



Yapay Zekâ Yöntemlerinde Entropi

Muhammed Ali KOŞAN^{a,*}, Aysun COŞKUN^b, Hacer KARACAN^c

^{a,*} Muş Alparslan Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, MUŞ, TÜRKİYE

^b Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, ANKARA, TÜRKİYE

^c Gazi Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, ANKARA, TÜRKİYE

MAKALE BİLGİSİ

Alınma: 29.10.2019
Kabul: 25.12.2019

Anahtar Kelimeler:
Bilginin Düzensizliği,
Entropi,
Yapay Zekâ

***Sorumlu Yazar**
e-posta:
ceo.muhammed@gmail.com

ÖZET

Bilginin tarihsel süreçteki kapasitesinin değişimi, bu bilgileri yönetmek ve yönlendirmek için oluşturulan sistemlerin gelişimi gibi birçok nokta dikkate alındığında, bilginin hem ölçümünün hem kalitesinin hem de anlaşıldırılmasının önemi daha iyi anlaşılmaktadır. Claude E. Shannon tarafından ortaya atılan Bilgi Kuramı da bu noktada bilginin kontrollü yönetimi için çığır açmıştır. Devamında gelişen birçok yöntem tarafından bilginin belirsizliği için kullanılan entropi kavramı ise Shannon tarafından ortaya atılan özellikler temel alınarak geliştirilmiştir. Bu bağlamda bilginin ölçümünde düzensizliğin ölçütü olan entropi birçok alanda önem arz etmektedir. Son yıllarda hızla büyüyen Yapay Zekâ alanı ise bunlardan biridir. Yapay Zekâ özellikle Büyük Veri ve Derin Öğrenme alanlarının gelişimi ile daha büyük veriler üzerinde işlem yapılabilir bir alan haline gelmiştir. Bu çalışmamızda Yapay Zekâ alanındaki yöntemlerden bazılarında kullanılan entropi kavramı üzerine bir inceleme çalışması yapılmıştır. Sonuç olarak temel mantık ve kavram açısından aynı olmakla birlikte yöntemsel uygulamada farklılıklar gözlemlenmiştir. Çalışmanın ana hedeflerinden biri de incelenen entropi ve yapay zekâ yöntemleri bağlamında yeni yöntemlerin geliştirilmesine ön ayak olmak için bir bakış açısı kazandırmaktır.

Entropy in Artificial Intelligence Methods

ARTICLE INFO

Received: 29.10.2019
Accepted: 25.12.2019

Keywords:
Disorder of Information,
Entropy,
Artificial Intelligence

***Corresponding Author**
e-mail:
ceo.muhammed@gmail.com

ABSTRACT

The importance of measurement and quality of information, as well as attributing a meaning to information, is understood better by considering many aspects such as the historical change in the capacity of information and the development of systems to manage and guide such information. Information Theory, put forward by Claude E. Shannon, has revolutionized the controlled management of knowledge. The concept of entropy, which is used for the uncertainty of information by many methods, has been developed on the basis of features introduced by Shannon. In this context, entropy, which is a measure of irregularity in the measurement of information, is important in many areas. In recent years, rapidly growing Artificial Intelligence field is one of such areas. Artificial Intelligence has become processable on larger data, especially with the development of Big Data and Deep Learning areas. In this study, the use of entropy concept in some of the methods in Artificial Intelligence field is investigated. As a result, it is observed that although there is uniformity in terms of basic logic and concept, there are differences in methodological application. One of the main objectives of the study is to gain an insight to develop novel methods in the context of entropy and artificial intelligence.

