaki görüntüde, aktif öğrenme tekniğinin sorgulama yöntemi kullanılarak küçük bir nokta alt kümesi seçilerek lojistik regresyon yöntemi de kullanıldığını göstermektedir. Belirsizlik örnekleme yöntemi kullanılarak aktif olarak sorgulanan 30 örnekle eğitildi ve %90 doğruluğa ulaştı. Bu gelişme, sınıflandırıcının çok iyi bir karar sınırı oluşturabilmesi için üstün veri noktalarının seçilmesinden kaynaklanmaktadır.

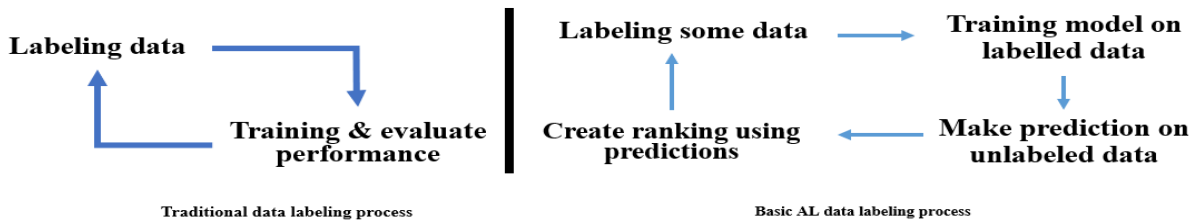
Aktif Öğrenme, daha az etiketlenmiş örneklerle makine öğrenme modelinin iyi performans göstereceği şekilde verilerin etiketlenmesine şansı verir. İyi performans için etiketlenecek en uygun numune grubunu seçer. Kısacası, bir modelin aktif olarak öğrenmesine yardımcı olan ve daha sonra ne öğreneceğini gören, tahmin ve sınıflandırmalarda daha iyi performans gösteren bir teknik olarak görülebilir. Ek verilerin rastgele etiketlenmesiyle karşılaştırıldığında, aktif öğrenme, modelin daha az ek etiketli örnek kullanarak daha yüksek performansa ulaşmasını sağlar ve böylece insan açıklayıcıların verimliliğini ve etkinliğini artırır [6].

3.2 Aktif Öğrenmenin Genel Yaşam Döngüsü.

Aktif öğrenmede kullanılan ana kavramsal adımlar şunlardır:

- Modeli mevcut etiketli verileri üzerinde eğitilmesi.
- Modeli tüm etiketlenmemiş numune üzerinde değerlendirilmesi.
- Değerlendirmeye dayanarak, etiketlenecek numuneleri/numune listesini seçilmesi
- Bu numunelerin etiketlenmesi ve etiketlenmiş verilere eklenmesi.
- Belirli bir koşul (durdurma kriteri) sağlanana kadar yukarıdaki adımları tekrarlanması.

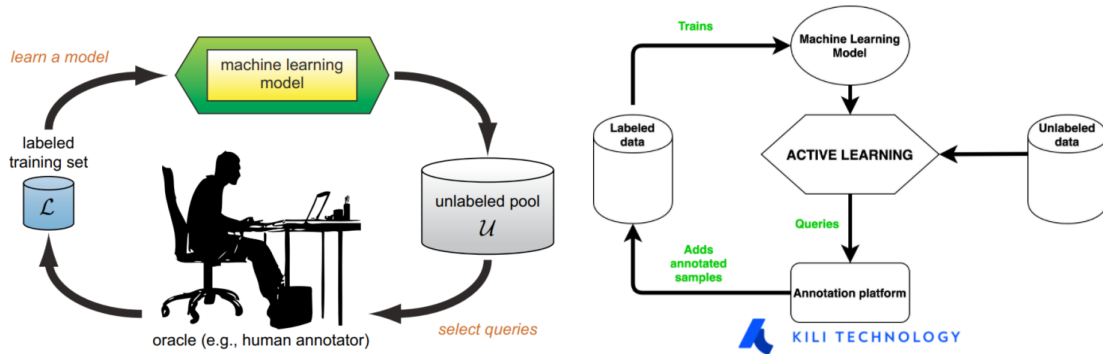
Aşağıdaki şema, aktif öğrenmeye dahil olan mantık sürecini açıklamaktadır.



Şekil 3.3 Aktif Öğrenme şeması

Şekil 3.4'teki her iki görüntü de en bilinen aktif öğrenme yaşam döngüsünü göstermektedir. Her aktif öğrenme senaryosunda, süreci etiketli örneklerin bir başlangıç kümesi L_0 ile başlar ve bunu, parametre tahminleri θ_0 ile bir sınıflandırma modelini eğitmek için kullanır. Daha sonra model, bir dizi etiketlenmemiş veri U_0 (etiketlenmemiş havuz) üzerine uygulanır ve U_0 'daki her x_i örneğine bir bilgilendiricilik (informativeness) ölçüsü atamak için bir $\phi(x_i | \theta_0)$ sorgulama stratejisi kullanır; bu bilgilendiricilik ölçüsü, numunenin o numune üzerinde eğitilmiş olsaydı, o numunenin sınıflandırma modeline ne kadar katkı sağlayacağını gösterir. Elde edilen doğruluk etiketleri kümesi, en yüksek bilgi değeri $\phi(x_i | \theta_0)$ ile U_0 'da n örneğinin bir alt kümesini oluşturur. Bu altküme, yeni daha büyük eğitim seti L_1 'i ve yeni daha küçük etiketlenmemiş U_0 setini oluşturmak için U_1 'den L_0 'a taşınır.

θ_1 parametre tahminlerine sahip yeni bir sınıflandırma modeli daha sonra L_0 üzerinde eğitilir ve U_1 'e uygulanır. Bir kez daha, sorgulama stratejisi $\phi(x_i | \theta_1)$ kullanarak U_1 'den etiketlemek için en bilgilendirici n örnekleri seçmek ve U_2 ve L_2 'yi oluşturmak için L_0 'e eklemek için kullanılır. Bu işlem, sınıflandırma modeli istenen performansa ulaşana veya U 'da hiç örnek kalmayana kadar tekrarlanır. Bir kez daha, sorgulama stratejisi $\phi(x_i | \theta_1)$, U_0 'den etikete en bilgilendirici n örneğini seçmek ve U_2 ve L_2 'yi oluşturmak için L_1 'e eklemek için kullanılır. Bu işlem, sınıflandırma modeli istenen performansa ulaşana veya U 'da hiç örnek kalmayana kadar tekrarlanır.



Şekil 3.4 Aktif Öğrenme detaylı şeması

3.3 Aktif Öğrenme Senaryoları

Aktif öğrenme, öğrenen tarafından problem alanına göre örnekleri (numuneler) sorgulamak için kullanılan farklı senaryoları uygular. Literatürde ele alınan ana ve yaygın senaryolar, üyelik sorgusu sentezi, buhar tabanlı seçici örnekleme ve havuz tabanlı örneklemedir [3]. Bu

senaryoların çoğu, insan etiketlenmemiş örnekler havuzundan her bir örneği etiketlediğinden, döngüdeki insan tabanlı süreçlerdir.

3.3.1 Üyelik sorgusu (Membership query)

Bu senaryoda, öğrenen modeli etiketleme için dağıtılmış örneklerden kendi örneklerini oluşturabilir. Bu yaklaşım, veri numuneyi oluşturmanın kolay olduğu durumlarda soğuk başlatma sorunlarını çözmeyi vaat eder.

