




# KİTLE FONLAMASINDAKİ PROJE METİN İÇERİKLERİNİN LSTM İLE ANALİZİ

## ANALYSIS OF PROJECT TEXT CONTENTS WITH LSTM IN CROWDFUNDING

Murat KILINÇ\*   
Can AYDIN\*\*   
Çiğdem TARHAN\*\*\* 

### Öz

Kitle fonlaması (KF) platformları, topluluklardan gelen finansal desteklerle projelerin fonlanarak hayata geçmesini sağlayan ve web ortamında kullanıcılara sunulan yeni nesil fonlama ve yatırım sistemleridir. Dünya çapında her yıl bu platformlar kullanılarak binlerce iş fikri çeşitli öznitelikler ile başarılı bir şekilde gerçekleştirilmektedir. KF başarısına en çok etki eden özniteliklerden birisi de projelerdeki metin içerikleridir. Bu doğrultuda yapılan araştırmada, Türkiye’de faaliyet gösteren KF platformlarındaki özetleyici proje metinleri veri kazıma teknikleriyle toplanmış ve analize hazır hale getirilmiştir. Sonrasında ise KF projelerinin metin içerikleri bir RNN modeli olan LSTM kullanılarak başarı etiketleriyle sınıflandırılmış ve değerlendirme metrikleriyle analiz edilmiştir. Parametre seçimleriyle birlikte kurulan modelin doğruluk oranı %96.18’dir. Çalışmanın sonuçları, KF projeleri için hazırlanan metinlerin karar destek sistemlerinde test edilebileceğini göstermektedir.

**Anahtar Kelimeler:** Kitle Fonlaması, metin madenciliği, sinir ağları, derin öğrenme, uzun-kısa süreli bellek  
**JEL Sınıflandırılması:** M13, O32, L26

### Abstract

Crowdfunding (CF) platforms are new generation funding and investment systems that enable projects to be funded with financial support from communities and offered to users on the web. Thousands of

\* **Sorumlu Yazar:** Lecturer, Manisa Celal Bayar University, Computer Research and Application Center, kilinc.murat@cbu.edu.tr, ORCID: 0000-0003-4092-5967.

\*\* Assoc. Prof., Dokuz Eylül University, Faculty of Economics and Administrative Sciences, Department of Management Information Systems, DEÜ-KALMER, can.aydin@deu.edu.tr, ORCID: 0000-0002-0133-9634.

\*\*\* Assoc. Prof., Dokuz Eylül University, Faculty of Economics and Administrative Sciences, Department of Management Information Systems, DEÜ-BİMER, cigdem.tarhan@deu.edu.tr, ORCID: 0000-0002-5891-0635.

**To cite this article:** Kılınç, M., Aydın, C. & Tarhan, Ç. (2022). Kitle fonlamasındaki proje metin içeriklerinin LSTM ile Analizi. *Journal of Research in Business*, 7(1), e48-59.

“Çalışmada etik kurul izni gerekmemektedir.”

**Başvuru:** 10.11.2021

**Düzeltilme:** 18.01.2022

**Kabul:** 02.03.2022

**Online Yayın:** 30.03.2022

business ideas are successfully implemented with various attributes by using these platforms worldwide every year. One of the attributes that most affect the success of CF is the text content in the projects. In the research conducted in this direction, the summary project texts in the CF platforms operating in Turkey were collected by data scraping techniques and made ready for analysis. Afterwards, the text contents of the CF projects were classified with success tags using an RNN model, LSTM, and analyzed with evaluation metrics. The accuracy rate of the model established with the parameter selections is 96.18%. The results of the study show that the texts prepared for CF projects can be tested in decision support systems.

**Keywords:** Crowdfunding, text mining, neural networks, deep learning, long-short term memory

**JEL Classification:** M13, O32, L26

## Extended Summary

Crowdfunding (CF) is a new generation funding and investment system based on small or large support from a large number of backers. CF platforms, which provide a great potential for the emergence of value-added products in the entrepreneurship ecosystem, also have a great impact on the realization of creative ideas. With this effect, CF has become a very popular area in the last 10 years, and every year billions of dollars of funds are transferred to the project owners through CF.

However, there has been a decrease in the success rates of the CF ecosystem from the past to the present. These decreases are due to the wrong preparation of the projects as well as the fact that the attributes are not determined correctly. When examined in detail, CF projects contain many attributes. Each attribute determined during project preparation can have an impact on success. The text content among these attributes is more difficult to analyze and classify. However, although it has a difficult structure to analyze, the content of the text is considered as the area where the project is explained to the supporters in detail and the supporters are convinced. For this reason, the binary classification (successful-unsuccessful) of the text prepared during the creation of CF projects before being presented to the supporters has the potential to provide decision support to the users.

In this direction, within the scope of our study, the text contents of the projects on the CF platforms operating in Turkey were collected by data scraping techniques, pre-processed and made ready for analysis. Afterwards, the classification of text contents was provided by using Long Short-Term Memory (LSTM), which is a type of Recurrent Neural Network (RNN). The values obtained in our study are presented in the findings section.