1. GİRİŞ (INTRODUCTION)

Birçok alanda pek çok farklı sistem ve bu sistemler tarafından üretilen devasa bilgiler bulunmaktadır. Var olan sistemlerin duruma göre düzensizliğinin ölçülmesi gerekir. İşte tam bu noktada entropi kavramı devreye girmektedir. İlk olarak Alman bilim adamı Rudolf CLAUSIUS tarafından 1865 yılında tanımı yapılmıştır [1]. Bu kavramın ilk kez kullanıldığı termodinamik alanının ikinci yasasına göre entropi zamanla azalmaz.

$$\partial S = \partial \frac{Q}{T} \quad (1)$$

Devamında 1872 ve 1878 yılları arasında Ludwig BOLTZMAN entropiyi istatistiksel olarak yorumlamıştır [2].

$$S = k_b (\ln W) \quad (2)$$

1880 yıllarında Josiah Willard GIBBS istatistik mekaniği üzerine çalışmalarında “Termodinamik denge maksimum entropiye karşılık gelir” tanımını yapmıştır [3]. Devamında 1924 yılında Harry NYQUIST “Telgraf Hızını Etkileyen Bazı Faktörler” adlı makalesinde telgraf üzerinden maksimum hızda ve bozulma olmadan mesaj göndermenin yolunu açıklamıştır [4]. Bu noktada kullanılan devreye göre iletişim kanalının bir limitinin olduğu belirlenmiştir. 1928 yılında ise Ralph HARTLEY’in “Bilginin İletimi” adlı çalışmasında bilgi ilk defa ölçülmeye çalışılmıştır [5]. Ölçüm yöntemi olarak s karakterlik bir alfabede n uzunluğundaki bir mesajın içerisindeki bilgi miktarına ait formül aşağıdaki şekilde belirlenmiştir.

$$H_H(s^n) = \log s^n = n \log s \quad (3)$$

Bilgi Kuramının babası olarak kabul edilen Claude E. SHANNON; NYQUIST ve HARTLEY’in kuramlarını genişleterek 1948 yılında “İletişimin Matematiksel Teorisi” [6] adlı çalışmayı yapmıştır. Bu çalışmaya göre Bilgi Teorisindeki entropinin tanımı yapılmış ve özellikleri belirtilmiştir.

Bu çalışmamızda ilk olarak entropinin temel yapısı ve özelliklerinden 2. bölümde bahsedilmektedir. Sonrasında, entropinin Yapay Zekâ tekniklerindeki kullanım çeşitleri üzerine 3. bölümde bir araştırma ve değerlendirme sunulmaktadır. Son olarak yapılan araştırma ve değerlendirme sonuçları 4. bölümde tartışılmaktadır.

2. ENTROPİ (ENTROPY)

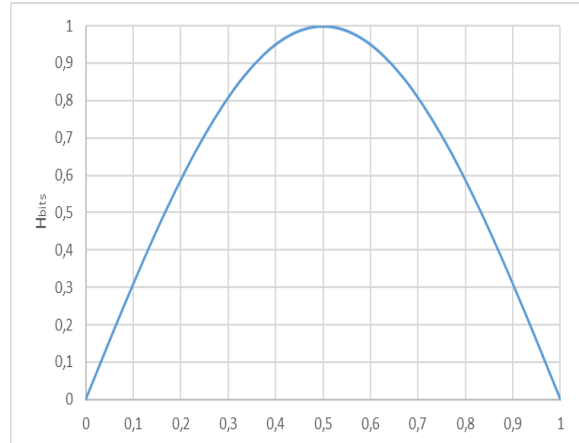
Entropi tanımlandığı her alandaki düzensizliğin ölçülmesi için kullanılmaktadır. Bilgi Teorisindeki tanımı ise “Bir veri kaynağındaki belirsizlik (rastgelelik) miktarını temsil eden bit sayısıdır” [6].

2.1. Bilgi Kuramında Entropi (Entropy in Information Theory)

Bilgi Kuramında entropi H sembolüyle gösterilmektedir. H, bir bilginin bir ortamdaki düzensizliğini belirlemek için kullanılmaktadır. Bu noktada bilginin aynı tek düze yapıda akması düzensizliği ortadan kaldırır. Entropi formülünde n verinin miktarını ve p_i ise ilgili verinin olma olasılığını ifade etmektedir.

$$H = - \sum_{i=1}^n p_i \log(p_i) \quad (4)$$

Yukarıda sunulan formül ve tanımlar temel alındığında, p ve 1-p olasılığa sahip iki durum varsayıldığında entropinin maksimum olduğu nokta iki olasılığın eşit olduğu durumdur. Şekil 1’de görüldüğü üzere $p=0.5$ olduğu durumda belirsizlik maksimumdur. Herhangi bir durumun olma olasılığı arttıkça belirsizlik ortadan kalkmaya başlamaktadır.