3.3.2 Buhar tabanlı seçici örnekleme (Steam-based selective sampling)

Bu senaryo, etiketlenmemiş örnekleri almanın ücretsiz olduğunu varsayar. Daha sonra buna dayanarak, veri kümesinden etiketlenmemiş her bir örneği birer birer seçmekten oluşur ve öğrenen model örneğin etiketini sorgulamanın veya bilgilendiriciliğine dayanarak reddetmenin yeterince yararlı olup olmayacağını belirlemesine olanak tanır. Model eğitilirken ve bir veri örneği ile sunulurken, etiketi görmek isteyip istemediğine hemen karar verir. Bu yaklaşımın dezavantajı, doğal olarak, veri bilimcisinin bütçesi dahilinde kalacağı garantisinin olmamasından kaynaklanmaktadır.

3.3.3 Havlu bazı Örnekleme (Pool based Sampling).

En iyi senaryo olarak bilinen, örneklere açıklama eklemek için en iyi sorguyu veya bir dizi en iyi sorguyu seçmeden önce tüm veri kümesini değerlendirir. Örneklemeden önce bazı bilgilendirici önlemler belirlenir. Daha sonra aktif öğrenen, modelin ilk versiyonunu oluşturan verilerin tamamen etiketlenmiş bir kısmı üzerinde eğitilir. Bu önlemler daha sonra, bir sonraki yineleme için eğitim kümesine hangi örneklerin enjekte edilmesinin en yararlı olacağını belirlemek için uygulanır.

3.4 Aktif Öğrenme Sorgu stratejileri

Sorgulama stratejisi, daha sonra hangi numune etiketleneceğini belirlemek için kullanılan yaklaşımdır. Daha önce de belirttiğimiz gibi, tüm Aktif öğrenme senaryoları, etiketlenmemiş verilerden örnekleri (numuneleri) seçmek için bilgilendiricilik ölçütlerini kullanır. Örnekleri geçmiş sorgulara dayalı olarak sorgulama ve bunlardan kaynaklanan bu yetenek, aktif öğrenme ile pasif öğrenme teknikleri arasındaki temel farktır. Ayrıca alt stratejilere ayrılan birçok strateji türü vardır. [1,2,3] Literatürde en yaygın olarak tartışılan yöntemler ise belirsizlik örnekleme (uncertainty sampling), çeşitlilik örnekleme (diversity sampling),

komiteye göre sorgulama ve yoğunluk ağırlıklı (common uncertainty-based query strategy) yöntemidir.

3.4.1 Belirsizlik örnekleme Stratejisi (Uncertainty sampling Strategy)

[11]'de bahsedildiği gibi, yaygın aktif öğrenme stratejilerinden biri belirsizlik örneklemesidir. Konsept, numunenin güvenine dayanmaktadır. Belirli bir örneğin güveni düşükse, model için daha bilgilendirici olacağı anlamına gelir. Aktif Öğrenme alanındaki yaygın olarak belirsizliğe dayalı sorgu stratejisi algoritmaları şunlardır:

- **En az güven (least confidence) (LC):** Bu stratejiyi kullanarak, öğrenen modeli en olası etiketine en az güvendiği örneği seçer. ϕ 'nin sınıflandırıcısının tahmin güvenine dayalı olarak hesaplandığı en basit algoritmadır. ϕ aşağıdaki gibi hesaplanır:

$$\phi^{LC}(x) = 1 - P(y^*/x; \theta)$$

y^* belirli bir örnek için tahmin edilen sınıftır, yani en yüksek softmax değerine sahip sınıf ve $P(y^*/x; \theta)$ o sınıfla ilişkili softmax değerini temsil eder. [2]

- **Marjinal örnekleme (MC) [32]:** LC yalnızca en olası etiketi dikkate alıp diğer etiket olasılıklarını göz ardı ederken, MC her tahmin edilen örnekten yalnızca en yüksek softmax değerini dikkate alır. Birinci ve ikinci en olası etiketler arasında en küçük farka sahip örneği seçerek LC'nin dezavantajını ortadan kaldırır.

$$\phi^{ES}(x) = Cc = 1 P(y_i|x; \theta) * \log P(y_i | x; \theta)$$

Burada $y^* 1$ ve $y^* 2$, sırasıyla en yüksek ve ikinci en yüksek softmax değerleriyle ilişkili iki sınıfı temsil eder.

- **Entropi Örnekleme (ES) [33]:** Belirsizliğe dayalı bir sorgu algoritması olarak yaygın olarak kullanılır, ana fikri tüm olası etiket olasılıklarını kullanmaktır. Her örneğe entropi formülü uygulanır ve en büyük değere sahip örnek sorgulanır. Bu yaklaşımda ϕ , bu denklemde tanımlanabilen entropi tabanlı bir metriktir:

$$\phi^{ES}(x) = C c=1 P(y_i|x; \theta) * \log P(y_i|x; \theta)$$

burada $C c=1$ tüm olası sınıfların toplamını temsil eder ve $P(y_i|x; \theta)$ ise y_i sınıfı ile ilişkili softmax değeri.

3.4.2 Çeşitlilik Örneklem Stratejisi (Diversity sampling Strategy (DS))

Eldeki eğitilmesi gereken veriler bazen çok çeşitli olabilir ve bu çeşitlilikle başa çıkmak başka bir sorundur. Örnekler arasındaki benzerlik ölçüsüne dayalı olarak, eğitim veri setinin çeşitliliğini en üst düzeye çıkarmak için çeşitlilik örneklem stratejisi kullanılır. Ayrıca derin öğrenme modelleriyle birleştirilmiştir. **Öklid Kümesi Tabanlı Örneklem (Euclidean Cluster-based Sampling (ES))** ve **Kosinüs Kümesi Tabanlı Örneklem (Cosine Cluster-Based Sampling (CC))**, kanser patoloji raporlarını sınıflandırmak için iki DS algoritması örneği olarak [2]'de kullanılmıştır.

3.4.3 Komiteye göre sorgulama örneklem stratejisi (QBC) (Query-by-committee sampling strategy (QBC))

QBC stratejisinin ana fikri, çoklu öngörücü modellerin (komite) eğitilmesi ve aralarındaki anlaşmazlığa dayalı olarak en iyi modeli elde etmek için değerlendirilmesidir. Örnekler, komite içinde ne kadar anlaşmazlık olduğuna göre sıralanır. Bu nedenle, en yüksek anlaşmazlıkla ilişkili Örnekler seçilir ve eğitim veri kümesine eklenir.

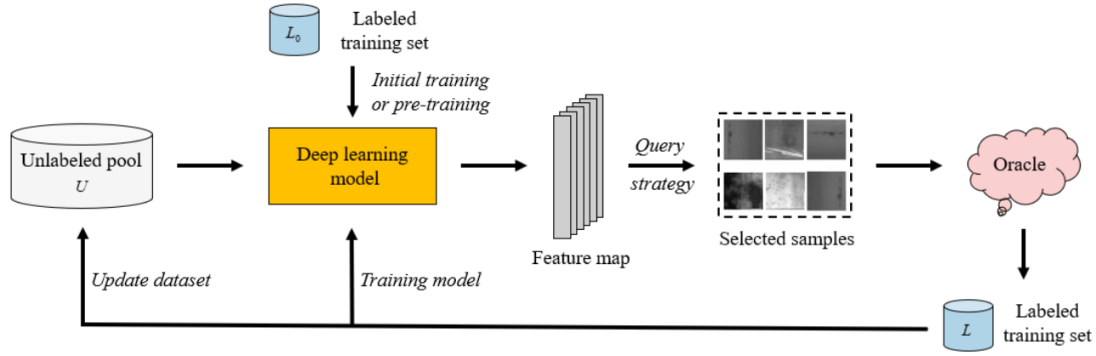
3.4.4 Çeşitlilik ağırlıklı yöntemi

[15]'te yapılan çalışmalar, belirsizlik örneklemesi ve BC stratejilerinin aykırı değerleri seçmeye yatkın olduğu sonucuna varmıştır. Bu sorunu çözmek için, Settles ve Craven [16] tarafından Bilgi Yoğunluğu (Information Density ID) yöntemi önerilmiştir. Bu yöntemin özelliği, belirsizlik örneklem tekniği ile atanan bilgilendiricilik puanlarını tartmasıdır. Pratikte bu yöntem, modelin belirsiz olduğu ancak veri setindeki diğer örneklere benzer örnekleri seçmeye çalışır [2].