## 1. Giriş

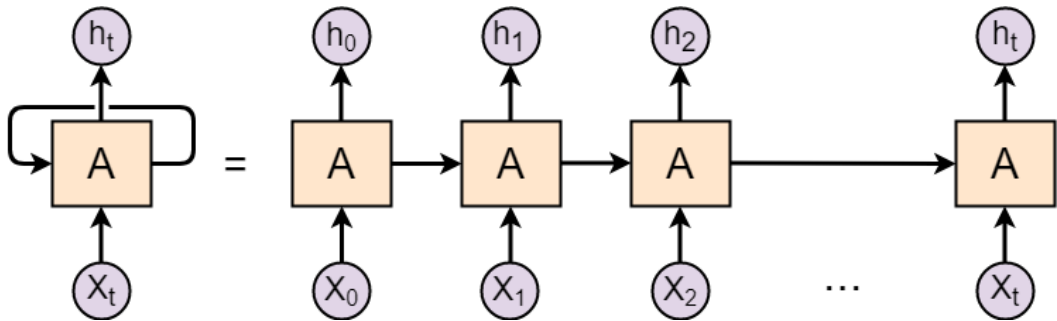
Kitle fonlaması (KF), çok sayıda destekçinin küçük ya da büyük desteklerine dayanan yeni nesil bir fonlama ve yatırım sistemidir. Girişimcilik ekosisteminde katma değerli ürünlerin ortaya çıkması için büyük bir potansiyel sağlayan KF platformlarının aynı zamanda yaratıcı fikirlerin hayata geçmesi noktasında da büyük bir etkisi vardır. Bu etkisi sayesinde KF, son 10 yıl içerisinde oldukça popüler bir alan haline gelmiştir (Borrero-Domínguez vd., 2020). Çünkü KF ekosistemi mikro finansa da benzemektedir. Yani, proje için gerekli fonlar tek bir kişi yerine birkaç kişiden toplanan ve nispeten düşük miktarlarda olan paranın bir araya getirilmesiyle oluşmuştur (Yan Li

vd., 2016). Bu sayede küçük yatırımcı ve destekçilerin ilgisi bu platformlar üzerinde yoğunlaşmıştır. Topluluklardan edinilen desteklerle yakın gelecekte KF platformlarındaki fonlama miktarının 90 milyar dolar civarına ulaşması beklenmektedir (Chakraborty ve Swinney, 2020). Sosyal girişim olarak da değerlendirilebilen KF'de bu kadar büyük miktarlarda fon toplanabilmesi sayesinde birçok inovatif fikir destek alabilmek için online sosyal platformları tercih etmektedir. Bunun da sebebi, COVID-19 salgını gibi büyük kriz durumlarında bile KF platformlarının girişimcilik ekosistemindeki olumsuz etkileri azaltabilmesinden kaynaklıdır (Farhoud vd., 2021). Bu doğrultuda, KF platformları sağladığı avantajlarıyla birlikte günümüzde birçok ülkede faaliyet göstermektedir. Türkiye'de faaliyet gösteren KF platformları bünyesinde birçok öznitelik barındırmaktadır. KF dinamiklerinin daha iyi anlaşılması ve kampanya başarı oranlarının iyileştirilmesi için, KF başarısına etki eden özniteliklerin bilinmesi gerekir (Lukkarinen vd., 2016). Bu öznitelikler arasında bulunan metin içerikleri analiz edilmesi ve sınıflandırılması daha zor bir yapıdadır. Fakat analizi zor bir yapıda olmasına rağmen, metin içerikleri projenin destekçilere detaylıca anlatıldığı ve destekçilerin ikna edildiği alan olarak değerlendirilmektedir. Bu sebeple, KF projeleri oluşturulurken hazırlanan metnin destekçilere sunulmadan önce başarı kapsamında ikili sınıflandırılması kullanıcılara karar desteği sağlama potansiyeline sahiptir. Bu potansiyel her iki taraf (proje oluşturucusu ve destekçi) için de fayda sağlar (Lukkarinen vd., 2016). Literatüre bakıldığında metin sınıflandırma ve analizi alanında birçok çalışma bulunmaktadır. 2021 yılında Moradi ve Badrinarayanan'ın yaptığı bir çalışmada KF finansmanında proje anlatımlarındaki dil stilleri ve anlatı uzunluğunun fonlama başarısına etkisi araştırılmıştır. Çalışmadaki 343 KF projesinden elde edilen verilerin analizi; bir markanın öne çıkarılmasının, anlatım uzunluğu ve dil stiline finansman başarısına olumlu bir şekilde etki ettiğini göstermektedir (Moradi ve Badrinarayanan, 2021). 2019 yılında Nergiz ve diğerlerinin yaptığı çalışmada Türkçe haber sitelerinde yer alan yayınlanan farklı kategorilerdeki haber içerikleri Long-Short Term Memory (LSTM) kullanılarak sınıflandırılmıştır. Doc2Vec, Word2Vec ve FastText modelleri uygulanan çalışmada başarı oranları karşılaştırılmış ve en iyi sonucu FastText modeli ortaya koymuştur (Nergiz vd., 2019). 2020 yılında Shneor ve diğerlerinin yaptığı çalışmada KF proje içeriklerindeki metin uzunluklarının proje başarısını olumlu yönde etkilediği görülmektedir (Shneor & Vik, 2020). Diğer bir deyişle proje destekçileri metinlerdeki detaylı açıklamalar ile fonlama için daha çok ikna olmaktadır. 2020 yılında Jang ve diğerlerinin yaptığı çalışmada internet film veritabanı (IMDb) içindeki film inceleme metinleri bi-LSTM+CNN hibrit modeliyle sınıflandırılmıştır. Çalışmada hibrit modelin metin sınıflandırmada doğruluk değeri daha yüksek sonuçlar elde ettiği görülmektedir (Jang vd., 2020). Elgenar ve diğerlerinin 2020 yılında yaptığı çalışmada arapça metin sınıflandırmada derin öğrenme yöntemleri kullanılmıştır. Çalışmada sınıflandırma performanslarının iyileştirilmesi için Word2Vec modeli kullanmanın etkisi incelenmiştir. Sonuçlara göre 10 kategori alt kümesi için %88,68 ile en yüksek genel sınıflandırma doğruluk değeri elde edilmiştir (Elnagar vd., 2020). Li ve diğerlerinin 2018'de yaptığı bir başka çalışmada Çince metinleri sınıflandırmak için LSTM ve evrişimli sinir ağı (CNN) kullanılmıştır. Hibrit bir şekilde uygulanan iki yöntemin çift yönlü (2 katman) LSTM, bir katman CNN şeklinde uygulanmasının (BLSTM-C) metin sınıflandırmada dikkate değer bir performans ortaya koyduğu görülmüştür (Yue Li vd., 2018). Bilgin ve Şentürk'ün 2017'de yaptığı çalışmada Twitter aracılığıyla ürünler ve firmalar hakkında kullanıcılardan toplanan geri bildirimler Doc2Vec modeliyle pozitif,