Şekil 1. p ve 1-p olasılığına sahip iki durumun H bitlik veri miktarındaki entropi değişim örneği
(Example of entropy change in the H-bit data amount of two states with the probability p and 1-p)

2.2. Entropinin Özellikleri (Properties of Entropy)

Bilgi Teorisinde entropi tanımlandığında bazı özelliklere göre çerçevesi belirtilmiştir. Bu özellikler sırası ile aşağıda açıklanmaktadır [6].

Özellik-1: $H=0$ ise bir durum dışındaki tüm durumların olasılığı sıfırdır. Geriye kalan bir durum

ise tüm olasılıkların tamamına sahiptir. Aksi durumda H pozitifdir. (Örneğin hileli bir zarın sürekli altı gelmesi)

Özellik-2: Belirli bir n için tüm p_i ler eşit olduğunda (tamamı 1/n) H maksimum değerinde olur ve $\log(n)^e$ eşit olur. Bu da sezgisel olarak en belirsiz durumdur.

Özellik-3: x ve y olarak iki olay olsun ve bunların olasılıkları sırası ile m ve n söz konusu olduğunda;

- İlki için i ve ikinci için j ortak oluşumunun olasılığı $p(i,j)$ 'dir. Bu ortak oluşumun entropisi;

$$H(x,y) = - \sum_{i,j} p(i,j) \log p(i,j) \quad (5)$$

$$H(x) = - \sum_{i,j} p(i,j) \log \sum_j p(i,j) \quad (6)$$

$$H(y) = - \sum_{i,j} p(i,j) \log \sum_i p(i,j) \quad (7)$$

- Aşağıdaki eşitsizlik, ancak olayların bağımsız olması durumunda gösterilebilir. Bu noktada bir ortak olayın belirsizliği bireysel belirsizliklerin toplamına eşit veya daha küçüktür denebilir.

$$H(x,y) \leq H(x) + H(y) \quad (8)$$

$$p(i,j) = p(i)p(j) \quad (9)$$

Özellik-4: Bir problemdeki p_i olasılıklarının eşit olmasına yönelik herhangi bir değişim H'ı artırır.

Örneğin $p_1 < p_2$ olsun. p_1 'i arttırmak p_2 'ye yaklaşmasını hatta eşit olmasını sağlamaya yönelik bir adımdır. İki durumun olma olasılığı $1/2$ 'ye eriştiğinde Entropi maksimum değerini alır.

Özellik-5: Üçüncü özelliğe olduğu gibi x ve y olayları olduğunu varsayalım.

- x için i ve y için j değerine göre iki olayın koşullu olasılığı $p_i(j)$ olarak tanımlanır ve formül:

$$p_i(j) = \frac{p(i,j)}{\sum_j p(i,j)} \quad (10)$$

- x'in elde edilme olasılığına göre ağırlıklandırılan x'in her bir değeri için y'nin entropi ortalaması olarak y'nin koşullu

entropisi $H_x(y)$ 'dir. Bu eşitlik x'i bildiğimiz zamandaki ortalama y'nin ne kadar belirsiz olduğunu ölçmek için kullanılmaktadır.

$$H_x(y) = - \sum_{i,j} p(i,j) \log p_i(j) \quad (11)$$

$$\begin{aligned} H_x(y) &= - \sum_{i,j} p(i,j) \log p(i,j) \\ &+ \sum_{i,j} p(i,j) \log \sum_j p(i,j) \end{aligned} \quad (12)$$

$$H_x(y) = H(x,y) - H(x) \quad (13)$$

$$H(x,y) = H(x) + H_x(y) \quad (14)$$

- x,y ortak olayının belirsizliği x'in belirsizliği ile x'in bilindiği durumdaki y'nin belirsizliğinin toplamına eşittir.