4. Derin Aktif öğrenme

4.1 Derin Aktif Öğrenme kavramı

Bu bölümde, derin aktif öğrenmenin temel yaşam döngüsünü ile ilgili ele alınacak, veri etiketlemede en son teknoloji optimal sonucu elde etmek için süreçte aktif öğrenme ve derin öğrenmenin nasıl birleştirildiğini açıklayacaktır. Aşağıdaki şekil 4.1, Derin aktif öğrenme modeli mimarisinin tipik bir örneğini göstermektedir.



Şekil 4.1 Derin aktif öğrenme modeli mimarisi.

Parametre θ^1 başlatmak için derin öğrenme modeli, ilk kullanılabilir etiketli eğitim seti L_0 ile eğitilir. Daha sonra, etiketlenmemiş U havuzundaki örnekler, Derin öğrenme modeli aracılığıyla özellik çıkarma için kullanılır. Aşağıdaki adımlar, Oracle'daki etiketi sorgulamak için sorgu stratejisine dayalı olarak veri örneklerini seçer ve bunu eğitim setinin yeni bir etiketini oluşturmak için kullanır. Daha sonra derin model L etiketli veri kümesi üzerinde eğitilir ve sonuç olarak U güncellenir. Bu işlem, etiket bütçesi tükenene veya önceden tanımlanmış sonlandırma koşullarına ulaşılan kadar tekrarlanır. [1]'de bu mimari açısından özetlendiği gibi, DAL çerçevesi, etiketlenmemiş veri kümesi üzerinde Aktif Öğrenme sorgu stratejisi ve derin öğrenme modelini eğitme yöntemi olmak üzere iki ana bölüme ayrılabilir. DAL uygulaması ile ilgili çalışmaların çoğu, sorgu stratejilerinin optimizasyonuna odaklanır ve optimizasyon problemi genellikle aşağıdaki gibi ifade edilir:

$$\arg \min_{L \subseteq U, (x,y) \in L} \mathbb{E}_{(x,y)}[\ell(f(x), y)],$$

Burada:

$$L^m = \{X, Y\}$$

: geçerli etiketli veri kümesi ile X numune uzayın m örneği.

$$U^n = \{X, Y\}$$

: n örnek içeren etiketlenmemiş veri kümesi havuzu

$$P(x, y)$$

: potansiyel dağılım nerede $x \in X, y \in Y$.

$$\ell(\cdot) \in \mathbb{R}^+$$

: kayıp denklemi.

Burdaki standart ana amaç, derin model işlevi f 'yi kullanarak bir sorgu stratejisi Q tasarlamaktır. Bu $f \in \mathcal{F}, f : \mathcal{X} \rightarrow \mathcal{Y}$ bölümünün aşağıdaki kısmı, $U^n \xrightarrow{Q} L^m$, sorgu stratejilerini optimize etmek için önerilen bazı sorgulama yöntemlerini tartışacaktır.

- **Batch Mode DAL yöntemi** [2], genel Aktif Öğrenmenin karşılaştığı aşırı uyum (overfitting) sorununu çözmek için kullanılır. Çoğu Aktif Öğrenme algoritmasının tek tek $\mathcal{B} = \{x_1, x_2, \dots, x_b\} \subseteq U$ sorgulama yönteminden farklı olarak BMDAL, toplu tabanlı örnek sorgulama kullanır.

Bir grup etiketlenmemiş aday olduğunu varsayalım $\mathcal{B}^* = \{x_1^*, x_2^*, \dots, x_b^*\}$. Her edinim adımında (acquisition step), \mathcal{B} 'den aday grubu, a edinme işlevine göre puanlanır, derin model yeni numune grubunu seçmek için L üzerinde eğitilir. Bu problem [2]'de şu şekilde formüle edilmiştir:

$$\mathcal{B}^* = \arg \max_{\mathcal{B} \subseteq U} a_{batch}(\mathcal{B}, f_{\theta}(L)).$$

Toplu yönetime dayalı olarak, [17], en yüksek puanlara sahip en iyi b örneklerini elde etmek için Anlaşmazlık yoluyla Bayesian Aktif Öğrenmeyi (BALD) kullanarak sorgulamaya karar verdi. Ancak, bilgi açısından zengin benzer örnek kümeleri seçiyor gibi görüldüğü için bu yöntemin verimsizliği bildirilmiştir.

- **Belirsizliğe dayalı ve karma (hybrid) sorgu stratejisi**, başka bir sorgu iyileştirici yöntemidir. Önceki bölümlerde bahsedildiği gibi, belirsizliğe dayalı yaklaşımın uygulanması basittir ve hesaplama karmaşıklığı düşüktür. Bu nedenle destek vektör makinesi (SVM) ve KNN ile [19] birçok madencilik yönteminde [18] yaygın olarak kullanılmaktadır çünkü iyi ve doğru modeller sağlayabilirler. Bununla birlikte, çoğu durumda belirsizliğe dayalı örnekleme stratejisinin kullanılması, yığın örneklerinde verimli çeşitlilik sağlamaz (verilerdeki bilgi dağılımı tam olarak kullanılmadığından), bu da düşük veya derin öğrenme modelinin önemine yol açar [2]. Örneklerdeki daha fazla bilgiden yararlanan bir derin öğrenme modelinin performansını iyileştirmek için, toplu örnek sorgu stratejisinin daha iyi optimizasyonu için belirsizliğe dayalı hibrit ile birleştiren bir yöntem kullanılır.

4.1 Literatür taraması

Bu kısımda temel olarak Derin Aktif Öğrenme için optimizasyon yaklaşımlarına ve geliştirilmiş ve önerilen ilgili çerçeveye odaklanacaktır.

Önceki bölümlerde bahsedilenlere bakıldığında, Aktif öğrenme ile Derin öğrenme arasındaki işleme hattında var olan bazı tutarsızlıklar fark edilebilir. Bu tutarsızlıkları çözmek için birçok çerçeve önerilmiştir. AL-DL [20], DBN'li Derin Öğrenme (DÖ) modelleri için bir Aktif Öğrenme yöntemi önerirken, Aktif Derin ağları [21] ayrıca duygu sınıflandırması için aktif bir derin ağ mimarisi önerdi.

CNN ile Aktif Öğrenme algoritması, captcha tanıma için [22]'de önerilmiştir. Bu yöntemlerin çoğu bir zamanda başarıya ulaşmış olsa da ana mimarileri AÖ ve DÖ modelini eğitimi iki bağımsız problem olarak ele alan ve sonuç olarak iki farklı probleme yol açan yaklaşımları kabul edilmiştir. Bunun için, DL ve AL'nin iki görevini yakından birleştirebilecek alternatif bir çerçeve geliştirmek, DAÖ model performanslarını iyileştirecektir.