negatif ve nötr etiketli veriler üzerinde çalıştırılmış ve sonuçlar kaydedilerek duyarlılık analizi yapılmıştır. Sonuçlara göre Doc2Vec için yapılandırılan dağıtılmış bellek (DM) ve dağıtılmış kelime torbası (DBoW) arasından DBoW'un Türkçe metin duygu analizinde daha başarılı olduğu gözlemlenmiştir (Bilgin ve Şentürk, 2017). Literatür özetlenecek olursa, farklı alanlardaki metinlerin LSTM, Word2Vec, Doc2Vec modellerinin kullanılmasıyla birlikte yüksek doğruluk oranlarıyla sınıflandırılabilirdiği görülmektedir. Çok farklı alfabeler kullanılsa bile kurulan modelin metinsel verilerle eğitilmesiyle başarılı sonuçlar alınabilmektedir. Bu doğrultuda çalışmamız kapsamında Türkiye'de faaliyet gösteren KF platformlarındaki projelerin metin içerikleri veri kazıma teknikleriyle toplanıp ön işlemeden geçirilerek analize hazır hale getirilmiştir. Daha sonrasında ise bir RNN türü olan LSTM kullanılarak metin içeriklerinin sınıflandırılması sağlanmıştır. Çalışmamızda elde edilen değerler bulgular kısmında sunulmuştur.

## 2. Yöntem

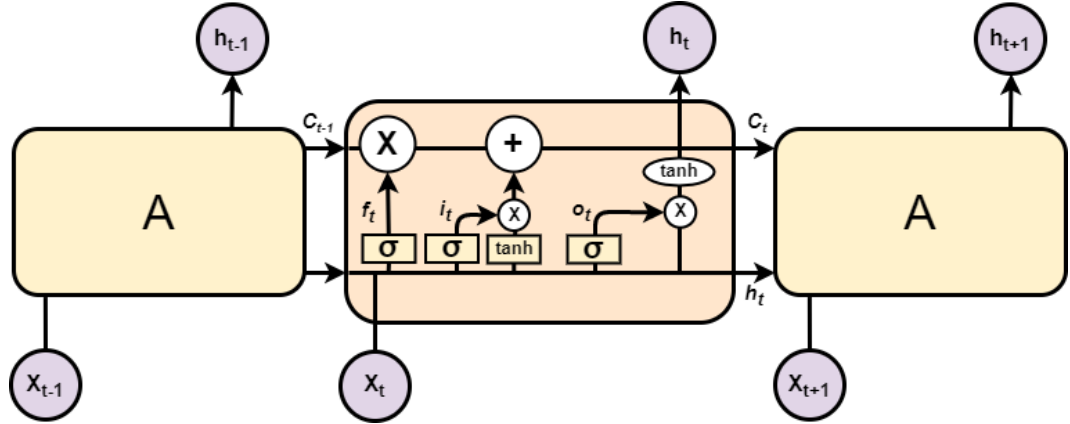
Sinir ağı modelleri çeşitli doküman, belge ya da metinlerde sınıflandırma açısından dikkate değer bir performans elde edebilmektedir (Zhou vd., 2015). Bu doğrultuda, tekrarlayan sinir ağı modeli olan RNN metin sınıflandırma alanında sıklıkla kullanılmakta ve gayet iyi işler çıkarabilmektedir (Şekil 1). Fakat bilgi teknolojilerinin ortaya çıkardığı veri sayısı çok fazla sayıda olduğu için RNN bazı durumlarda yetersiz kalabilmektedir. Diğer bir deyişle RNN, daha kısa bir yapıda hazırlanan cümleleri sınıflandırmada başarılı olurken, paragraf ya da daha uzun metinleri sınıflandırmada başarısız olmaktadır. Çünkü aktivasyon fonksiyonlarıyla birlikte oluşan girdilerdeki değişimler, metin uzun bir yapıya sahip olduğunda katmanlardaki öğrenmenin azalmasına sebep olarak kısa süreli bir öğrenme durumu ortaya çıkartabilir. Dolayısıyla bazı sinir ağı katmanlarında öğrenme olmadığı için RNN, uzun olan metinleri sınıflandırmada yetersiz kalmaktadır (Akça, 2021). Bu doğrultuda, çalışma kapsamında incelenen metinlerde daha uzun bir yapıya sahip olduğu için bir RNN türü olan ve uzun vadeli bağımlılıkları öğrenebilen LSTM kullanılmıştır.