Özellik-6: Üç ve beşinci özelliklerde tanımladığımız formüller temel alındığında aşağıdaki yapı oluşmaktadır.

$$H(x) + H(y) \geq H(x,y) = H(x) + H_x(y) \quad (15)$$

Burdan yola çıkılarak aşağıdaki eşitsizlik tanımlanabilmektedir.

$$H(y) \geq H_x(y) \quad (16)$$

Böylelikle y'nin belirsizliğinin asla x bilgisiyle artmayacağı söylenebilmektedir. Hatta, x ve y bağımsız olaylar olmadıkça azalacaktır.

3. YAPAY ZEKA YÖNTEMLERİNDE ENTROPİ KULLANIMI (THE USE OF ENTROPY IN ARTIFICIAL INTELLIGENCE METHODS)

Entropi kullanıldığı her bilgi ölçütü yönteminde farklı şekilde ortaya çıkmaktadır. Yapay Zekâ Yöntemlerinde ise genel olarak kullanılan entropi Çapraz Entropi'dir. Bu bölümde Çapraz Entropi ve kullanıldığı Yapay Zekâ Yöntemleri detaylandırılmaktadır.

3.1. Çapraz Entropi (Cross Entropy)

Bilgi Kuramında çapraz entropi özellik-3 ile tanımlanmıştır. Buradaki temel amaç p yerine yanlış kodlama şeması q'yu kullanırsak q'dan bir olay kodlamak için gerekli ortalama bit sayısını elde etmektir. İki durum arasındaki olayın düzensizlik temelinde değerlendirilmesine çapraz entropi

denilmektedir ve aşağıdaki formül ile gösterilmektedir. [6, 7]

$$H(p, q) = E_p[-\log(q)] \\ = - \int_x p(x) \log(q(x)) dx \quad (17)$$

Sınıflandırma yöntemlerinde tahmin edilen ile olması gereken bilginin karşılaştırmasında yine entropi temelli çalışmalar yapılmaktadır. Bu noktada oluşturulan formül aşağıdaki şekilde tanımlanabilir. Aşağıdaki formülde y durumu 1 olduğunda \hat{y} , 0 olduğunda da $1 - \hat{y}$ durumu belirsizliğin ölçümünde önem ifade etmektedir.

$$H(p, q) = - \sum_i p_i \log q_i \\ = -y \log \hat{y} - (1 - y) \log(1 - \hat{y}) \quad (18)$$

3.2. Entropi'nin Yapay Zeka'da Kullanılan Diğer Türevleri (Other Derivatives of Entropy Used in Artificial Intelligence)

Bilginin ve kalitesinin önemli olduğu Yapay Zekâ yöntemlerinde bilginin kalitesinin ve doğruluğunun ölçülmesi gerekmektedir. Bu noktada birçok farklı yöntemler olmakla birlikte entropi de bu kullanılan yöntemlerden biridir. Bu bölümde entropinin yapay zekâ yöntemlerinde farklı şekillerde görülen türevleri açıklanmaktadır.

3.2.1. Maksimum Entropi (Maximum Entropy)

Elimizde bulunan bir veri kümesi üzerinde modelleme yapmak için en uygun dağılımı bulmak gerekebilir. Bu durumda önceki bilgileri de işin içine katarak aralarında en yüksek entropiyi sağlayan dağılımın en uygun dağılım olduğu çıkarımı yapılmaktadır. Özellikle Bayes Çıkarımı temelli yöntemlerde kullanılmaktadır. Kullanılan temel formül aşağıda sunulmakla birlikte, ana amaç entropiyi maksimum tutan dağılımı bulmaktır. [8]

$$S = - \sum_i p_i \log \frac{p_i}{m_i} \quad (19)$$

Bayes Çıkarımı ise rastgele iki olay arasındaki koşullu olasılıklar temelindeki ilişkiyi belirlemek için kullanılan bir teoremdir. Bu çıkarım üzerinden Naif Bayes gibi birçok yöntem türetilmiştir. Bilginin başka bir bilgi üzerindeki düzensizliği temelinde hareket eden çıkarım, entropi temelli hareket etmektedir. İlgili formül aşağıda sunulmaktadır. P(A), A'nın olma olasılığını, P(A|B) ise B için A'nın koşullu olasılığını ifade etmektedir.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (20)$$