Derinlik görüntü sınıflandırma problemini çözmek için AÖ ve DÖ'yi birleştiren ilk çalışmalardan biri CEAL [23] tarafından gerçekleştirildi. Bu yeni DAÖ çerçevesi, derin evrişimli sinir ağlarını AL ile birleştirdi. Bu konsept mimarisi adım adım CNN'e etiketlenmemiş veri kümesinden örnekler göndermektir, ardından CNN sınıflandırıcı iki tür örnek üretecektir: az sayıda belirsiz örnek ve yüksek tahmin güvenilirliğine sahip çok sayıda örnek. Belirsiz örneklerin az bir kısmı oracle tarafından etiketlenir. Bu adım AL sürecinden oluşur. Bir sonraki adım, çok sayıda yüksek tahminli güvenilir örneğe otomatik olarak sahte etiketler (pseudo-labels) atayan CNN sınıflandırıcısına döner. AÖ ve DÖ'nin bu kombinasyon süreci ile birleştirilir ve CNN'ye ince ayar yapmak için iki tür numune kullanılır ve güncelleme işlemi tekrarlanır.

[24]'te önerilen HDAL adlı benzer bir çerçeve, yüz tanıma işlemleri için kullanıldı. Özellik öğrenmeyi ve Aktif Öğrenme sorgu modeli eğitimi entegre etmek için AÖ'yi derin CNN modeliyle birleştirdi.

DAL için ortak genel çerçeve kullanan diğer ilgili çalışmalar arasında, Aktif öğrenme ile [25]'te geliştirilen tam evrişimli ağı (FCN) kullanan daha özel çalışma [26]'da önerilen çerçeveydi. Tıbbi görüntü bölütleme probleminin az sayıda ek açıklama kullanarak çözülmesi önerildi. Çerçeve mimarisi süreci şu şekilde kullandı: FCN ilk olarak az sayıda etiketlenmiş veri kümesi

üzerinde eğitildi. İkinci olarak, FCN aracılığıyla etiketlenmemiş veri kümelerinden öznitelikler çıkarılır. Daha sonra bu özellikler, etiketlenmemiş örneklerin belirsizliğini ve benzerliğini tahmin etmek için kullanılır. Bu strateji, eğitimin bir sonraki aşamasında etiketlenmiş veri kümesine eklenecek oldukça belirsiz ve çeşitli örneklerin seçilmesi için yüksek bir olasılık sağlar.

Avuç içi tanıma problemini çözmek için başka bir DAL çerçevesi Aktif Avuç İzi Tanıma [27] olarak önerildi. [28]'de önerilen kusur tespiti için DAL çerçevesi, açıklama için aday örneklerin bir listesini oluşturmak amacıyla tespit modeli tarafından özelliklerin çıktısına dayalı olarak belirsizlik örnekleme gerçekleştirme içeren bir yaklaşımdır.

[2]'de belirtildiği gibi, farklı çalışma analizlerine dayalı olarak, Derin Öğrenme modelinin nihai çıktısı, örneklerin belirsizliğini veya çeşitliliğini belirlemek için temel olarak kullanılır. Ancak, derin ve sığ modeller arasındaki öğrenme paradigmalarındaki farklılık nedeniyle, geleneksel belirsizliğe dayalı sorgulama stratejisi doğrudan DL modeline uygulanamaz. Derin model, öznitelik çıkarma aşaması ve görev öğrenme aşaması olmak üzere iki aşamalı olarak kabul edilirdi. DL modelinin bu ayrı bileşimi nedeniyle, bu aşamalar tahmini etkileyeceğinden, örnek belirsizlik tahmini için kullanılamaz. Bu sorunu çözmek için AL-MV [29] önerildi. Bu mimari, CNN'nin ortasındaki farklı gizli katmanlardan gelen özelliği, her iki aşamanın belirsizliğini hesaba katarak çoklu görüntü verileri olarak ele alır. AL-MV algoritması, örnekleme belirsizliğinin daha doğru bir ölçümünü sağlamak için her katmanın belirsizliğinin uyarlanabilir ağırlıklandırmasını uygulamak üzere tasarlanmıştır.

LLAL [30] da benzer bir fikir kullanmıştır. Daha spesifik olarak, LLAL, kayıp tahmin modülünün girişi olarak hedef ağın çoklu gizli katmanlarının çıkışını kullanarak, hedef ağa bağlanmak için kayıp tahmin modülünün küçük bir parametre modülünü tasarladı. Kayıp tahmin modülü, etiketlenmemiş veri kümesinin hedef kaybını tahmin etmek için öğrenilirken, sorgu örneklerini seçmek için top-K stratejisi kullanılır. LLAL, görevden bağımsız AL çerçeve tasarımını küçük bir parametre maliyetiyle başardı ve ayrıca çeşitli ana görsel görevlerde (yani, görüntü sınıflandırma, hedef algılama ve insan pozu tahmini) rekabetçi performans elde etti. Benzer şekilde, [31] parmak kemiği segmentasyon görevleri için bir DAL çerçevesi uygulamak için benzer bir strateji kullandı.

DÖ ve AÖ'nin mevcut birleşiminde, DÖ esas olarak özellik çıkarmadan sorumludur, AL ise esas olarak örnek sorgulamadan sorumludur; bu nedenle, daha derin ve daha sıkı bir füzyon, DAÖ'nin daha iyi performans elde etmesine yardımcı olacaktır.

5. Tezin ve genel DAÖ problemi

Son literatür araştırma sonuçlarına dayanarak, DAÖ'nin derin öğrenme ve Aktif öğrenmenin ortak avantajlarını birleştirdiği bilinmektedir. DAÖ, yüksek boyutlu veri noktalarını işleme yeteneği ile verilerden özellik çıkarma (feature selection) kabiliyetini ve AÖ'nin veri açıklama maliyetlerini verimli bir şekilde azaltma potansiyelini devralır. Bu nedenle, DAÖ, bu gerçeklere dayanarak, özellikle etiketlerin yüksek düzeyde uzmanlık gerektirdiği ve elde edilmesinin zor olduğu alanlarda araştırmacıların dikkatine çekmiştir.

DAL yöntemleriyle ilgili mevcut araştırma yönü çoğunlukla aşağıdakilere odaklanır ancak bunlarla sınırlı değildir:

- AL seçim stratejilerinin iyileştirilmesi.
- Eğitim yöntemlerinin optimizasyonu
- Ve yeni DAÖ çerçevesine göre modellerin geliştirilmesi.

6. Makine öğrenmede Evrimsel Algoritmalar (EA) ve Veri İşleme

Evrimsel algoritmaların, yüksek ölçeklenebilirlik ve esneklikleri, global optimizasyon problemlerini çözme yetenekleri ve özellik seçimi, örnek seçimi ve diğer veri azaltma problemleri için çoklu kısıtlamaları optimize etme yetenekleri nedeniyle optimizasyon için güçlü avantajlar sağladığı bilinmektedir. Evrimsel Makine öğrenmenin veri analizinde daha önemli hale geldiğini kanıtlamak için birçok örnek sunulmuştur ve büyük veriler için daha fazla gelişme görmesi beklenmektedir [4].

Büyük verideki en büyük sorunları çözmek için kullanılan çoklu ve farklı yöntemlerin test edilmesi ve modern yaklaşımların değerlendirilmesine dayanarak, önerilen çözümler hala yeterli değildir. Ayrıca, mühendislik problemlerine karşı verimlilikleri için Evrimsel Algoritma olarak bilinen algoritmalar vardır, ancak bazen büyük veri işleme için hafife alınırlar. Bununla birlikte, öznelik seçimi, örnek seçimi ve eksik değerlerin ele alınması gibi bilim adamlarının çoğunlukla odaklandığı belirli veri işleme problemlerini çözmek için kullanılabilirler.

Evrimsel algoritmaları, diğer optimizasyon algoritmalarına göre birçok avantaj sağlar [4]:

- Çok Boyutlu problemleri başarılı bir şekilde çözebilirler (örneğin genetik algoritmalar).
- Çeşitli arama uzaylarında kullanıldıklarında, ilgili öznitelik vektörlerini veya bilgilendirici örnekleri bulma yeteneğine sahiptirler.
- Farklı türden birkaç kısıtlamayı tek tek optimize edebilirler.

