



# İş Tatmini Faktörlerini Belirlemeye Ve Analiz Etmeye Yönelik Olarak Çalışanların Çevrimiçi Değerlendirmelerinin Sınıflandırıcı Topluluklarına Dayalı Analizi

Ali Özdemir<sup>1</sup>, Aytuğ Onan<sup>2</sup>, Vildan Çınarlı Ergene<sup>3\*</sup>

<sup>1</sup>Matematik /Fen Edebiyat Fakültesi, Manisa Celal Bayar Üniversitesi, Manisa, Türkiye, (ORCID: 0000-0001-9330-7084), [acaozdemir@gmail.com](mailto:acaozdemir@gmail.com)

<sup>2</sup>Bilgisayar Mühendisliği / Mühendislik Fakültesi, İzmir Katip Çelebi Üniversitesi, İzmir, Türkiye, (ORCID: 0000-0002-9434-5880), [aytugonan@gmail.com](mailto:aytugonan@gmail.com)

<sup>3\*</sup> Matematik / Fen Bilimleri Enstitüsü, Manisa Celal Bayar Üniversitesi, Manisa, Türkiye (ORCID: 0000-0002-1220-3337), [vildan.cinarli@gmail.com](mailto:vildan.cinarli@gmail.com)

(1st International Conference on Applied Engineering and Natural Sciences ICAENS 2021, November 1-3, 2021)

(DOI:10.31590/ejosat.1008691)

**ATIF/REFERENCE:** Özdemir, A., Onan, A., & Çınarlı Ergene, V. (2021). İş Tatmini Faktörlerini Belirlemeye Ve Analiz Etmeye Yönelik Olarak Çalışanların Çevrimiçi Değerlendirmelerinin Sınıflandırıcı Topluluklarına Dayalı Analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (28), 531-538.

## Öz

Bu çalışma kapsamında iş tatmini faktörlerini belirlemeye ve analiz etmeye yönelik olarak çalışanların çevrimiçi değerlendirmelerinin sınıflandırıcı topluluklarına dayalı analizi yapılmıştır. Metin sınıflandırma işlemi, önceden tanımlanmış kategoriler veya sınıflar altında metin belgelerinin otomatik olarak sınıflandırılmasıdır. Bu kapsamda makine öğrenmesi teknikleri, dokümanlara anahtar kelimeler atamak ve bunları belirli kategorilerde sınıflandırmak için kullanılır. Makine öğrenmesi, belgeleri otomatik olarak kategorilere ayırmamızı sağlar.

Metin madenciliği, metinleri organize etmek amacıyla önceden tanımlanmış kategorilere dahil etme işlemidir. Son yıllarda dijital ortamdaki verinin artması, bu verilerden bilgi çıkarımını ve bilginin kullanılabilirliğinin artırılması üzerine olan çalışmaları hızlandırmıştır. Metinlerin verimli kullanılabilmesi ve aranılan bilgiye hızlıca ulaşılabilmesi için metinlerin kategorize edilmesi önem kazanmıştır. Büyük belge koleksiyonlarını elle düzenlemek son derece zaman alıcıdır ve hataya açıktır. Otomatik metin sınıflandırma uygulamaları bu noktada hızlı ve maliyeti düşük çözümler sunar. Metin sınıflandırma işleminin, istenmeyen mesaj veya elektronik postaları filtreleme, bilimsel makalelerin otomatik indekslenmesi, web sayfaların etiketlenmesi, belge türünün tanımlanması gibi uygulama alanları mevcuttur.

Çalışanların çevrimiçi değerlendirmelerinin sınıflandırıcı topluluklarına dayalı analizi yapılması amacıyla çalışan kişilerin çalışmış oldukları firmalar hakkındaki düşünceleri tr.indeed.com'da yayınlanan toplam 18 farklı firmadan elde edilen 3023 adet çevrimiçi değerlendirme elde edilmiştir. Her bir veri seti için, öznitelik temsili yöntemlerinden terim sıklığı (TF), terim varlığı (TP) ve TF-IDF ölçütleri için 1-gram, 2-gram ve 3-gram temsilleri ile toplam dokuz farklı veri temsili elde edilmiştir. Yapılan deneysel çalışmalarda beş temel sınıflandırma algoritması topluluk öğrenme yöntemleri ile birleştirilerek doğru sınıflandırma performansları incelenmiştir.