Şekil 1: RNN Mimarisi (Nergiz vd., 2019)

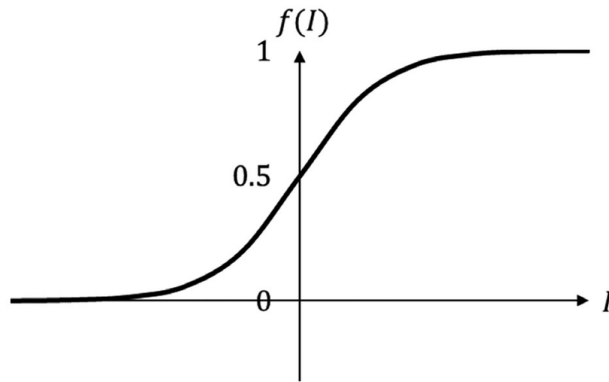
LSTM, depolama işlevi sayesinde standart olarak sunulan RNN yapısının kısa vadeli bellek problemini ortadan kaldırmayı sağlamaktadır (Akköse, 2020). LSTM modelleri kelime dizileri arasındaki uzun

vadeli bağımlılıkları yakalayarak metin sınıflandırması için daha iyi bir çözüm sunmaktadır (Jang vd., 2020). Başka bir deyişle, LSTM akışı içerisinde gelecekteki ve geçmişteki bilgiler anlamlandırılarak öznitelikler yinelenerek taşınmaktadır (Kızrak ve Bolat, 2019). Ayrıca tekrarlayan sinir ağlarının bünyesinde sadece bir tanjant kapısı bulunur (Şekil 2).



Şekil 2: LSTM Mimarisi (Basiri vd., 2021; Hu vd., 2020)

LSTM'de ise tekrarlayan bir yapıda olacak şekilde 3 farklı kapı mevcuttur. Bu kapılar, forget gate, input gate ve output gate olarak sıralanmaktadır. İlk olarak forget gate katmanında LSTM akışındaki hangi bilgilerin silineceğine, hangi bilginin unutulacağına karar verilir. Kararın verileceği aşamada sigmoid fonksiyonu ( $\sigma$ ) kullanılır. Sigmoid fonksiyonu tarafından 0-1 arası değerlere dönüştürülen bilgiler ile değer sıfır ise tamamen unutulması, değer 1 ise bilgide değişiklik yapılmadan bir sonraki katmana iletilmesi durumu gerçekleştirilir (1).



Şekil 3: Sigmoid Fonksiyonu (Matsubara vd., 2019)

$\sigma$  sigmoid fonksiyonu,  $W$  ağırlık matrisi,  $\tanh$  aktivasyon fonksiyonu olmak üzere,

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t * \tanh(C_t) \quad (5)$$

Bir sonraki adım olan input gate katmanında, LSTM döngüsü içerisinde gelen bilgilerden hangilerinin saklanacağına karar verilir. Bu kapı teknik olarak 2 kısımdan oluşmaktadır. Sigmoid fonksiyonuyla kullanılacak değerleri belirleyen input geçidi ( $i_t$ )(2), tanjant fonksiyonunun oluşturduğu vektörle birleşerek cell state ( $\tilde{C}_t$ ) üzerine eklenmekte ve güncellenmektedir (3). Son adım olarak ise output kapısı oluşturulmaktadır. Bu aşamada önce sigmoid fonksiyonu, sonrasında tanh fonksiyonu ile filtreleme işlemi uygulanır (4)(Akdoğan, 2020). Tüm işlemler uygulandıktan sonra cell state LSTM akışındaki diğer bir hücreye aktarılmaya hazır bir hale gelir (5).

**Tablo 1:** Kurulan Modele Ait Seçilmiş Parametreler

Optimizier Seçimi	Loss Parametresi	Aktivasyon Fonksiyonu
Adam*	Binary_crossentropy	Sigmoid

\*: *Adagrad, rmsprop ve sgd seçimlerine kıyasla daha iyi sonuçlar verdiği için Adam algoritması seçilmiştir.*

Diğer taraftan, kullanılan yöntemde daha iyi sınıflandırma yapılabilmesi için parameter seçimleri yapılmıştır (Tablo 1). Bu doğrultuda, kayıp fonksiyonu parametresi olarak seçilen binary\_crossentropy elde edilen örneğin ikili olarak sınıflandırıldığı problemlerde kullanılmaktadır. Optimizasyon algoritması seçiminde ise derin öğrenme çalışmalarında adagrad, rmsprop, sgd gibi algoritmalara kıyasla daha iyi sonuçlar veren adam algoritması seçilmiştir (Seyyarer vd., 2020). Son olarak, aktivasyon fonksiyonu seçiminde problemin doğrusal olmamasından dolayı sigmoid tercih edilmiştir. Sigmoid, elde edilecek çıktının hangi sınıf içerisinde olduğuna dair bir değer ortaya koymaktadır.