Bu nedenle, bu önemli özelliklere ve belirtilmeyen diğer yardımcı programlara dayanarak, EA veri işleme için daha faydalı olabilir ve bu nedenle evrimsel makine öğrenimi çerçevesi ve araçlarında veriye dayalı model optimizasyonunu geliştirmek için kullanılabilir. Aşağıdaki tablo, büyük verinin farklı konularında EA 'nın temel avantajlarını ayrıntılı olarak sunmaktadır.

Büyük veri problemi		EA kullanışlı kalitesi
Özellik seçimi		<ul style="list-style-type: none"> - Bir özellik vektörünün ikili gösterimi nedeniyle özelliklerin alt sistemlerini karşılaştırma imkânı. - EA'lar, özellik alt sistemlerinin yüksek boyutlu arama uzayında etkin bir şekilde çalışır. - EA'lar aynı anda bir dizi kriteri optimize edebilir.
Örnek seçimi	Eğitim verisi rotasyonu	<ul style="list-style-type: none"> - Evrimsel Algoritmaların popülasyon tabanlı tabiatına bağlı hesaplamaları paralelleştirme imkânı. - Birkaç paralel popülasyonla işbirlikçi modifikasyonlar, veri rotasyon ile kullanılabilir
	Uyarlanabilir örnek seçimi	<ul style="list-style-type: none"> - Popülasyon tabanlı yöntemler için örnek seçimi geliştirilmiştir.

		- Örnek seçimi, bir modelin eğitimine kolayca yerleştirilebilir
Eksik değerlerin ele alınması		- Eksik değerleri işleyebilen birkaç model yalnızca EA'lar tarafından oluşturulabilir

Tablo 6.1: Evrimsel Algoritmaların farklı büyük veri konularında ana avantajları.

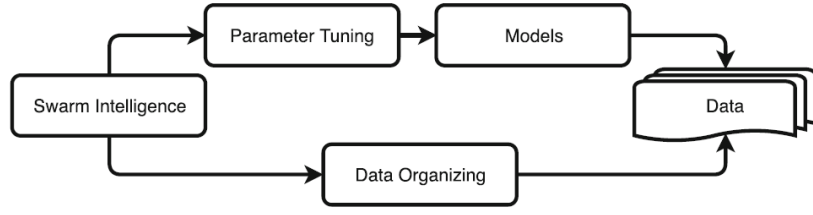
7. Sürü zekâ algoritmalar ve Makine öğrenme

Veri bilimi alanındaki ana görev, çoğunlukla veri işleme ve analitik üzerine odaklanır. Daha iyi kararlar için büyük verilerden değerli bilgiler elde etmeyi amaçlar. Bununla birlikte, veri işleme için kullanılan geleneksel yaklaşımlar, verimsizliklerine neden olan çeşitli sorunlarla karşılaşmaktadır. Bu nedenle, veri analitiği ile ilgili daha zor problemleri çözmek için başka yeni etkili ve verimli yöntemler geliştirilmelidir [34].

Sürünün doğasından ve biyolojik davranışından ilham alan sürü zekâ (SZ) algoritmalarının, karmaşık optimizasyon problemlerini çözmek için başarılı ve kapsamlı bir yaklaşım sağladığı kanıtlanmıştır. Bu algoritmalar, biyolojik bir sürünün ortaya çıkan davranışlarını taklit eder ve simüle eder.

Mevcut ve bilinen geleneksel veri yönetimi ve analizi yöntemleri, yalnızca sürekli ve farklı işlemlere uygulanabilir ve bu da onları çok büyük ve karmaşık miktarda veriye karşı kullanılamaz hale getirir [5]. Bu yöntemlerden farklı olarak, SZ algoritmaları, veri ile ilgili karmaşık problemleri çözmede iyi olan meta-sezgisel ve popülasyon tabanlı yaklaşımlardır ve veri bilimi alanında ilgili bir görev için büyük potansiyele sahip olduklarını kanıtlamıştır. Veri analitiğinde büyük ölçekli, dinamik ve çok amaçlı sorunları çözmesinde kullanılabilir [35].

Şekil 7.1'de gösterildiği gibi, veri analitiği için SZ algoritmalarının kullanıldığı iki ana kategori vardır. SZ, veri madenciliği teknolojilerinin parametre ayarlama/optimize etme yöntemi olarak kullanılır veya uygun bir kümelemeye ulaşmak veya verinin boyutluluğunu azaltmak için veri örneklerini düşük boyutlu bir özellik uzayı üzerinde hareket ettirmek gibi veri organizasyonunda uygulanır [5].



Şekil 7.1 Veri bilimi için sürü zekâ 'ın iki yaklaşımı

Sürü zekâ, düşük karmaşıklıkta bireylerden oluşan bir popülasyonda kolektif zekâyı inceleyen, doğadan ilham alan bir arama ve optimizasyon teknikleri grubudur [35]. SZ algoritmaları, sürü popülasyonunun problemlerini çözmek için kullandığı akıllı teknikleri simüle eder. SZ genellikle, akıllı organizasyon, rekabet, esneklik ve popülasyon içindeki iş birliği özelliklerini içeren bir grup veya birkaç birey grubuyla bireyler arasındaki etkileşimlerden ilham alır [36].

7.1 SI algoritmalarının genel prosedürü

Daha önce açıklandığı gibi SZ algoritmaları, sürülerin biyolojik ve akıllı davranışlarından ilham alır. İyi ölçeklenebilirlik, uyarlanabilirlik, sağlamlık ve bireysel basitlik ile kendi kendine organizasyon ve merkezi olmayan kontrol ile karakterize edilirler.

Herhangi bir SZ algoritmasında, sürü popülasyonu, eldeki problem için potansiyel çözümleri gösteren bireylerden oluşur. Bu adaylar (çözümler) birbirleriyle rekabet eder ve kendi aralarında iş birliği yapar, daha sonra istatistiksel olarak ve potansiyel olarak yineleme üzerinden daha iyi çözümler haline gelir [37]. SZ algoritmalarının genel prosedürü aşağıdaki şekilde temsil edilir:

Algorithm 1. General procedure of swarm intelligence algorithms

- 1 Population Initialization: Generate random solutions for an optimized problem, repair solutions if solutions violate any of the constraints;
- 2 Evaluate all initialized individuals;
- 3 **while not terminated do**
- 4 Reproduce individuals to form a new population;
- 5 Evaluate the fitness of each solution;
- 6 Select solutions with better fitness values;
- 7 Update solutions in the archive;
- Result:** Relatively good solution(s)

Şekil 7.2 Sürü zekâsı algoritmalarının genel prosedürü

Tablo 02'de özetlendiği gibi, son yıllarda insan toplumu, hayvanlar, böcekler, piller vb. dahil olmak üzere çok sayıda ve farklı ilham kaynaklarına dayalı olarak birçok sürü zekâ algoritmalar geliştirilmiştir.