3.2.2. Bilgi Kazanımı (Information Gain)

Karar ağaçları, kök düğümden başlanarak aşağıya doğru karar verici yönlendirmeler ile oluşturulan ağaçlardır. Ağacı inşa etmek için entropi kullanılmaktadır. Aslında kullanılan yöntemi entropi oluşturmaktadır. Entropiyi kullanan bu yöntem Bilgi Kazancı (Information Gain - IG) denmektedir. Bilgi Kazancı ise durumlar arası entropi farkından oluşmaktadır. ID3 ve C4.5 gibi karar ağacı yöntemleri tarafından kullanılmaktadır [9]. Aşağıda Bilgi Kazancına ait formül sunulmaktadır.[10]

$$IG(S, A) = H(S) - \sum_{t \in T} p(t)H(t) \\ = H(S) - H(S|A) \quad (21)$$

3.2.3. Kayıp Fonksiyonu (Loss Function)

Yapay Sinir Ağları gibi sınıflandırma yöntemlerinde bir durumun olup olmama durumu arasındaki farkın hesaplanmasında kayıp fonksiyonları kullanılmaktadır. Birçok farklı çeşidi olmakla birlikte en fazla kullanılan Ortalama Kare Hatası (Mean Squared Error)'dır. Bu yapıda çapraz entropide bahsettiğimiz yöntemden farklı olarak n örnek için toplanan çapraz entropi değerlerinin ortalamasının alınmasıdır. Bu şekilde bilgide oluşan düzensizliğin ortalaması üzerinden değerlendirme ve işlemler yapılabilmektedir. Formül aşağıda sunulmaktadır. i durum sayısını n ise toplam örnek sayını ifade etmektedir. [11]

$$L = - \frac{1}{n} \sum_{j=1}^n \sum_i p_i \log q_i \quad (22)$$

3.2.4. Lojistik Regresyon (Logistic Regression)

Temel sınıflandırma yöntemlerinden biri olarak kullanılan yöntemdir. Bir olayın gerçekleşme ve gerçekleşmeme durumlarının oranlarının logaritması olarak tanımlanmaktadır. Doğrusal Regresyondan farklı olarak kategorik bağımlı değişkenlere sahip olan yöntemin formülü aşağıda sunulmaktadır. [12]

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) \quad (23)$$

3.2.5. Kullback Leibler Uzaklığı (Kullback Leibler Distance)

Bilgi kazanımı, çapraz entropi ile aynı şey olmakla birlikte bir olayın olma durumu ile ikinci bir beklenen

olayın olasılıkları arasındaki düzensizliği hesaplamak için kullanılmaktadır. Formül olarak aşağıdaki şekilde sunulmaktadır. n tane örnek için yapılan hesaplama ortalama olarak alınmaktadır. [13]

$$D_{KL}(p||q) = H(p) - H(p, q) \quad (24)$$

$$D_{KL}(p||q) = \frac{1}{n} \sum_{i=1}^n (p_i \log \left(\frac{p_i}{q_i} \right)) \quad (25)$$

$$D_{KL}(p||q) = \frac{1}{n} \sum_{i=1}^n (p_i \log(p_i)) - \frac{1}{n} \sum_{i=1}^n (p_i \log(q_i)) \quad (26)$$

3.2.6. Renyi'nin Entropisi (Renyi's Entropy)

Alfred RENEY tarafından ortaya atılan entropi tanımıdır. Ekoloji ve istatistikte genelleştirme ve çeşitlilik indeksi olarak kullanılmakta ve genelleştirmedeki başarımı nedeni ile Yapay Zeka içinde kümeleme algoritmalarında kullanılmaktadır. $\alpha=2$ olduğunda oluşan formül ise Renyi'nin ikinci dereceden entropisi olarak çağrılmaktadır. [14]

