



Türkçe Metin Madenciliği için Dikkat Mekanizması Tabanlı Derin Öğrenme Mimarilerinin Değerlendirilmesi

Aytuğ Onan*¹

¹* İzmir Katip Çelebi Üniversitesi, İzmir, Türkiye, (ORCID: 0000-0002-9434-5880), aytugonan@gmail.com

(2nd International Conference on Applied Engineering and Natural Sciences ICAENS 2022, March 10-13, 2022)

(DOI: 10.31590/ejosat.1082379)

ATIF/REFERENCE: Onan, A. (2022). Türkçe metin madenciliği için dikkat mekanizması tabanlı derin öğrenme mimarilerinin değerlendirilmesi. *European Journal of Science and Technology*, (34), 403-407.

Öz

Son yıllarda, hesaplama maliyetlerindeki düşüşler ve veri hacmindeki artışlar ile, büyük veri üzerinde etkin öğrenme modelleri oluşturmaya yönelik derin sinir ağı modelleri alanında önemli başarımlar elde edilmeye başlanmıştır. Derin sinir ağı mimarileri, duygu analizi, metin sınıflandırma, makine çevirisi gibi doğal dil işleme problemlerinin yanı sıra, konuşma tanıma, bilgisayarla görme gibi birçok alanda başarıyla uygulanmaktadır. Metin sınıflandırma problemlerinde, tekrarlayan sinir ağları, uzun kısa süreli bellek mimarisi ve geçitli tekrarlayan birim gibi mimariler etkin bir biçimde kullanılmaktadır. Tekrarlayan sinir ağları (RNN), uzun süreli bağımlılıkları öğrenirken verimli değildir. Bu mimaride, belirteçler sırayla işlenir ve her belirteçten sonra görülen verileri temsil eden bir durum vektörü elde edilir. RNN tabanlı mimariler, kaybolan gradyan sorunu nedeniyle, uzun tümcelerin sonundaki durumun genellikle erken gözlemlenen belirteçlere ilişkin bilgi içermemesi sorununa neden olur. Dikkat mekanizması, bir modelin tümcedeki daha önceki bir noktanın durumuna doğrudan bakmasını ve bundan yararlanabilmesini olanaklı kılar. Böylelikle, tüm durumlara erişilebilir ve mevcut belirteçle öğrenilmiş bazı ilgi ölçütlerine göre, çok uzaktaki ilgili belirteçler hakkında daha net bilgi sağlanarak, daha etkin bir metin modelleme gerçekleştirilebilir. Dikkat mekanizması, makine çevirisi, metin özetleme başta olmak üzere birçok doğal dil işleme görevinde başarıyla uygulanmaktadır. Bu çalışmada, Türkçe duygu analizi için dikkat mekanizması tabanlı derin öğrenme mimarilerinin başarımları karşılaştırmalı olarak incelenmektedir. Deneysel analizlerde, tekrarlayan sinir ağı, uzun kısa süreli bellek mimarisi (LSTM), geçitli tekrarlayan birim mimarisi (GRU), küresel dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM, hiyerarşik dikkat mekanizması tabanlı LSTM, küresel dikkat mekanizması tabanlı GRU, kendine dikkat mekanizması tabanlı GRU ve hiyerarşik dikkat mekanizması tabanlı GRU değerlendirilmiştir. Dikkat mekanizması tabanlı mimarilerin, duygu analizi için doğru sınıflandırma oranını önemli ölçüde artırdığı gözlenmektedir.

Anahtar Kelimeler: Uzun kısa süreli bellek, Dikkat Mekanizması, Metin Madenciliği, Derin Öğrenme, Makine Öğrenmesi.