Algoritmalar	Algoritmaların İlham kaynağı	
Beyin Fırtınası Opt., (Brainstorm Opt.) BSO [38] Teaching-learning-based Opt. (Öğretme-öğrenme-tabanlı), TLBO [39]	İnsanın beyin fırtınası süreci Öğretmenler öğrencileri nasıl etkiler	İnsan toplumu
Particle Swarm Opt. (Parçacık Sürüsü Opt), PSO [40] Ant Colony Opt. (Karıncı Kolonisi Opt.), ACO [41] Artificial Bee Colony (Yapay Arı Kolonisi), ABC [42] Firefly Algorithms FA (Ateşböceği Algoritmaları), [43] Glowworm Swarm Opt., GSO (Glowworm Swarm Opt) [44] Cuckoo Search, CS (Guguk Kuşu Arama) [45]	Kuş sürüsü ve yiyecek arama Karıncalar yiyecek arama mekanizmaları Bal arılarının yiyecek arama davranışı Ateş böceklerinin biyoluminesansı Luciferin tarafından uyarılan bir ateş böceği parıltısı Guguk kuşlarında zorunlu kuluçka parazitizmi	Böcekler, kuşlar vb.
Bat Algorithm, BA (Yarasa Algoritması) [46] Grey Wolf Opt., GWO (Gri Kurt Opt) [47]	Mikro yarasaların ekolokasyon davranışları Gri kurtların liderliği ve avı Maymunların kullandığı tırmanma teknikleri	Hayvanlar

Monkey Algorithm, MA (Maymun Algoritması) [48]	Aslanların işbirliği özellikleri	
Lion Opt. Algorithm, LOA (Aslan Op. Algoritma) [49]		
Bacterial Foraging Opt, BFO (Bakteriyel Toplayıcılık) [50]	Bakterilerin grup arama davranışı	Mikroskobik
Fireworks Algorithm, FWA (Havai Fişek Algoritması) [51]	Havai fişek patlaması	Diğer

Tablo 7.1: İlham kaynağı olan bazı sürü zekâsı algoritmaları

SZ algoritmaları, neredeyse tüm bilim, mühendislik ve endüstriyel alanlardaki doğrusal olmayan problemleri çözmenin pratik uygulamasında uygulanabilirliğini bulmuştur: Veri madenciliğinden optimizasyona, hesaplama zekasına, iş planlamasına, biyoinformatik ve endüstriyel uygulamalara [5]. Bu bilimsel ve mühendislik sorunları az veya çok veri ile ilgili sorunlar olduğundan, veri işleme sorunlarını çözmek için SZ algoritmalarını entegre edebilen yeni yöntemler ve yaklaşımlar geliştirmek geri bırakılmamalıdır.

“Sürü zekâsı, veri ile ilgili uygulamalarda birçok başarılı uygulama ulaştı. Bu arada, artan dinamikler, gürültüler (noises) ve işlemlerin karmaşıklığı ile birlikte, veri bilimleri ile sürü zekâsı uygulamalarında zorluklarla birlikte hala birçok fırsat var” [5]

8. Araştırmanın amaçları

Bu araştırma, DAL uygulamasına aşağıdakiler gibi yeni özellikler eklemeyi amaçlamaktadır:

- Doğadan ilham alan bir algoritma tarafından desteklenen yeni bir Derin aktif öğrenme (DAÖ) çerçevesinin geliştirilmesi ve iyileştirilmesi.
- Genetik algoritma (GA), Yapay Arı Kolonisi (ABC) algoritması ve Parçacık Sürü Optimizasyonu (PSO) vb. gibi sürü zekâ algoritmalarına dayalı yeni en iyi DAÖ çerçevesi önerilmesi ve makine öğrenme modeli performansındaki etkinliklerini karşılaştırılması.
- Veri bilimi ve makine öğrenme alanındaki veri etiketleme sürecini azaltacak DAÖ mimarisini önerilmesi.

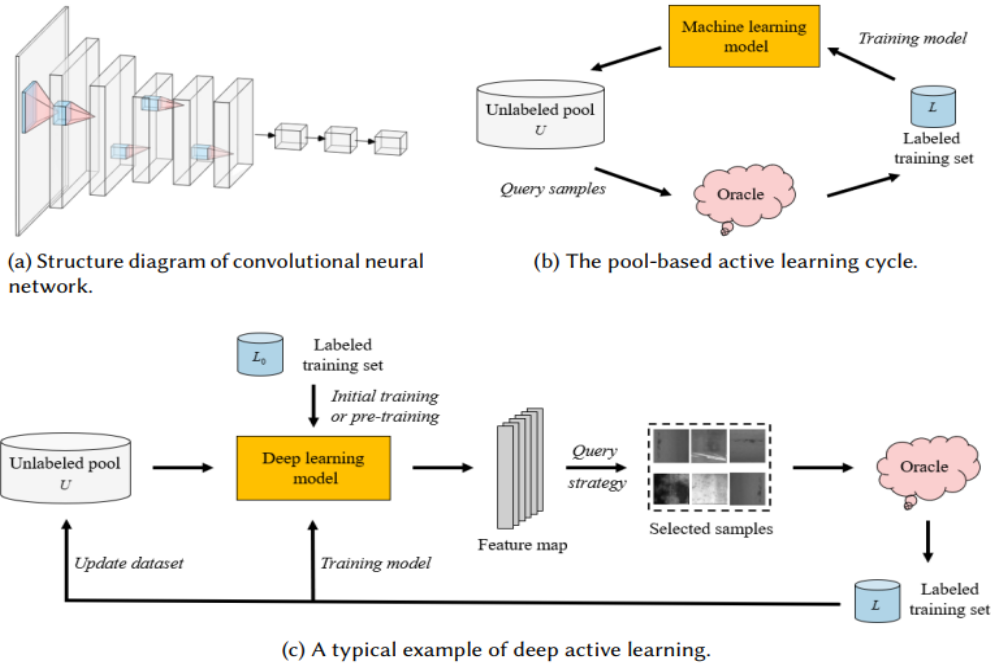
- Önerilen yeni çerçeveyi görüntüler, videolar, metinler gibi çok sayıda ve farklı türdeki gerçek dünya veri kümelerinde uygulama çalışılması ve makine verisi etiketleme yaşam döngüsündeki etkinliğini ve performansını değerlendirilmesi.
- Seçilen özellikleri optimize etmek için özellik seçim teknikleri ve doğadan ilham alan algoritmalar sürecinden (örn. evrimsel ve sürüler) yararlanan hibrit bir yaklaşımla veri açıklamasında özellik seçimi optimizasyonunu uygulayabilen yeni bir Derin aktif öğrenme çerçevesi sunulması.

9. Önerilen Çerçeve

Şekil 8.1, genel DAL çerçevelerinin bir gelişimi olarak bu araştırmada izlenecek çerçevenin genel yapısını göstermektedir. Bu çerçeve, evrimsel algoritmaların (EA) ve sürü akıllı algoritmalarının (SZ) DAL sürecine entegrasyonundan oluşacaktır. Örnek olarak, EA'dan Genetik algoritma ve SZ'dan Yapay Arı Kolonisi seçilmiştir. Bu algoritmalar, Derin öğrenme modeli için en iyi özellik setinin seçimini ve giriş veri setine dayalı Aktif öğrenme bölümü için en iyi örneklerin seçimini kolaylaştırmak için DAL sürecine entegre edilecektir.