### 3.1. Veri ve Değişkenler

Çalışmada kullanılan veri seti Türkiye'de faaliyet gösteren KF platformlarından veri kazıma teknikleriyle oluşturulmuştur. Python ile derlenen BeautifulSoup ve Scrapy kütüphanelerinin yanısıra Octoparse yazılımıyla veri kazıma yapılmış, sorun olması durumunda ise veriler manuel olarak veri setine eklenmiştir. Ayrıca, çalışma için toplanan veriler halka açık platformlardan elde edilmiş olup (Fongogo, Crowdfon, Ideanest, FonBulucu) kazıma işleminde kullanıcıların herhangi

bir üyelik ya da gizlilik gerektirmeden ulaşabileceği meta alanları tercih edilmiştir (Tablo 2). Veriler eğitim aşamasında test verisi 0.33, eğitim verisi ise 0.67 olacak şekilde bölünmüştür.

**Tablo 2:** Veri Seti Örnek Gösterim

Proje Açıklaması*	Karakter Sayısı	Başarı Durumu
Bu proje bir sanat, yoga, yolculuk ve bunların birleşimiyle yolları birleştirme, hayalleri gerçekleştirme projesidir.	104	Başarısız
Dizimizin ilk bölümünü çekmek için kamera, ışık ve ses ekipmanları ile mekan kiralamanın yanı sıra, dekor, kostüm ve oyuncular için bu bütçeye ihtiyacımız var.	136	Başarılı
CoderDojo ile Çocuklar Kod Yazmayı Öğreniyor! CoderDojo Türkiye'nin Vakıf olmasına Destek olun.	84	Başarılı
Orman yangınlarında daha etkili müdahale edebilecek, her türlü arazi koşuluna uygun, zırhlı ve paletli itfaiye aracı ile artık ormanlarımız yok olmayacak.	134	Başarısız
Bu zor günlerde hayallerimi gerçekleştirdiğim giyim butiğimin ayakta kalmasını sağlayabilirsin. Maaş ve diğer giderlerimizi karşılamak için yanımızda olabilirsin.	145	Başarısız

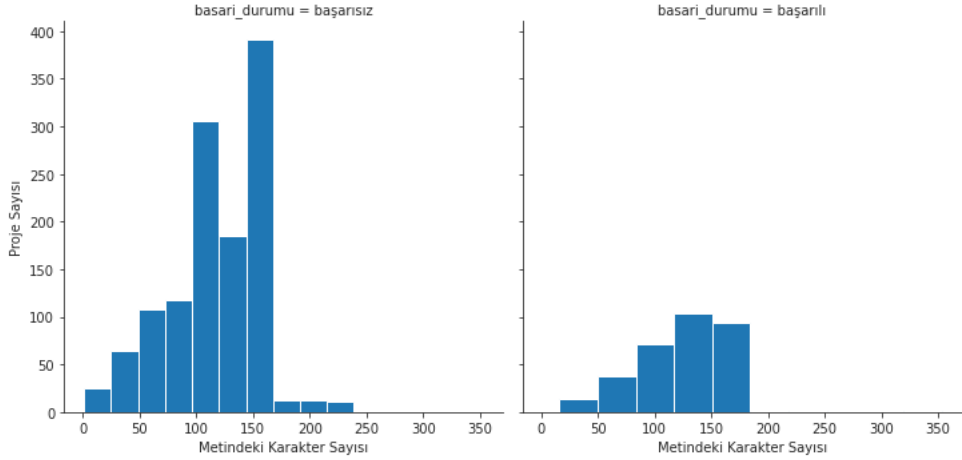
\*: Platformlarda yer alan projelerin destekçilere tanıtıldığı kısım olarak tanımlanmaktadır.

### 3.2. Veri Önışleme

Kullanılan veri setindeki metinler tek bir standarta sahip değildir. Bu şekliyle analiz edilmesi zor olduğu için proje metinleri içindeki noktalama işaretleri, emojiler ve özel semboller veri seti içerisinden çıkarılmış ve tüm karakterler küçük harf şeklinde tutulmuştur. Ayrıca Türkçe metinler içerisinde analiz sonuçlarına etki etmeyen ve yok sayılan stopwords'ler (etkisiz kelimeler) proje açıklamaları içerisinden çıkarılarak önışleme aşaması tamamlanmıştır.

### 4. Araştırma Bulguları

Türkiye'deki KF platformlarında destekçilere sunulan projeler detaylı metin açıklamalarının yanı sıra genellikle proje özetleyen bir yapıya sahip olan kısa açıklamalara da sahiptir. Açıklama içerisinde projenin neden desteklenmesi gerektiği, ne gibi destekler talep edildiği kullanıcılar tarafından belirtilerek özetlenmektedir. Veri kazıma ile oluşturulan veri setinde başarısız projeler genellikle 100-150 karakter içermektedir. Başarılı projelerde ise projelerin karakter sayısı büyük çoğunlukla 100-200 karakter arasına dağılmıştır. Başka bir deyişle, başarılı projelerin açıklama metinleri başarısız projelere göre daha uzun bir yapıdadır (Şekil 4).