Evaluation of Attention Mechanism-Based Deep Learning Architectures for Turkish Text Mining

Abstract

In recent years, with the decrease in computational costs and the increase in data volume, significant achievements have been done in the field of deep neural network models for constructing effective learning models on big data. Deep neural network architectures have been successfully applied in many fields such as speech recognition, computer vision, as well as natural language processing problems such as sentiment analysis, text classification, and machine translation. In text classification problems, architectures such as recurrent neural networks, long-short-term memory architecture and gated recurrent unit are effectively used. Recurrent neural networks (RNN) are not efficient when learning long-term dependencies. In this architecture, tokens are processed sequentially, resulting in a state vector representing the data seen after each token. RNN-based architectures suffer from the problem that the state at the end of long sentences often does not contain information about early observed markers, due to the vanishing gradient problem. The attention mechanism makes it possible for a model to look directly at the state of an earlier point in the sentence and take advantage of it. Thus, all cases can be accessed and a more effective text modeling can be performed by providing clearer information about very distant related markers, according to some criteria of relevance learned with the current marker. The attention mechanism has been successfully applied in many natural language processing tasks, especially machine translation and text summarization. In this study, the performances of attention mechanism-based deep learning architectures for Turkish sentiment analysis are comparatively examined. In experimental analysis, recurrent neural network, long short-term memory architecture (LSTM), gated recurrent unit architecture (GRU), global attention mechanism-based LSTM, self-attention mechanism-based LSTM, hierarchical attention mechanism-based LSTM, global attention mechanism-based GRU, self-attention mechanism-based GRU and hierarchical attention mechanism-based GRU were evaluated. Attention mechanism-based architectures are observed to significantly increase the rate of correct classification for sentiment analysis.

Keywords: Long short-term memory, Attention mechanism, Text mining, Deep learning, Machine learning.

* Sorumlu Yazar: g.kocamanoglu@hotmail.com

1. Giriş

Bilgi ve iletişim teknolojilerindeki ilerlemeler ile, metin belgesi türündeki veri hacminde önemli artışlar meydana gelirken, buna paralel olarak, hesaplama maliyetinde de önemli düşüşler meydana gelmiştir. Büyük veri üzerinde, etkin öğrenme modelleri oluşturulmasına yönelik derin sinir ağı modelleri alanında önemli başarımlar elde edilmektedir. Derin sinir ağı mimarileri, duygu analizi, metin sınıflandırma, makine çevirisi gibi doğal dil işleme problemlerinin yanı sıra, konuşma tanıma, bilgisayarla görme gibi birçok alanda başarıyla uygulanmaktadır [1]. Metin sınıflandırma, metin belgelerinin önceden belirlenmiş sınıf etiketlerinden birine atanmasına yönelik bir doğal dil işleme (DDİ) çalışma alanıdır. Görüş madenciliği olarak da adlandırılan duygu analizi, kaynak materyallerdeki bir varlığa yönelik görüş, duygu, tutum gibi öznel bilgileri çıkarmaya yönelik bir disiplinlerarası araştırma alanıdır [2, 3]. Duygu analizi kullanılarak insanların belirli bir olaya yönelik görüşlerinin belirlenmesi, karar verme sürecinin faktörleri ve etkileycilerinin belirlenmesi, yönetim bilimleri, siyaset bilimi, ekonomi ve sosyal bilimler gibi birçok disiplin için oldukça önemli olabilmektedir [4].

Metin sınıflandırma problemlerinde, tekrarlayan sinir ağları, uzun kısa süreli bellek mimarisi ve geçitli tekrarlayan birim gibi mimariler etkin bir biçimde kullanılmaktadır. Tekrarlayan sinir ağları (RNN), uzun süreli bağımlılıkları öğrenirken verimli değildir. Bu mimaride, belirteçler sırayla işlenir ve her belirteçten sonra görülen verileri temsil eden bir durum vektörü elde edilir. RNN tabanlı mimariler, kaybolan gradyan sorunu nedeniyle, uzun tümcelerin sonundaki durumun genellikle erken gözlemlenen belirteçlere ilişkin bilgi içermemesi sorununa neden olur. Dikkat mekanizması, bir modelin tümcedeki daha önceki bir noktanın durumuna doğrudan bakmasını ve bundan yararlanabilmesini olanaklı kılar. Böylelikle, tüm durumlara erişilebilir ve mevcut belirteçle öğrenilmiş bazı ilgi ölçütlerine göre, çok uzaktaki ilgili belirteçler hakkında daha net bilgi sağlanarak, daha etkin bir metin modelleme gerçekleştirilebilir. Dikkat mekanizması, makine çevirisi, metin özetleme başta olmak üzere birçok doğal dil işleme görevinde başarıyla uygulanmaktadır. Bu çalışmada, Türkçe duygu analizi için dikkat mekanizması tabanlı derin öğrenme mimarilerinin başarımları karşılaştırmalı olarak incelenmektedir. Deneysel analizlerde, tekrarlayan sinir ağı, uzun kısa süreli bellek mimarisi (LSTM), geçitli tekrarlayan birim mimarisi (GRU), küresel dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM, hiyerarşik dikkat mekanizması tabanlı LSTM, küresel dikkat mekanizması tabanlı GRU, kendine dikkat mekanizması tabanlı GRU ve hiyerarşik dikkat mekanizması tabanlı GRU değerlendirilmiştir.