**Anahtar Kelimeler:** Sınıflandırıcı topluluğu, Metin madenciliği, Makine öğrenmesi

## Analysis of Employees' Online Assessments Based on Classifier Ensembles to Identify and Analyze Job Satisfaction Factors

### Abstract

Within the scope of this study, an analysis of the online evaluations of employees based on classifier communities was made in order to identify and analyze job satisfaction factors. Text classification is the automatic classification of text documents under predefined categories or classes. In this context, machine learning techniques are used to assign keywords to documents and classify them into specific categories. Machine learning allows us to automatically categorize documents.

Text mining is the process of including texts into predefined categories for the purpose of organizing them. The increase in data in the digital environment in recent years has accelerated the studies on extracting information from these data and increasing the usability of information. It has become important to categorize the texts in order to use the texts efficiently and to reach the sought information quickly. Manually editing large document collections is extremely time-consuming and error-prone. Automatic text classification applications offer fast and cost-effective solutions at this point. Text classification has application areas such as filtering spam or e-mails, automatic indexing of scientific articles, tagging web pages, and defining document type.

\* Sorumlu Yazar: [vildan.cinarli@gmail.com](mailto:vildan.cinarli@gmail.com)

In order to analyze the online evaluations of the employees based on their classifier communities, 3023 online evaluations obtained from a total of 18 different companies, which were published on tr.indeed.com, were obtained. For each data set, a total of nine different data representations were obtained with 1-gram, 2-gram and 3-gram representations for term frequency (TF), term presence (TP) and TF-IDF criteria from the feature representation methods. In the experimental studies, the correct classification performances were examined by combining five basic classification algorithms with ensemble learning methods.

**Keywords:** Classifier ensemble, Text mining, Machine learning

## 1. Giriş

İşletmelerin hedeflerine ulaşması ve kar elde etmesi öncelikle çalışanlar sayesinde olur. Yani işini seven, işini iyi yapan, yaptığı işten gurur duyan, işletmenin hedeflerini benimsemiş ve amaçlarını işletme amaçlarıyla bütünleştirmiş çalışanlara sahip işletmeler başarılı olacaklardır. Bu da çalışanların işinden duydukları tatmin derecesiyle ilişkilidir. İş tatmini yöneticiler için çalışanların işlerine karşı tutumlarının performans ve verimlilik üzerindeki etkisi açısından önemlidir. Çünkü iş tatminsizliğinin doğrudan neden olduğu davranışlar; işten ayrılma, devamsızlık, performans düşüklüğü, ruhsal ve fiziksel sağlıkta bozulma şeklinde görülmektedir [1].

Günümüzde modern yönetim anlayışı çalışanın iş tatmini üzerinde önemle durmaktadır. Tüm kurumlarda kurum başarısı çalışanın iş doyumunu ile doğru orantılıdır. Kurumların kullandıkları kaynaklar içinde en karmaşık ve değişken olanı insan gücüdür. Örgütsel davranış alanında iş tatmini, en çok önem verilen ve sıklıkla araştırılan bir davranış olarak görülmektedir [2].

Yapay zekâ çözümleri her sektörde artarak devam etmektedir. Metin madenciliği son yıllarda hızla gelişen bir alan haline geldi ve bu alandaki uygulamalar yapay zekâ alanındaki uygulamaların büyük bir bölümünü oluşturmaktadır. Diğer fonksiyonel departmanlar gibi, insan kaynakları departmanlarının işletme yöneticileri çeşitli konularda kararlar almalıdır. Metinsel veriler de dahil olmak üzere daha fazla veri oluşturulup biriktikçe şirketleri ile ilgili karar verirken metin madenciliği kullanılarak bu tür verilerden uygulanabilir bilgi elde edilir [3].