Şekil 4: Karakter Sayısı ve Başarı Durumlarına Göre Metinlerin Karşılaştırılması

LSTM modeli, Google tarafından geliştirilen Colab üzerinde analiz edilmiştir. Google Colab, sağladığı TPU, Tesla K80 GPU ve normal CPU desteği sayesinde web üzerinde derin öğrenme yöntemlerinin kolaylıkla çalışmasını sağlamaktadır. Çalışma içerisinde veri boyutu çok büyük olmadığı için GPU kullanılmıştır.

Tablo 3: LSTM ile Metin Sınıflandırma Sonuçları

	Sayı	Doğruluk Oranı	Kayıp Değeri (Loss)	Precision	Recall	F1 Değeri
Eğitim Verisi	996					
Test Verisi	495	0.9618	0.0919	0.9690	0.9630	0.9570
Eğitim Adımı (Epoch)/Durma	30/26					

Not: Model eğitim süreci, sabitleme gerçekleştiği için 26. adımda durdurulmuştur.

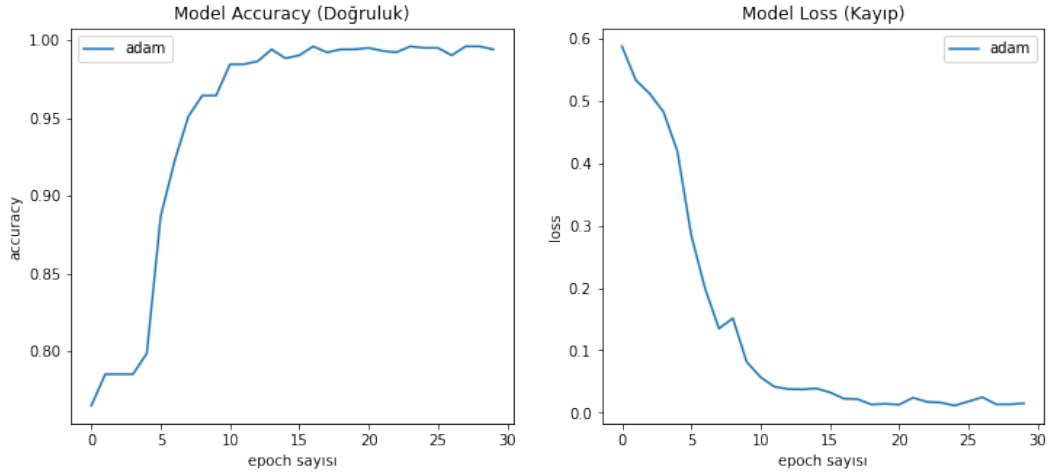
$$A_{cc} = \frac{TP + TN}{TP + FP + FN + TN} \quad (6), \quad F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (8), \quad Recall = \frac{TP}{TP + FN} \quad (9)$$

Elde edilen bulgulara göre, KF proje metinlerinin başarı sınıflandırması için LSTM doğruluk oranı %96.18, kayıp değeri %9.19'dur. Model değerlendirme metrikleri olan ve karmaşıklık matrisine göre belirlenen Precision (kesinlik), Recall (duyarlılık) ve F1 Skorları ise sırasıyla %96.90, %96.30 ve



%95.70 olarak bulunmuştur (Tablo 3). Metriklerin formülleri denklem (6-9) arasında belirtilmiştir. Ayrıca, eğitim için 996 veri kullanılarak 30 eğitim adımından oluşan öğrenme süreci oluşturulmuş ve 495 veriyle model test edilmiştir. Test verilerinin doğruluk oranı ise %78.2 olarak bulunmuştur. Tablo 1'deki parametrelerle uygulanan LSTM modelinin doğruluk ve kayıp grafikleri ise aşağıdaki gibidir (Şekil 5).



Şekil 5: Eğitim Adımlarına Göre Doğruluk ve Kayıp Grafikleri

## 5. Sonuç ve Tartışma

Kısa vadeli bellek problemini ortadan kaldıran LSTM, KF metinlerini başarı anlamında ikili sınıflandırmada geleneksel sinir ağları modelleri ve makine öğrenmesi algoritmalarına göre daha iyi sonuçlar vermektedir. Fakat her ne kadar kurulan model yüksek doğruluk oranıyla sınıflandırma yapsa da aşırı öğrenmeden kaçınmak için veri setindeki dengenin iyi kurulmuş olması gerekmektedir. Çünkü ikili sınıflandırmada, sınıflara ait veri sayısının eşit tutulması gerektiği literatürde de tavsiye edilmektedir (Ryoba vd., 2020). Çalışmamızdaki temel kısıt bu noktada ortaya çıkmaktadır. Türkiye'deki KF platformlarındaki proje verisi sayısı 1500-2000 civarındadır. Bu projelerden yaklaşık 500 tanesi başarılı geriye kalanlar ise başarısız projelerdir. Yani, çalışmada kullanılan veri setindeki sınıf dengesi yaklaşık olarak %30-%70 şeklindedir. Bu nedenle, eğitim verileriyle %96.18 doğruluk oranıyla sınıflandırma yapılabilirken, test verileriyle sınıflandırma yapıldığında bu oran %78.2'ye kadar düşmüştür. Bu sorun gelecek çalışmalarda veri sayısının artırılıp sınıf dengesinin kurulmasıyla çözülebilir. Sonuç olarak, Türkiye'de faaliyet gösteren KF projelerinin daha iyi bir şekilde hazırlanabilmesi için başarı tahmini odaklı daha fazla yaklaşıma ihtiyaç vardır. Bu ihtiyaç, KF projelerinin başarı oranının geçmişe göre giderek düşmesine odaklıdır. Dolayısıyla, KF kullanıcıları için oluşturulabilecek web tabanlı karar destek sistemlerinde proje metinlerinin analizi kesinlikle yapılmalıdır. Çünkü proje destekçilerinin fonlama için ikna edilmesinde en çok etkisi olan