Çalışmanın geri kalanı şu bölümlerden oluşmaktadır: İkinci bölümde, ilgili çalışmalara değinilmektedir, üçüncü bölümde çalışmanın metodolojisini oluşturan derin öğrenme mimarileri ve dikkat mekanizması yaklaşımları tanıtılmaktadır. Dördüncü bölümde, deneysel sonuçlar ve tartışma, son bölümde ise çalışmanın temel sonuçları sunulmaktadır.

2. İlgili Çalışmalar

Dikkat mekanizmasına dayalı mimariler, duygu analizi, metin sınıflandırma, makine çevirisi gibi birçok doğal dil işleme görevlerinde başarıyla uygulanmaktadır. Örneğin, Munkhdalai

vd. [5] tarafından tekrarlayan sinir ağı ve yinelemeli sinir ağı tabanlı iki derin sinir ağı mimarisi geliştirilmiştir. Geliştirilen mimari, bir öncül ve bir hipotez temsilini öğrenen ve daha sonra bunları bir dikkat mekanizması aracılığıyla birleştiren sözdizimsel ayrıştırmadan bağımsız ağaç yapılı bir modeldir. Bir diğer çalışmada, Yang vd. [6] tarafından metin belgelerinin hiyerarşik yapısını modellemek amacıyla, hiyerarşik dikkat ağı adı verilen bir mimari geliştirilmiştir. Geliştirilen mimeri, sözcük ve tümcelerin bilgi vericilik açısından farklılık göstermesi nedeniyle sözcük ve tümce düzeyinde uygulanan iki adet dikkat mekanizması uygulamaktadır. Mimarinin temelinde kodlayıcı olarak, geçitli tekrarlayan kapı mimarisi kullanılmıştır. Önerilen yöntem, test edildiği altı farklı doğal dil işleme görevinde, güncel mimari ve yöntemlere kıyasla daha yüksek başarımlar elde etmiştir. Bir başka çalışmada, Yin vd. [7] önceki modeli genişleterek, en-boy tabanlı duygu analizi görevi için hiyerarşik bir yinelemeli dikkat modeli önermiştir. En-boy temelli model, bir makine anlama problemi olarak modellenmiştir. Model, hiyerarşik dikkat ağı tabanlı mimariye kıyasla daha yüksek başarımlar elde etmiştir. Lin vd. [8] tarafından gerçekleştirilen çalışmada, kendine dikkat tabanlı bir mimari geliştirilmiştir. Bu model, tümce gömme için yeni bir model önermektedir. İlgili temsilde, her gömme, matrisin her satırının tümcenin farklı bir kısmına katıldığı 2 boyutlu bir matrise karşılık gelir. Öncelikle tümce bir RNN'den geçirilir. Ardından, her RNN durumu için birden fazla dikkat değeri öğrenilir ve ardından her bir dikkat vektörü bir ağırlık terimi eklenerek tümcenin farklı bölümlerine odaklanır. İlgili model hem duygu analizi hem de metin sınıflandırma problemlerinde yüksek başarımlar elde etmiştir. Chen vd. [9] tarafından gerçekleştirilen bir başka çalışmada, genel kullanıcı tercihini ve ürün özelliklerini belirlemeye yönelik olarak LSTM ve dikkat mekanizmasına dayalı bir mimari gerçekleştirimi yapılmıştır. Liu vd. [10] tarafından iki yönlü LSTM'e dayalı ve gizli katmanlardan gelen bilgi çıkışına odaklanmak amacıyla dikkat mekanizması kullanan bir model geliştirilmiştir. Bu mimaride, daha yüksek seviyeli ifade temsillerini çıkarabilmek amacıyla, LSTM tabanlı mimari, evrişimli sinir ağı ile birleştirilmiştir. Fu vd. [11] tarafından geliştirilen başka bir çalışmada ise, LSTM tabanlı derin sinir ağı mimarisine, sözlük tabanlı bir dikkat mekanizması dahil edilmiştir. Böylelikle, belirli bir Genel olarak, sözlük, belirli bir hedefe sahip olmak yerine tüm metnin genel bilgisini kullanan bir dikkat mekanizması ile kelime yerleştirmelerinin kelimeleri temsil etme yeteneğini geliştirir.