İnsan kaynakları yönetiminde yapay zekâ teknolojilerinin kullanılması; iş için doğru adayları seçmekte ve yetenek kazanmalarına yardımcı olmakta, işyerinde çalışanların elde tutulma oranının tahmin edilmesine ve şirketteki idari personel üzerindeki yükü azaltmaya yardımcı olmaktadır. Aynı zamanda, hata olasılığını minimum seviyeye indirmeye, çalışanların daha yüksek performans ile katılım sağlamasına ve karar vermede önyargı davranışını en aza indirmeye yardımcı olmaktadır [4].

## 2. Materyal ve Metot

Sınıflandırma işlemi daha önce görülmemiş ve kategorisi bilinmeyen her bir örneğin, eğitim verisi kategorilerinden en uygun olan kategoriye atanması işlemidir [5]. Makine öğrenmesi açısından düşünüldüğünde ise, yapılan son araştırmalar, metin madenciliği gibi yöntemler üzerine olmuştur. Makine öğrenmesine ihtiyaç duyulmasının nedeni, elle kategorizasyonun pahalı ve zaman tüketen bir iş oluşudur ki, ayrıca elle sınıflandırmada, sınıflandırmayı yapan uzmanların vermiş oldukları kararlara bağlı olarak sonuçlar da değişmektedir [6].

Çalışanların şirketleri hakkında yaptıkları yorumlardan 1 ve 2 puan olanlar negatif, 3,4 ve 5 puanlar pozitif olarak değerlendirilmiş olup toplamda 18 farklı firmadaki 2321 pozitif yorum ve 702 negatif yorum tr.indeed.com'dan elde edilmiştir.

Elde edilen 3023 adet çevrimiçi değerlendirme ele alınarak sınıflandırıcı topluluklarına dayalı analizleri gerçekleştirilmiştir.

Öncelikle sistemde kullanılacak olan verilerin WEKA programında hataları düzeltilerek ön işlemden geçirilmiştir. Ön işlemden geçirilen çevrimiçi değerlendirmeler, öznelik temsili yöntemlerinden terim sıklığı (Term Frequency-TF), terim varlığı (Term Present-TP) ve TF-IDF ölçütleri için N-gram temsil modellerinden olan 1-gram, 2-gram ve 3-gram temsilleri ile toplam dokuz farklı veri temsili elde edilmiştir. Daha sonra çıkarılan öznelikler ağırlıklandırılmıştır. Son olarak elde edilen veriler arff (Attribute Relational File Format) formatına dönüştürülerek sınıflandırma işlemi gerçekleştirilmiştir. Çevrimiçi değerlendirme yapılan firmaların adı ve kaç adet pozitif ve negatif yorum yapıldığını belirten bilgiler Tablo 1'de verilmiştir:

Tablo 1. Çevrimiçi değerlendirme yapılan firmalara ait yorum sayıları

Firma adı	Negatif (1-2 puan)	Pozitif (3-4-5 puan)	Toplam Firma Sayısı
VESTEL	16	85	101
BOSCH	3	37	40
ARÇELİK	23	151	174
SIEMENS	2	33	35
MC DONALDS	54	151	205
BURGER KING	49	170	219
DOMİNOS	21	13	89
KOTON	43	0	173
LC WAIKIKI	49	337	386
DE FACTO	32	189	221
STARBUCKS	8	86	94
KAHVE DÜNYASI	16	38	54
TURKCELL	24	0	158
TURK TELEKOM	33	0	212
VODAFONE	23	0	161
BİM A.Ş.	39	113	152
A101	260	248	508
MIGROS	7	34	41

Çalışmamızda 5 temel sınıflandırıcı olarak NaiveBayesMultinomial, lazy.IBk, Random.forest, functions.Logistic ve functions.LibLINEAR algoritmaları kullanmış olup kullanılan öğrenme yöntemleri aşağıdaki gibidir;