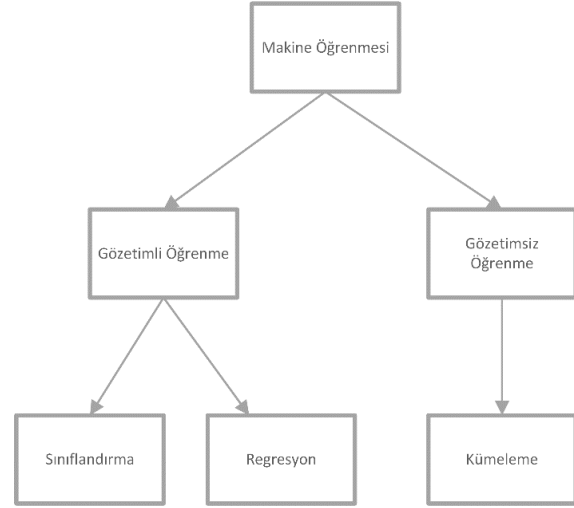
$$H_R(X) = \frac{1}{1-\alpha} \log \int f_X^\alpha dx, \alpha > 0, \alpha \neq 1 \quad (27)$$

3.3. Yapay Zekâ Yöntemleri (Artificial Intelligence Methods)

Zekâ, insanların temel davranışlarını önceki öğrenme dürtülerine göre şekillendirebilmesi için önem arz etmektedir. Bu noktada bilgiyi depolamada insanlardan daha iyi olan bilgisayarlarında bu bilgileri kullanarak yapacakları işleri şekillendirmede büyük başarılar ulaşması için çalışmalar yapılmaktadır. Günümüzde gelinen noktada belirli bir eşik aşılsa da halen insan beyninin olayları şekillendirmedeki becerisine teknoloji temelli tam olarak ulaşamamıştır.

Yapay Zekâ için bir problemin çözümünde önceden denenmiş veya probleme dayalı bilgilerin derlendiği bir veri kümesi olmalıdır. Bu veri kümesinin yapısı, içeriğine göre farklı önışlem adımları ve yöntemler kullanılmaktadır. Farklı birçok problemin çözümünde kullanılan Yapay Zekâ tekniklerinde özellikle Makine Öğrenmesi, gözetimli ve gözetimsiz öğrenme olmak üzere Şekil 2'de gösterildiği üzere iki temel kategoride incelenmektedir. Bu kategoriler ise üç alt kategoride ayrılmaktadır. Bunlar Sınıflandırma, Kümeleme ve Regresyondur. Sınıflandırmada öncelikle sistemin eğitilmesi gerekmektedir. Eğitim aşamasında verilen bilgiler belirli bir sonucu ifade

etmelidir. İfade ettiği sonuca göre verilen diğer parametrelere dayalı kullanılan yöntem kendisine ait bir model oluşturmaktadır. Sonrasında verilen her problem sorusuna önceki eğitim deneyimi kullanılarak yanıt verilmektedir. Kümelemede ise verilen bilgilerin uygun kategorilerde gruplanması beklenmektedir. Son olarak Regresyon yöntemlerinde ise verilen bilgilerin dağılımındaki ayırım önem arz etmektedir. Asıl problem bu ayırımın doğru şekilde en az yanlış sonuç verecek şekilde şekillendirilmesidir.



Şekil 2. Makine öğrenmesi kategorik yapısı
(Categorical structure of machine learning)

Var olan yöntemler temel alınarak literatürde yapılan entropi temelli Yapay Zekâ çalışmalarının entropiye dayalı kategorik yapısı Tablo 1'de sunulmaktadır. Yöntemin kullanıldığı bilgi türüne göre kategorisi ve kullandığı entropi türüne göre çeşitlendirilmiş formülü gösterilmektedir.

4. SONUÇ (CONCLUSION)

Günümüzde bilgi en değerli maden olarak görülmektedir. Fakat ham bilgi kendi başına yeterli değildir. Anlamlandırılması ve değerli hale dönüştürülmesi gerekmektedir. Bunun içinde ilk olarak ölçülebilirliktir. Bilginin değerini ölçmenin en temel yollarından biri olan entropi bu çalışmada tüm açılardan ele alınmaya çalışılmıştır. Özellikle çalışmanın ana konularından biri olan Yapay Zekâ içindeki kullanım çeşitliliğinin sunulması ve bunun üzerine yapılan tartışma, geliştirilmesi muhtemel Yapay Zekâ yöntemleri veya uygulama modellerine bir bakış açısı kazandıracaktır.