3. Metodoloji

Bu bölümde, deneysel analizlerde kullanılan kelime gömme yöntemi, derin öğrenme algoritmaları ve dikkat mekanizması yaklaşımları sunulmaktadır.

3.1. word2vec Modeli

Kelimelerin sürekli vektörler olarak gösterilmesine dayanan word2vec modeli, en temel kelime gömme modellerinden biridir [12]. Word2vec modelinin temel fikri, model eğitimi sırasında sürekli vektörlerle çoklu benzerlik ilişkilerinin yakalanabilmesidir. Word2vec, girdi katmanı, çıktı katmanı ve gizli katmanı içeren yapay sinir ağı tabanlı bir dil modelleme yöntemidir. Word2vec, kelime vektörlerini eğitmek için iki temel algoritma içerir: sürekli kelime torbası (CBOW) ve atlama gramı (SG). CBOW modelinde bağlam vektörü, bağlam penceresindeki kelime vektörlerinin toplamından elde edilir. SG algoritmasında hedef kelimeyi çevreleyen kelimeler, hedef

kelime üzerinden belirlenir. CBOW algoritması, az miktarda veri ile etkili bir şekilde çalışabilir. SG mimarisi hesaplama açısından daha maliyetlidir ve büyük veri kümelerinde daha etkili sonuçlar verir [13].

3.2. Derin Öğrenme Algoritmaları

Tekrarlayan sinir ağı (RNN), sıralı verileri işlemek için kullanılan bir derin öğrenme mimarisi türüdür. RNN'de, nöronlar arasındaki bağlantılar yönlendirilmiş bir grafik oluşturur. Bu mimaride, girdi dizisini işlemek için dahili durum kullanılmıştır. Bu nedenle mimari, konuşma tanıma gibi sıralı görevler için başarıyla kullanılabilir. RNN'de, her bir çıktı, dizinin örnekleri üzerinde aynı görevin tekrar tekrar işlenmesiyle belirlenir. Daha önceki tüm hesaplamalara dayanarak, çıktı belirlenir [14, 15].

Uzun kısa süreli bellek ağları (LSTM), tekrarlayan sinir ağlarına dayalı başka bir derin öğrenme mimarisidir. Geleneksel RNN mimarisinde patlayan veya kaybolan gradyan problemi gözlemlenmektedir. RNN'de, oldukça uzun giriş dizileri uygun şekilde işlenemez. Yanıt olarak, LSTM sorunların üstesinden gelmek için unutma kapılarını kullanır. LSTM mimarisinde, sınırlı sayıda zaman adımına kadar hatanın geri yayılmasına izin verilmiştir. Tipik bir LSTM ünitesi için, bir hücrede, giriş kapısı, çıkış kapısı ve unutma kapısı olmak üzere üç çeşit kapı vardır. Kapılar üzerindeki açma ve kapama işlemleri, hangi bilgilerin korunması gerektiğini ve bilgilere ne zaman erişilmesi gerektiğini kontrol etmek amacıyla kullanılmaktadır [16].

Geçitli tekrarlayan birim mimarisi (GRU), tekrarlayan sinir ağlarına dayalı başka bir derin öğrenme mimarisidir [17]. Tipik bir GRU mimarisinde, sıfırlama kapısı ve güncelleme kapısı olmak üzere iki kapı bulunmaktadır.