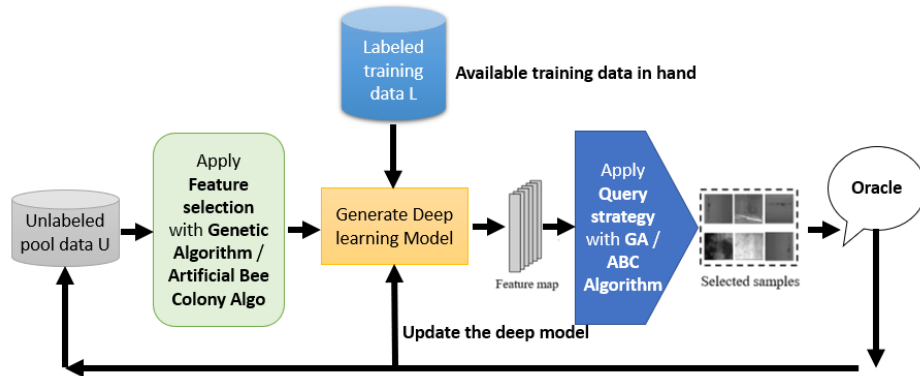
DAÖ'nin ana genel çerçeve süreci [1]'de şekil 9.1'deki gibi sunulmuştur (a) Yaygın bir DÖ modeli: Evrişimli Sınır Ağı. (b) Havuz tabanlı AÖ döngüsü: Etiketlenmemiş U havuzundaki numuneyi sorgulamak ve etiketleme için oracle'ya teslim etmek için sorgulama stratejisini kullanır, ardından sorgulanan numuneyi etiketli eğitim veri kümesi L'ye eklenmesi ve eğitilmesi ve ardından bir sonraki sorgulama turu için yeni öğrenilen bilgileri kullanması. Etiket bütçesi tükenene veya önceden tanımlanmış durdurma koşullarına ulaşılan kadar bu işlemi tekrarlanması. (c) Tipik bir DAÖ örneği: DÖ modelinin parametreleri θ , L_0 etiket eğitim setinde inisyalize edilir veya önceden eğitilir ve etiketlenmemiş U havuzunun örnekleri, DÖ modeli aracılığıyla özellikleri çıkarmak için kullanılır.

Ardından, ilgili sorgu stratejisine göre örnekleri seçilesi ve yeni bir etiket eğitim seti L oluşturmak için sorgulamada etiketi sorgulanması, ardından DÖ modelini L üzerinde eğitilmesi ve aynı anda U'yu güncellenmesi. Etiket bütçesi tükenene veya önceden tanımlanmış sonlandırma koşullarına ulaşılan kadar bu işlemi tekrarlanır.” [1]



9.1 DÖ, AÖ ve DAÖ'nin tipik mimarilerinin karşılaştırılması.

Bu araştırma çalışmasında önerilen yeni geliştirilmiş çerçeve, şekil 9.1'de sunulan çerçeveye dayanacaktır. Mevcut literatürde, AI model eğitimi için veri açıklamalarında derin aktif öğrenmeyi evrimsel veya sürü akıllı algoritmalarla bir araya getiren bir çalışmaya henüz çalışılmamıştır. Bu çalışmanın amacı, Derin öğrenme için özellik seçimini ve Aktif öğrenme için örnek sorgulamayı optimize etmek için genetik algoritmayı (GA) ve / veya Yapay Arı Kolonisi algoritmasını DAÖ sürecine entegre etmeye çalışacaktır. Çerçeve aşağıdaki şekilde gösterilmiştir:



Şekil 9.2 Evrimsel ve sürü zekâsı algoritmasıyla entegre yeni Derin Aktif öğrenme çerçevesi

10. Metodoloji ve Uygulama

Bu araştırmanın amaçlarına ulaşmak için belirli bir metodoloji gereklidir. Herhangi bir makine öğrenme projesinin başarılı olması için, algoritma çerçevesi iyi tasarlanmalı ve hedef çalışma alanıyla ilgili mevcut verilerle eğitilmelidir. Ancak bu çalışmada kullanılacak olan etki alanının kapsamı daha sonra DAL çerçevesi ile işlenecek verilerin mevcudiyetine göre belirlenecektir.

Python, makine öğreniminde kullanılan en güçlü programlama dillerinden biridir. Python, gerekli tüm yerleşik kitaplıklarıyla birlikte önerilen çerçeveyi geliştirmek için kullanılacaktır.

11. Yazım planı

- Giriş
- Genel bilgi
- Geliştirme ve uygulama
- Sonuçların değerlendirilmesi
- Karşılaştırma analizi
- Tartışma

12. Çalışma takvimi

Bu tezin 24 ay içinde tamamlanması planlanmaktadır. İlk altı ay, literatür taramasının analizine ve DAL ile ilgili teori bilgilerinin derinlemesine anlaşılmasına ve bunların evrimsel algoritmalar ve sürü akıllı algoritmalarla olası entegrasyonuna odaklanacak; gerçek dünya alanıyla ilgili veri toplama ve bunların işlenmesi ikinci altı ay içinde gerçekleştirilecektir; sonraki altı ay için önerilen yeni çerçevenin uygulanması ve simülasyonu yapılacaktır; son altı ay, geliştirme raporlarının yazılması ve tez yazımının tamamlanması sonuçlarının değerlendirilmesine odaklanmalıdır.

Gelişme raporları dönemi	Yapılacak iş	Süre
Birinci gelişme raporu dönemi	Literatür analizi ve teorik çerçevenin yazımı	6 ay
İkinci gelişme raporu dönemi	Verilerin analizi ve bir kaç yazılım yardımıyla simülasyon yapılması	6 ay

Üçüncü gelişme raporu dönemi	Sonuçların değerlendirilmesi ve çıktılarından makale hazırlanması	6 ay
Dördüncü gelişme raporu dönemi	Tez raporunun yazım kurallarına göre yazımı ve teslimi	6 ay
Toplam		24 ay

Kaynaklar

1. Pengzhen R. et al Survey of Deep Active Learning 2020.
2. Kevin De Angeli et al. Deep active learning for classifying cancer pathology reports. 2021.
3. Burr Settles Active learning Literature review survey Computer Sciences Technical Report 1648, University of Wisconsin–Madison Updated on: January 26, 2010.
4. Vladimir St. et al. Why don't you use Evolutionary algorithms in Big Data, IOP Conference, International workshop on Mathematical Model and their Application.
5. Yang J. et al. (2020) Swarm Intelligence in Data Science: Applications, Opportunities and Challenges. In: Tan Y., Shi Y., Tuba M. (eds) Advances in Swarm Intelligence. Lecture Notes in Computer Science, vol 12145. Springer, Cham. https://doi.org/10.1007/978-3-030-53956-6_1, ICSI 2020
6. Settles, Burr, Active learning literature survey. University of Wisconsin-Maison Department of Computer Science, 2009.
7. Frederic Brenton Fitch. 1944. McCulloch Warren S. and Pitts Walter. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics , vol. 5 (1943), pp. 115–133. Journal of Symbolic Logic 9 (1944), 49–50.
8. Alawad M, Gao S, Qiu JX, Yoon HJ, Blair Christian J, Penberthy L, Mumphrey B, Wu X-C, Coyle L, Tourassi G. Automatic extraction of cancer registry reportable information from free-text pathology reports using multitask convolutional neural networks. J Am Med Inform Assoc., 27(1):89–98. <https://doi.org/10.1093/jamia/ocz153>, 2019
9. Gao S, Qiu JX, Alawad M, Hinkle JD, Schaefferkoetter N, Yoon H-J, Christian B, Fearn PA, Penberthy L, Wu X-C, Coyle L, Tourassi G, Ramanathan A. Classifying cancer pathology reports with hierarchical self-attention networks. Artif Intell Med. 2019;101:101726. <https://doi.org/10.1016/j.artmed.2019.101726>, 2019