Entropi bilginin olasılıksak ölçümü olmakla birlikte Yapay Zekâ içindeki kullanım şekilleri problem uzaylarının karakterine göre şekillenmiştir. Fakat genel olarak belirli özelliklere sahip bir varlığın

Tablo 1. Entropinin kullanıldığı yapay zekâ yöntemleri
(Artificial intelligence methods using entropy)

Ref.	Yöntem	Kategori	Entropi Türü	Açıklama
[14]	Renyi'nin Entropisi ile Kümeleme	Kümeleme	Renyi'nin Entropisi	İlgili çalışmada Renyi entropisi kullanılarak yeni bir kümeleme algoritması geliştirilmiş ve hem yapay hemde gerçek veri kümeleri üzerinde test edilmiştir.
[15]	Maksimum Entropi ile Bulanık Kümeleme	Kümeleme	Maksimum Entropi	Maksimum entropi kullanılarak yeni bir bulanık kümeleme yöntemi önerilmiştir. Bulanık c-means yöntemi ile kıyaslandığında daha iyi sonuçlar gözlemlenmektedir.
[16]	Geliştirilmiş ID3 Karar Ağacı Algoritması	Sınıflandırma	Bilgi Kazanımı	ID3 algoritmasının yetersizliği belirtildikten sonra ortaya konan yeni yöntem ID3 ile Ortaklık Fonksiyonunun birleştirilmesinden elde edilmektedir. Bu noktada değişimler Bilgi Kazanımı üzerinde de yapılmıştır. Normal yöntem ile arasındaki fark yapılan testlerde ağacın oluşturulması açısından iyi sonuç üretecek formda sunulmaktadır.
[17]	Genetik Algoritma – Lojistik Regresyon Modeli	Sınıflandırma ve Regresyon	Lojistik Regresyon	Lojistik Regresyon, Geri Yayılımlı Sinir Ağı ve Lojistik Regresyon tabanlı Genetik Algoritma yöntemlerinin kıyaslandığı bir çalışmadır. Yapılan testlerde, istatistik modeli olarak bilinen Lojistik Regresyonun Yapay Zeka yöntemlerinden Genetik Algoritma ile kullanımının başarıyı arttırdığı gözlemlenmiştir.
[18]	Yapay Sinir Ağları	Sınıflandırma	Kayıp Fonksiyonu	Üç farklı yapay sinir ağı üzerinde rüzgar hızı tahmini üzerine yapılan çalışmada 3 farklı kayıp fonksiyonu üzerinden test yapılmıştır. Bunlar Ortalama Mutlak Hata (Mean Absolute Error), Kök Ortalama Kare Hatası (Root Mean Squared Error) ve Ortalama Mutlak Yüzde Hatası (Mean Absolute Percentage Error)'dur. Yapılan farklı model testlerinde her ölçüm yönteminin değişiminin farklı olduğu gözlemlenmektedir.
[19]	KL Uzaklığı ile Derin Sinir Ağları Adaptasyonu	Sınıflandırma ve Kümeleme	Kullback Leibler Uzaklığı	Sözlük konuşma tanıma üzerine oluşturulan Derin Sinir Ağı'nın KL Uzaklığı ile adaptasyonu üzerine bir yöntem önerilmektedir. Hem gözetimli hem de gözetimsiz öğrenme şeklinde uygulanan adaptasyonun farklı boyutlardaki veriler üzerindeki testlerinde en iyi kelime hata oranı 200 veri boyutuna sahip testlerde gözlemlenmiştir.

kategorizasyonu problem olarak ele alındığından çapraz entropinin büyük çoğunlukla etkisinin görüldüğü gözlemlenmiştir. Aslında sunulan problemlerin çözümleri ele alındığında temel amaç, yöntemin geliştirilmesi aşamasında oluşturulan matematiksel modelde bilginin çeşitliliğine göre oluşan bilgi kaybının minimize edilerek en mantıklı ve doğru sonuca sistemi ulaştırmaktır. Bu yapıda ise en uygun yöntemlerden biri entropi temelli bilginin düzensizliğini ölçmektir.