3.3. Dikkat Mekanizması Yöntemleri

Küresel dikkat mekanizması yönteminde, dikkat ağırlık vektörü α ve ağırlıklı gizli temsil r , $H \in R^{d \times N}$, derin sinir ağı gizli vektörlerini $[h_1, \dots, h_N]$, d , gizli katman boyutunu, W , eğitim sürecinde öğrenilen ağırlık değerlerini ve v_α kelime gömmeyi temsil etmek üzere, sırasıyla Eşitlik 2 ve Eşitlik 3'e göre belirlenir [18, 19]:

$$M = \tanh \left(\begin{bmatrix} W_h H \\ W_v v_\alpha \otimes e_N \end{bmatrix} \right)$$

$$\alpha = \text{softmax} (W^T M)$$

$$r = H \alpha^T$$

Küresel dikkat mekanizması yönteminde, her bir t zaman adımında, mevcut hedef durumu h_t ve tüm kaynak durumları h_s^* değişken uzunlukta bir hizalama ağırlık vektörü olan α_t vektörünü oluşturur. Küresel bağlam vektörü (c_t) daha sonra tüm kaynak durum üzerinden α_t vektörüne göre ağırlıklı ortalama olarak hesaplanır. Böylelikle, kodlayıcı RNN yapısının tüm gizli durumlarına dayalı ağırlıklandırılmış bir bağlam vektörü elde edilmesi amaçlanır. Küresel dikkat mekanizması, tüm girdi durum uzayını dikkate alır.

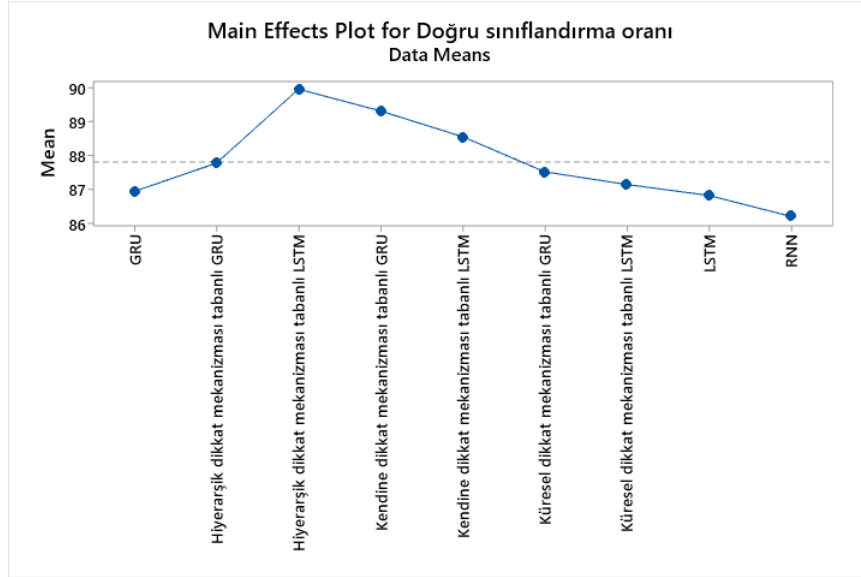
Kendine dikkat mekanizması yönteminde, dikkat mekanizmasının tümcenin bölümlerine odaklanmasına izin verilerek RNN tabanlı mimarilerin daha önce değinilen

problemlerinin ortadan kaldırılması amaçlanır. Burada, bölümün uygunluğu modelleme yapılan göreve yapılan katkıya göre belirlenir. Giriş dizisinden türetilen aynı gizli durum uzayının farklı konumları, birden çok bileşenin birlikte bir dizinin genel semantiğini oluşturduğu argümanına dayalı olarak farklı konumlarla ilişkilidir. Bu yaklaşım, birden fazla atlamının dikkati yoluyla farklı şekilde konumlanmış bilgileri bir araya getirir [19, 20]. Bu mimaride, gömülü belirteçler, LSTM katmanlarına beslenir. Gizli durumlar, sınıflandırma için bir girdi olarak kullanılan rafine bir tümce temsili elde etmek amacıyla bir dikkat vektörü ile ağırlıklandırılır. Birden fazla dikkat vektörünün benzer veya fazla olmasını önlemek için kendine dikkat matrisine dayalı bir ceza terimi kullanılır [20].