- Bu 5 temel algoritmanın tümü Meta Adaboost algoritması için ayrı ayrı denenerak 9 konfigürasyon için 5 farklı analiz yapıldı.
- Benzer şekilde 5 temel algoritmanın tümü Meta Bagging algoritması için ayrı ayrı denenerak 9 konfigürasyon için 5 farklı analiz yapıldı.
- 5 temel algoritmanın tümü Meta Random Subspace algoritması için ayrı ayrı denenerak 9 konfigürasyon için 5 farklı analiz yapıldı.
- Stacking (L1 regularized L2 loss support vector classification) algoritması 5 algoritma kullanılarak Logistic ile çalıştırılarak 1 analiz yapıldı.
- Meta Vote algoritması ise içerisindeki 5 tane farklı kombinasyon (Liblinear+L1 regularized logistic

regression) ile (meta.Vote AVG- meta.Vote PROD- meta.Vote MAJORITY- meta.Vote MIN- meta.Vote MAX ) çalıştırıldı.

Buna göre toplamda  $21 \cdot 9 = 189$  tane deney elde edilmiş oldu.

### 3. Araştırma Sonuçları ve Tartışma

Sınıflandırmaya ait deneysel sonuçlar Weka tabanlı Java uygulaması ile elde edilmiş olup beş temel sınıflandırma algoritması topluluk öğrenme yöntemleri ile birleştirilmiştir. Bu deneylerin sonuçlarına ait veriler tablolar ile gösterilmiştir (Tablo 2-6).

Tablo2. Temel algoritmaların AdaBoost sınıflandırıcı topluluğuna dayalı analizleri

Algoritma	Veri seti için elde edilen sonuçlar				
	Correctly Classified Instances	Precision	Recall	F-Measure	Time taken to build model:
<b>meta.AdaBoostM1-bayes.NaiveBayesMultinomial</b>					
TF_bigram.arff	75.1902%	0,734	0,752	0,741	0.05 seconds
TF_trigram.arff	74.8594%	0,707	0,749	0,717	0.03 seconds
TF_unigram.arff	75.2564%	0,761	0,753	0,756	0.05 seconds
TFIDF_bigram.arff	71.8822%	0,732	0,719	0,725	0.03 seconds
TFIDF_trigram.arff	71.8492%	0,693	0,718	0,704	0.03 seconds
TFIDF_unigram.arff	74.0655%	0,753	0,741	0,746	0.04 seconds
TP_bigram.arff	72.8746%	0,732	0,729	0,73	0.04 seconds
TP_trigram.arff	74.6279%	0,707	0,746	0,718	0.04 seconds
TP_unigram.arff	74.5617%	0,754	0,746	0,749	0.04 seconds
<b>meta.AdaBoostM1-functions.LibLINEAR</b>					
TF_bigram.arff	74.3632%	0,726	0,744	0,733	0.21 seconds
TF_trigram.arff	75.3887%	0,701	0,754	0,708	0.14 seconds
TF_unigram.arff	73.5693%	0,741	0,736	0,738	1.55 seconds
TFIDF_bigram.arff	73.4370%	0,72	0,734	0,726	1.5 seconds
TFIDF_trigram.arff	74.5617%	0,691	0,746	0,703	1.29 seconds
TFIDF_unigram.arff	73.4039%	0,741	0,734	0,738	2.33 seconds
TP_bigram.arff	73.6024%	0,722	0,736	0,728	0.31 seconds
TP_trigram.arff	74.9917%	0,698	0,75	0,708	0.18 seconds
TP_unigram.arff	73.4039%	0,741	0,734	0,737	2.25 seconds
<b>meta.AdaBoostM1-functions.Logistic</b>					
TF_bigram.arff	70.7906%	0,705	0,708	0,706	1939.32 seconds
TF_trigram.arff	71.5184%	0,681	0,715	0,695	267.4 seconds
TF_unigram.arff	67.4165%	0,705	0,674	0,687	4692.38 seconds
TFIDF_bigram.arff	70.9229%	0,706	0,709	0,708	1365.47 seconds
TFIDF_trigram.arff	71.5845%	0,682	0,716	0,695	178.94 seconds
TFIDF_unigram.arff	67.4496%	0,708	0,674	0,688	2137.57 seconds
TP_bigram.arff	70.7575%	0,704	0,708	0,706	1223.26 seconds
TP_trigram.arff	71.3860%	0,68	0,714	0,693	194.26 seconds