özniteliklerden birisi proje metnidir. Bu doğrultuda, gelecek çalışmalarda veri sayısının arttırılmasının yanısıra, yapay veri üretme metotları incelenebilir, Türkçe metinler için hazırlanan detaylı stopwords kütüphaneleri benzer çalışmalar içerisinde dahil edilebilir. Ayrıca kelime gömme modellerinin veri setine uygulanmasıyla daha iyi öğrenme modelleri kurulabilir. Bu bilgiler doğrultusunda çalışmamız, proje başarı tahmini için oluşturulacak web tabanlı karar destek sistemlerinde metin sınıflandırma için derin öğrenme tabanlı tekniklerin entegre edilmesini önermektedir. Bu sayede, KF ekosisteminde daha kaliteli projeler oluşturulabilmesi için bir karar desteği oluşturulabilecektir.

### Yazar Katkısı

KATKI ORANI	AÇIKLAMA	KATKIDA BULUNANLAR
Fikir veya Kavram	Araştırma fikrini veya hipotezini oluşturmak	Murat KILINÇ Can AYDIN
Literatür Taraması	Çalışma için gerekli literatürü taramak	Murat KILINÇ
Araştırma Tasarımı	Çalışmanın yöntemini, ölçeğini ve desenini tasarlamak	Murat KILINÇ Can AYDIN Çiğdem TARHAN
Veri Toplama ve İşleme	Verileri toplamak, düzenlemek ve raporlamak	Murat KILINÇ
Tartışma ve Yorum	Bulguların değerlendirilmesinde ve sonuçlandırılmasında sorumluluk almak	Murat KILINÇ Can AYDIN Çiğdem TARHAN

### Çıkar Çatışması

Çalışmada yazarlar arasında çıkar çatışması yoktur.

### Finansal Destek

Bu araştırma TÜBİTAK tarafından 121E363 proje numarasıyla desteklenmiştir.

### Kaynakça

- Akça, M. F. (2021). *LSTM Nedir? Nasıl Çalışır?* Erişim Tarihi: 12.07.2021, Erişim Linki: <https://mfakca.medium.com/lstm-nedir-nasil-calisir-326866fd8869>.
- Akdoğan, A. (2020). *Uzun Kısa Vadeli Hafıza Ağları*. Erişim Tarihi: 18.07.2021, Erişim Linki: <https://medium.com/bilisim-hareketi/uzun-kisa-vadeli-hafiza-aglari-lstm-95cbe7d51b44>.
- Akköse, O. (2020). *Uzun-Kısa Vadeli Bellek (LSTM)*. Erişim Tarihi: 12.07.2021, Erişim Linki: <https://medium.com/deep-learning-turkiye/uzun-kisa-vadeli-bellek-lstm-b018c07174a3>.

- Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Generation Computer Systems*, 115, 279–294. <https://doi.org/10.1016/j.future.2020.08.005>.
- Bilgin, M., & Şentürk, İ. F. (2017). Sentiment analysis on Twitter data with semi-supervised Doc2Vec. *2nd International Conference on Computer Science and Engineering, UBMK 2017*, 661–666. <https://doi.org/10.1109/UBMK.2017.809.3492>.
- Borrero-Domínguez, C., Cordon-Lagares, E., & Hernández-Garrido, R. (2020). Analysis of success factors in crowdfunding projects based on rewards: A way to obtain financing for socially committed projects. *Heliyon*, 6(4). <https://doi.org/10.1016/j.heliyon.2020.e03744>.
- Chakraborty, S., & Swinney, R. (2020). Signaling to the Crowd : Private Quality Information and Rewards-Based Crowdfunding. *Manufacturing & Service Operations Management*, April, 0–15.
- Elnagar, A., Al-Debsi, R., & Einea, O. (2020). Arabic text classification using deep learning models. *Information Processing and Management*, 57(1), 102121. <https://doi.org/10.1016/j.ipm.2019.102121>.
- Farhoud, M., Shah, S., Stenholm, P., Kibler, E., Renko, M., & Terjesen, S. (2021). Social enterprise crowdfunding in an acute crisis. *Journal of Business Venturing Insights*, 15(November 2020), e00211. <https://doi.org/10.1016/j.jbvi.2020.e00211>.
- Hu, J., Wang, X., Zhang, Y., Zhang, D., Zhang, M., & Xue, J. (2020). Time Series Prediction Method Based on Variant LSTM Recurrent Neural Network. *Neural Processing Letters*, 52(2), 1485–1500. <https://doi.org/10.1007/s11063.020.10319-3>.
- Jang, B., Kim, M., Harerimana, G., Kang, S. U., & Kim, J. W. (2020). Bi-LSTM model to increase accuracy in text classification: Combining word2vec CNN and attention mechanism. *Applied Sciences (Switzerland)*, 10(17). <https://doi.org/10.3390/app10175841>.
- Kızrak, M. A., & Bolat, B. (2019). Uçak Motoru Sağlığı için Uzun-Kısa Süreli Bellek Yöntemi ile Öngörücü Bakım. *Bilişim Teknolojileri Dergisi*, 103–109. <https://doi.org/10.17671/gazibtd.495730>.
- Li, Yan, Rakesh, V., & Reddy, C. K. (2016). Project success prediction in crowdfunding environments. *WSDM 2016 – Proceedings of the 9th ACM International Conference on Web Search and Data Mining*, 247–256. <https://doi.org/10.1145/2835.776.2835791>.
- Li, Yue, Wang, X., & Xu, P. (2018). Chinese text classification model based on deep learning. *Future Internet*, 10(11). <https://doi.org/10.3390/fi10110113>.
- Lukkarinen, A., Teich, J. E., Wallenius, H., & Wallenius, J. (2016). Success drivers of online equity crowdfunding campaigns. *Decision Support Systems*, 87, 26–38. <https://doi.org/10.1016/j.dss.2016.04.006>.
- Matsubara, N., Teramoto, A., Saito, K., & Fujita, H. (2019). Generation of Pseudo Chest X-ray Images from Computed Tomographic Images by Nonlinear Transformation and Bone Enhancement. *Medical Imaging and Information Sciences*, 36(3), 141–146. <https://doi.org/10.11318/mii.36.141>.
- Moradi, M., & Badrinarayanan, V. (2021). The effects of brand prominence and narrative features on crowdfunding success for entrepreneurial aftermarket enterprises. *Journal of Business Research*, 124(November 2020), 286–298. <https://doi.org/10.1016/j.jbusres.2020.12.002>.
- Nergiz, G., Safali, Y., Avaroglu, E., & Erdogan, S. (2019). Classification of Turkish News Content by Deep Learning Based LSTM Using Fasttext Model. *2019 International Conference on Artificial Intelligence and Data Processing Symposium, IDAP 2019*, 1–6. <https://doi.org/10.1109/IDAP.2019.887.5949>.
- Ryoba, M. J., Qu, S., & Zhou, Y. (2020). Feature subset selection for predicting the success of crowdfunding project campaigns. *Electronic Markets*, 1–14. <https://doi.org/10.1007/s12525.020.00398-4>.
- Seyyarer, E., Ayata, F., Uçkan, T., & Karcı, A. (2020). Derin Öğrenmede Kullanılan Optimizasyon Algoritmalarının Uygulanması Ve Kıyaslanması. *Anatolian Journal of Computer Sciences*, 2, 90–98.

- Shneor, R., & Vik, A. A. (2020). Crowdfunding success: a systematic literature review 2010–2017. In *Baltic Journal of Management* (Vol. 15, Issue 2, pp. 149–182). Emerald Group Publishing Ltd. <https://doi.org/10.1108/BJM-04-2019-0148>.
- Zhou, C., Sun, C., Liu, Z., & Lau, F. C. M. (2015). *A C-LSTM Neural Network for Text Classification*. <http://arxiv.org/abs/1511.08630>.

## Özgeçmiş

**Murat KILINÇ (Öğr. Gör.)**, Manisa Celal Bayar Üniversitesi Bilgisayar Araştırma ve Uygulama Merkezi'nde öğretim görevlisi olarak görev yapmaktadır. Dokuz Eylül Üniversitesi Yönetim Bilişim Sistemleri doktora öğrencisidir. Kitle fonlaması, veri madenciliği, iş zekası, veri bilimi, yapay sinir ağları, makine öğrenmesi, web teknolojileri gibi konularda araştırmalar yapmaktadır.

**Can AYDIN (Doç. Dr.)**, Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Yönetim Bilişim Sistemleri Bölümünde Doç. Dr. olarak görev yapmaktadır. Aynı zamanda Dokuz Eylül Üniversitesi Bölgesel Kalkınma ve İşletme Bilimleri Araştırma ve Uygulama Merkezi (BİMER) müdür yardımcısıdır. Dokuz Eylül Üniversitesinde Coğrafi Bilgi Sistemleri Doktorasını almıştır. Anlamsal web, büyük veri, programlama, makine öğrenmesi, yapay sinir ağları, iş zekası gibi konularda araştırmalar yapmaktadır.

**Çiğdem TARHAN (Doç. Dr.)**, Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Yönetim Bilişim Sistemleri bölümünde Doç. Dr. olarak görev yapmaktadır. Aynı zamanda Dokuz Eylül Üniversitesi Bölgesel Kalkınma ve İşletme Bilimleri Araştırma ve Uygulama Merkezi (BİMER) müdürüdür. İzmir Yüksek Teknoloji Enstitüsünde Şehir Planlama Doktorasını almıştır. Coğrafi bilgi sistemleri, bilgi yönetim sistemleri, veri madenciliği, veri tabanı yönetimi, iş analitiği, dijital bölünme gibi konularda araştırmalar yapmaktadır.