Hiyerarşik dikkat mekanizması yönteminde, metin belgelerinde yer alan hiyerarşik yapıyı modellemek amacıyla genellikle aşağıdan yukarıya bir yöntem izlenir [21]. Bu mimaride, dikkat mekanizması öncelikle sözcük ve ardından tümce seviyelerinde uygulanır. Böylelikle, metin belgesi temsili oluşturulurken daha fazla önemli ve daha az önemli içeriklerin farklı şekilde dikkate alınması olanaklı hale gelir. Hiyerarşik dikkat mekanizması, ilki kelime kodlayıcı ve ikincisi tümce kodlayıcı olmak üzere iki seviyeli geleneksel bir yapay sinir ağı mimarisine sahiptir. Bu mimaride, tümcenin anlamı için önemli olan sözcükleri çıkarabilen ve bu bilgilendirici sözcüklerin temsili bir tümce vektörü oluşturmak üzere toplayabilen bir sözcük tabanlı dikkat mekanizması uygulanır. Ardından, açıklamalar tekrar eğitilebilir bir bağlam vektörü ile çarpılır ve her bir sözcük için önem ağırlığına göre normalleştirilir. Önceden hesaplanmış bağlam açıklamalarıyla birleştirilen önem ağırlıklarının toplamı, tümce vektörünü oluşturur.

4. Deneysel Sonuçlar ve Tartışma

Deneysel analizde, Türkçe Twitter mesajları içeren, terim ağırlıklandırma yöntemlerinin duygu analizi alanındaki etkinliklerini değerlendirmek amacıyla oluşturulmuş bir veri seti kullanılmıştır [22]. Veri seti, 10,500'ü olumlu ve 10,500'ü olumsuz olmak üzere 21000 Twitter mesajı içermektedir. Twitter mesajları üzerinde, kök bulma, bilgi vericiliği olmayan kelimelerin çıkarılması gibi ön işleme adımları uygulanmıştır. Ham Twitter mesajlarına "olumlu" ya da "olumsuz" olmak üzere sınıf etiketlerinin atanabilmesi için (1) iki uzmandan yararlanılmıştır. Sınıf etiketlerinin uygun atanıp atanmadığını belirlemek amacıyla Cohen'in kappa (κ) metriği (2) hesaplanmıştır. Derlem için, açıklayıcılar arasında mükemmel bir uyum olduğunu gösteren 0.82'lik bir puan elde edilmiştir. Veri seti üzerinde [22]'de belirtilen ön işleme adımları uygulanarak veri seti, derin öğrenme modelleri ile işlemeye hazır hale getirilmiştir. Derin öğrenme mimarilerinin performansını değerlendirmek amacıyla, doğru sınıflandırma oranı, F-ölçütü, geri çağırma ve hassasiyet ölçütleri kullanılmıştır. Mimarilerin değerlendirilmesinde, 5-kat çapraz geçişleme kullanılmıştır. Gerçekleştirimler, Tensorflow derin öğrenme çerçevesi kullanılarak Python dilinde yapılmıştır. Eniyileme algoritması olarak Adam optimizör kullanılmıştır, toplu iş boyutu 256, nöron sayısı 256 ve öğrenme oranı 0.001 alınarak derin öğrenme mimarileri eğitilmiştir. Kelime gömme vektörü olarak, önceden eğitilmiş 300 boyutlu bir word2vec modeli kullanılmıştır.



Şekil 1. Karşılaştırılan yöntemlere ilişkin temel etki diyagramı

Tablo 1. Derin öğrenme mimarileri ile elde edilen doğru sınıflandırma, geri çağırma, hassasiyet ve F-ölçütü değerleri

Derin öğrenme mimarisi	Doğru sınıflandırma oranı	Geri Çağırma	Hassasiyet	F-ölçütü
RNN	86,20	0,87	0,88	0,88
LSTM	86,82	0,88	0,89	0,88
GRU	86,94	0,88	0,89	0,88
Küresel dikkat mekanizması tabanlı LSTM	87,14	0,88	0,89	0,88
Kendine dikkat mekanizması tabanlı LSTM	88,55	0,89	0,90	0,90
Hiyerarşik dikkat mekanizması tabanlı LSTM	89,96	0,91	0,92	0,91
Küresel dikkat mekanizması tabanlı GRU	87,52	0,88	0,89	0,89
Kendine dikkat mekanizması tabanlı GRU	89,31	0,90	0,91	0,91
Hiyerarşik dikkat mekanizması tabanlı GRU	87,77	0,89	0,90	0,89