10. Young T, Hazarika D, Poria S, Cambria E., Recent trends in deep learning based natural language processing, arXiv:1708.02709, 2017
11. Lewis DD, Gale WA., A Sequential Algorithm for Training Text Classifiers, arXiv:cmp-lg/9407020, 1994
12. Schein AI, Ungar LH. Active learning for logistic regression: an evaluation. Mach Learn, 68:235–65. <https://doi.org/10.1007/s10994-007-5019-5>, 2007
13. Shannon CE. A mathematical theory of communication. Bell Syst Tech J., 27:379–423, 1948
14. Seung HS, Oppor M, Sompolinsky H. Query by committee. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. COLT'92. Association for Computing Machinery, New York, NY, USA. 1992, p. 287–94. <https://doi.org/10.1145/130385.130417>., 1992
15. Roy N, McCallum A. Toward optimal active learning through sampling estimation of error reduction. In: Proceedings of the Eighteenth International Conference on Machine Learning. ICML'01. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA., p. 441–8, 2001
16. Settles B, Craven M. An analysis of active learning strategies for sequence labeling tasks. In: Proceedings of the conference on empirical methods in natural language processing. EMNLP '08. Association for Computational Linguistics, USA., p. 1070–1079, 2008
17. Neil Houlsby, Ferenc Huszar, Zoubin Ghahramani, and Mate Lengyel. 2011. Bayesian Active Learning for Classification and Preference Learning. arXiv: Machine Learning, 2011.
18. Simon Tong and Daphne Koller, Support vector machine active learning with applications to text classification, Journal of Machine Learning Research 2, 1, 45–66 2002,
19. Prateek Jain and Ashish Kapoor, Active learning for large multi-class problems. (2009), 762–769, 2009
20. Dan Wang and Yi Shang. 2014. A new active labeling method for deep learning, 112–119, 2014
21. Shusen Zhou, Qingcai Chen, and Xiaolong Wang. 2010. Active Deep Networks for Semi-Supervised Sentiment Classification, 1515–1523, 2010
22. Fabian Stark, Caner Hazırbaş, Rudolph Triebel, and Daniel Cremers, Captcha recognition with active deep learning. In Workshop new challenges in neural computation, Vol. 2015. Citeseer, 94, 2015

23. K. Wang, D. Zhang, Y. Li, R. Zhang, and L. Lin. Cost-Effective Active Learning for Deep Image Classification. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 12, 2591–2600, 2017
24. Ya Li, Keze Wang, Lin Nie, and Qing Wang, Face Recognition via Heuristic Deep Active Learning. (2017), 97–107, 2017
25. Jonathan Long, Evan Shelhamer, and Trevor Darrell, Fully convolutional networks for semantic segmentation. (2015), 3431–3440, 2015
26. Lin Yang, Yizhe Zhang, Jianxu Chen, Siyuan Zhang, and Danny Z Chen, Suggestive Annotation: A Deep Active Learning Framework for Biomedical Image Segmentation. (2017), 399–407, 2017
27. Xuefeng Du, Dexing Zhong, and Huikai Shao, Building an Active Palmprint Recognition System. In 2019 IEEE International Conference on Image Processing, ICIP 2019, Taipei, Taiwan, September 22-25, 2019. 1685–1689, 2019
28. Xiaoming Lv, Fajie Duan, Jiajia Jiang, Xiao Fu, and Lin Gan, Deep Active Learning for Surface Defect Detection. *Sensors* 20, 6 (2020), 1650, 2020
29. Tao He, Xiaoming Jin, Guiguang Ding, Lan Yi, and Chenggang Yan, Towards Better Uncertainty Sampling: Active Learning with Multiple Views for Deep Convolutional Neural Network. (2019), 1360–1365, 2019
30. Donggeun Yoo and In So Kweon. 2019. Learning Loss for Active Learning, 93–102, 2019
31. Ziyuan Z. et al, Deeply Supervised Active Learning for Finger Bones Segmentation. *arxiv*, 2020.
32. Schein AI, Ungar LH. Active learning for logistic regression: an evaluation. *Mach Learn.*, 68:235–65. <https://doi.org/10.1007/s10994-007-5019-5>, 2007
33. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.*, 27:379–423. 1948
34. Cheng, S., Shi, Y., Qin, Q., Bai, R.: Swarm intelligence in big data analytics. In: Yin, H., et al. (eds.) *IDEAL 2013. LNCS*, vol. 8206, pp. 417–426. Springer, Heidelberg, https://doi.org/10.1007/978-3-642-41278-3_51, 2013
35. Martens, D., Baesens, B., Fawcett, T.: Editorial survey: swarm intelligence for data mining. *Mach. Learn.* 82(1), 1–42, <https://doi.org/10.1007/s10994-010-5216-5>, 2011
36. Kennedy, J., Eberhart, R., Shi, Y.: *Swarm Intelligence*. Morgan Kaufmann Publisher, San Francisco, 2001
37. Eberhart, R.C., Shi, Y., Kennedy, J.: *Swarm Intelligence*. Elsevier, London, 2001

38. Panigrahi, B.K., Shi, Y., Lim, M.H.: Handbook of Swarm Intelligence: Concepts, Principles and Applications, vol. 8. Springer, Heidelberg , [https://doi.org/ 10.1007/978-3-642-17390-5](https://doi.org/10.1007/978-3-642-17390-5), 2011
39. Shi, Y.: Brain storm optimization algorithm. In: Tan, Y., Shi, Y., Chai, Y., Wang, G. (eds.) ICSI, LNCS, vol. 6728, pp. 303–309. Springer, Heidelberg ,[https://doi.org/10.1007/978-3-642-21515-5 36](https://doi.org/10.1007/978-3-642-21515-5_36), 2011.
40. Rao, R.V., Savsani, V.J., Vakharia, D.: Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Comput. Aided Des.* 43(3), 303–315, 2011.
41. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of ICNN 1995-International Conference on Neural Networks, vol. 4, pp. 1942–1948. IEEE, 1995.
42. Dorigo, M., Birattari, M., Stutzle, T.: Ant colony optimization. *IEEE Comput. Intell. Mag.* 1(4), 28–39, 2006.
43. Karaboga, D., Alper B, Bahriye A.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J. Global Optim.* 39(3), 459–471, <https://doi.org/10.1007/s10898-007-9149-x>, 2007
44. Yang, X.-S.: Firefly algorithms for multimodal optimization. In: Watanabe, O., Zeugmann, T. (eds.) SAGA 2009. LNCS, vol. 5792, pp. 169–178. Springer, Heidelberg, [https://doi.org/10.1007/978-3-642-04944-6 14](https://doi.org/10.1007/978-3-642-04944-6_14), 2009
45. Krishnanand, K., Ghose, D.: Glowworm swarm optimization for simultaneous capture of multiple local optima of multimodal functions. *Swarm Intell.* 3(2), 87–124, 2009.
46. Yang, X.S., Deb, S.: Cuckoo search via l’evy flights. In: 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC), pp. 210–214. IEEE, 2009
47. Yang, X.S., Gandomi, A.H.: Bat algorithm: a novel approach for global engineering optimization. *Eng. Comput.* 29(5), 464–483, 2012.
48. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. *Adv. Eng. Softw.* 69, 46–61, 2014.
49. Zhao, R.Q., Tang, W.S.: Monkey algorithm for global numerical optimization. *J. Uncertain Syst.* 2(3), 165–176, 2008,
50. Yazdani, M., Jolai, F.: Lion optimization algorithm (LOA): a nature-inspired metaheuristic algorithm. *J. Comput. Des. Eng.* 3(1), 24–36, 2016.
51. Das, S., Biswas, A., Dasgupta, S., Abraham, A.: Bacterial foraging optimization algorithm: theoretical foundations, analysis, and applications. In: Abraham, A., Hassanien, A.E.,

- Siarry, P., Engelbrecht, A. (eds.) Foundations of Computational Intelligence, vol. 3, pp. 23–55. Springer, Heidelberg. https://doi.org/10.1007/978-3-642-01085-9_2, 2009
52. Tan, Y., Zhu, Y.: Fireworks algorithm for optimization. In: Tan, Y., Shi, Y., Tan, K.C. (eds.) ICSI 2010. LNCS, vol. 6145, pp. 355–364. Springer, Heidelberg. https://doi.org/10.1007/978-3-642-13495-1_44, 2010

Maliki MOUSTAPHA (Öğrenci)

Doç. Dr. Celal ÖZTÜRK (Danışman)