Sunulan çalışmada entropiye dayalı Yapay Zekâ yöntemleri farklı bir bakış açısı ile sunulmaya çalışılmıştır. Entropinin kullanıldığı tek alan olmamakla birlikte bilginin ve onu anlamlandırma ihtiyacının olduğu her alanda kullanımlarını görmek mümkündür. Gelecek çalışmalarda Yapay Zekâ içindeki kullanıldığı yöntemlerdeki entropi türünün dışında diğer türlerin kullanılabilirliği ve kullanıldığındaki yöntem başarımları üzerine araştırma yapılması planlanmaktadır.

KAYNAKLAR (REFERENCES)

- [1] R. Clausius, "Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie," *Annalen der Physik*, vol. 201, pp. 353-400, 1865.
- [2] M. Badino, "Mechanistic slumber vs. statistical insomnia: the early history of Boltzmann's H-theorem (1868–1877)," *The European Physical Journal H*, vol. 36, pp. 353-378, 2011.
- [3] J. W. Gibbs, *On the equilibrium of heterogeneous substances*, 1879.
- [4] H. Nyquist, "Certain factors affecting telegraph speed," *Transactions of the American Institute of Electrical Engineers*, vol. 43, pp. 412-422, 1924.
- [5] R. V. Hartley, "Transmission of information 1," *Bell System technical journal*, vol. 7, pp. 535-563, 1928.

- [6] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, pp. 379-423, 1948.
- [7] C. E. Shannon, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on information theory*, vol. 26, pp. 26-37, 1948.
- [8] S. F. Gull, "Bayesian inductive inference and maximum entropy," in *Maximum-entropy and Bayesian methods in science and engineering*, ed: Springer, 1988, pp. 53-74.
- [9] B. Hssina, A. Merbouha, H. Ezzikouri, and M. Erritali, "A comparative study of decision tree ID3 and C4. 5," *International Journal of Advanced Computer Science and Applications*, vol. 4, 2014.
- [10] A. Janecek, W. Gansterer, M. Demel, and G. Ecker, "On the relationship between feature selection and classification accuracy," in *New Challenges for Feature Selection in Data Mining and Knowledge Discovery*, 2008, pp. 90-105.
- [11] N. Murata, S. Yoshizawa, and S.-i. Amari, "Network information criterion-determining the number of hidden units for an artificial neural network model," *IEEE Transactions on Neural Networks*, vol. 5, pp. 865-872, 1994.
- [12] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *The journal of educational research*, vol. 96, pp. 3-14, 2002.
- [13] C. Liu and H.-Y. Shum, "Kullback-leibler boosting," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 2003, pp. I-I.
- [14] R. Jenssen, K. Hild, D. Erdogmus, J. C. Principe, and T. Eltoft, "Clustering using Renyi's entropy," in *Neural Networks, 2003. Proceedings of the International Joint Conference on*, 2003, pp. 523-528.
- [15] R.-P. Li and M. Mukaidono, "A maximum-entropy approach to fuzzy clustering," in *Fuzzy Systems, 1995. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE Int*, 1995, pp. 2227-2232.
- [16] C. Jin, L. De-Lin, and M. Fen-Xiang, "An improved ID3 decision tree algorithm," in *Computer Science & Education, 2009. ICCSE'09. 4th International Conference on*, 2009, pp. 127-130.
- [17] C.-L. Chang and M.-Y. Hsu, "The study that applies artificial intelligence and logistic regression for assistance in differential diagnostic of pancreatic cancer," *Expert Systems with Applications*, vol. 36, pp. 10663-10672, 2009.
- [18] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Applied Energy*, vol. 87, pp. 2313-2320, 2010.
- [19] D. Yu, K. Yao, H. Su, G. Li, and F. Seide, "KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 7893-7897.