Tablo 1’de Türkçe duygu analizi veri seti üzerinde tekrarlayan sinir ağı (RNN), uzun kısa süreli bellek mimarisi (LSTM), geçitli tekrarlayan birim mimarisi (GRU), küresel dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM, hiyerarşik dikkat mekanizması tabanlı LSTM, küresel dikkat mekanizması tabanlı GRU, kendine dikkat mekanizması tabanlı GRU ve hiyerarşik dikkat mekanizması tabanlı GRU ile elde edilen sonuçlar sunulmaktadır. Tablo 1’de sunulan sonuçlar incelendiğinde, karşılaştırılan derin öğrenme mimarileri arasında en düşük sonucun geleneksel tekrarlayan sinir ağı (RNN) ile alındığı ve bu mimarinin %82,68’lik bir başarımla elde ettiği görülmektedir. RNN’in ardından ikinci en düşük doğru sınıflandırma oranı değeri, geleneksel uzun kısa süreli bellek mimarisi (LSTM) ile alınmaktadır. Geçitli tekrarlayan birim mimarisi (GRU), LSTM’e kıyasla daha basit yapısı olmasına karşın daha yüksek başarımla elde edilmesine neden olmaktadır. Hiyerarşik dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM ve küresel dikkat mekanizması tabanlı LSTM mimarilerine kıyasla daha yüksek başarımla göstermektedir. GRU ile dikkat mekanizması yöntemlerinin birarada kullanılması incelendiğinde ise kendine dikkat mekanizması tabanlı GRU mimarisinin hiyerarşik dikkat mekanizması tabanlı GRU ve küresel dikkat mekanizması tabanlı GRU mimarilerinin performansını doğru sınıflandırma, geri çağırma, hassasiyet ve F-ölçütü bakımından geçtiği

görülmektedir. Şekil 1’de karşılaştırılan yöntemlere ilişkin temel etki diyagramı sunulmaktadır. Tablo 1’de ve Şekil 1’de sunulan sonuçlardan da görülebileceği gibi karşılaştırmada kullanılan tüm derin öğrenme mimarileri arasında en yüksek başarımla %89,96’lık doğru sınıflandırma oranı ile hiyerarşik dikkat mekanizması tabanlı LSTM ile elde edilmiştir.

5. Sonuç

Son yıllarda, hesaplama maliyetlerinin düşmesi ve veri hacimlerinin artmasıyla birlikte, büyük veri üzerinde verimli öğrenme modelleri geliştirmek için derin sinir ağı modelleri alanında önemli ilerlemeler kaydedilmiştir. Derin sinir ağı mimarileri, konuşma tanıma, bilgisayarla görme ve duygu analizi, metin kategorizasyonu ve makine çevirisi dahil olmak üzere doğal dil işleme görevleri dahil olmak üzere çeşitli alanlarda başarıyla kullanılmıştır. Tekrarlayan sinir ağları, uzun-kısa süreli bellek mimarileri ve geçitli tekrarlayan birimlerin tümü, metin kategorizasyonu konularında faydalıdır. Uzun vadeli bağımlılıkları öğrenmek söz konusu olduğunda, tekrarlayan sinir ağları (RNN) etkisizdir. Belirteçler bu tasarımda sırayla işlenir ve her belirteçten sonra görüntülenen verileri temsil eden bir durum vektörü elde edilir. Dikkat mekanizması, bir modelin tümcedeki daha önceki bir noktanın durumuna doğrudan bakmasını ve bundan yararlanabilmesini

olanaklı kılar. Böylelikle, tüm durumlara erişilebilir ve mevcut belirteçle öğrenilmiş bazı ilgi ölçütlerine göre, çok uzaktaki ilgili belirteçler hakkında daha net bilgi sağlanarak, daha etkin bir metin modelleme gerçekleştirilebilir. Dikkat mekanizması, makine çevirisi, metin özetleme başta olmak üzere birçok doğal dil işleme görevinde başarıyla uygulanmaktadır. Bu çalışmada, Türkçe duygu analizi için dikkat mekanizması tabanlı derin öğrenme mimarilerinin başarımları karşılaştırmalı olarak incelenmektedir. Deneysel analizlerde, tekrarlayan sinir ağı, uzun kısa süreli bellek mimarisi (LSTM), geçitli tekrarlayan birim mimarisi (GRU), küresel dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM, hiyerarşik dikkat mekanizması tabanlı LSTM, küresel dikkat mekanizması tabanlı GRU, kendine dikkat mekanizması tabanlı GRU ve hiyerarşik dikkat mekanizması tabanlı GRU değerlendirilmiştir. Hiyerarşik dikkat mekanizması tabanlı LSTM, kendine dikkat mekanizması tabanlı LSTM ve küresel dikkat mekanizması tabanlı LSTM mimarilerine kıyasla daha yüksek başarımlar göstermektedir. GRU ile dikkat mekanizması yöntemlerinin birarada kullanılması incelendiğinde ise kendine dikkat mekanizması tabanlı GRU mimarisinin hiyerarşik dikkat mekanizması tabanlı GRU ve küresel dikkat mekanizması tabanlı GRU mimarilerinin performansını doğru sınıflandırma, geri çağırma, hassasiyet ve F-ölçütü bakımından geçtiği görülmektedir. Karşılaştırmada kullanılan tüm derin öğrenme mimarileri arasında en yüksek başarımlar %89,96'lık doğru sınıflandırma oranı ile hiyerarşik dikkat mekanizması tabanlı LSTM ile elde edilmiştir.

Teşekkür

Bu araştırma, İzmir Kâtip Çelebi Üniversitesi Bilimsel Araştırma Koordinasyon birimi (BAP) tarafından desteklenmiştir (Proje no: 2022-GAP-MÜMF-0030).

Kaynakça

- Deng, L., & Liu, Y. (Eds.). (2018). *Deep learning in natural language processing*. Springer.
- Onan, A., & Korukoğlu, S. (2017). A feature selection model based on genetic rank aggregation for text sentiment classification. *Journal of Information Science*, 43(1), 25-38.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- Tan, S., & Zhang, J. (2008). An empirical study of sentiment analysis for chinese documents. *Expert Systems with applications*, 34(4), 2622-2629.
- Munkhdalai, T., & Yu, H. (2017, April). Neural tree indexers for text understanding. In *Proceedings of the conference. Association for Computational Linguistics. Meeting* (Vol. 1, p. 11). NIH Public Access.
- Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016, June). Hierarchical attention networks for document classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies* (pp. 1480-1489).
- Yin, Y., Song, Y., & Zhang, M. (2017, September). Document-level multi-aspect sentiment classification as machine comprehension. In *Proceedings of the 2017 conference on empirical methods in natural language processing* (pp. 2044-2054).
- Lin, Z., Feng, M., Santos, C. N. D., Yu, M., Xiang, B., Zhou, B., & Bengio, Y. (2017). A structured self-attentive sentence embedding. *arXiv preprint arXiv:1703.03130*.
- Chen, H., Sun, M., Tu, C., Lin, Y., & Liu, Z. (2016, November). Neural sentiment classification with user and product attention. In *Proceedings of the 2016 conference on empirical methods in natural language processing* (pp. 1650-1659).
- Liu, G., & Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337, 325-338.
- Fu, X., Yang, J., Li, J., Fang, M., & Wang, H. (2018). Lexicon-enhanced LSTM with attention for general sentiment analysis. *IEEE Access*, 6, 71884-71891.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5, 135-146.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.
- ONAN, A. (2021). Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach. *Computer Applications in Engineering Education*, 29(3), 572-589.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Kardakis, S., Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2021). Examining attention mechanisms in deep learning models for sentiment analysis. *Applied Sciences*, 11(9), 3883.
- Cheng, J.; Dong, L.; Lapata, M. Long short-term memory-networks for machine reading. *arXiv 2016*, arXiv:1601.06733.
- Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, CA, USA, 12-17 June 2016; pp. 1480-1489.
- Onan, A. (2021). Ensemble of classifiers and term weighting schemes for sentiment analysis in Turkish. *Scientific Research Communications*, 1(1).