TP_unigram.arff	67.1849%	0,706	0,672	0,686	2318.24 seconds
<b>meta.AdaBoostM1-trees.RandomForest</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	75.5872%	0,727	0,756	0,735	628.27 seconds
TF_trigram.arff	75.9510%	0,706	0,76	0,709	953.79 seconds
TF_unigram.arff	81.2107%	0,798	0,812	0,812	120.28 seconds
TFIDF_bigram.arff	75.8518%	0,729	0,759	0,737	461.39 seconds
TFIDF_trigram.arff	76.0172%	0,707	0,76	0,709	941.56 seconds
TFIDF_unigram.arff	80.6153%	0,79	0,806	0,774	117.38 seconds
TP_bigram.arff	76.4472%	0,736	0,764	0,743	453.26 seconds
TP_trigram.arff	75.9180%	0,705	0,759	0,708	926.5 seconds
TP_unigram.arff	80.2845%	0,785	0,803	0,77	161.31 seconds
<b>meta.AdaBoostM1-lazy ibk</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	72.5438%	0,69	0,725	0,703	12.22 seconds
TF_trigram.arff	75.4879%	0,682	0,755	0,689	9.6 seconds
TF_unigram.arff	69.5336%	0,669	0,695	0,68	27 seconds
TFIDF_bigram.arff	72.2461%	0,687	0,722	0,7	10.28 seconds
TFIDF_trigram.arff	75.5210%	0,683	0,755	0,69	9.26 seconds
TFIDF_unigram.arff	70.7244%	0,666	0,707	0,682	29.05 seconds
TP_bigram.arff	74.9587%	0,701	0,75	0,711	10.44 seconds
TP_trigram.arff	75.7195%	0,686	0,757	0,69	9.21 seconds
TP_unigram.arff	70.7244%	0,666	0,707	0,682	26.76 seconds

Tablo3. Stacking sınıflandırıcı topluluğuna dayalı analizler

Algoritma	Veri seti için elde edilen sonuçlar				
meta.Stacking	Correctly Classified Instances	Precision	Recall	F-Measure	Time taken to build model:
TF_bigram.arff	80.2845%	0,786	0,803	0,768	1334.02 seconds
TF_trigram.arff	78.2666%	0,77	0,783	0,714	2308.68 seconds
<b>TF_unigram.arff</b>	<b>83.6586%</b>	<b>0,826</b>	<b>0,837</b>	<b>0,823</b>	<b>541.96 seconds</b>
TFIDF_bigram.arff	79.3913%	0,77	0,794	0,768	1343.01 seconds
TFIDF_trigram.arff	77.0427%	0,724	0,77	0,693	2352.92 seconds
TFIDF_unigram.arff	82.5339%	0,814	0,825	0,816	525.22 seconds
TP_bigram.arff	80.4499%	0,786	0,804	0,776	1329.51 seconds
TP_trigram.arff	78.2005%	0,758	0,782	0,72	2327.57 seconds
TP_unigram.arff	83.3278%	0,822	0,833	0,821	523.04 seconds

Tablo4. Temel algoritmaların Bagging sınıflandırıcı topluluğuna dayalı analizleri

Algoritma	Veri seti için elde edilen sonuçlar				
meta.Bagging-bayes.NaiveBayesMultinomial	Correctly Classified Instances	Precision	Recall	F-Measure	Time taken to build model:
TF_bigram.arff	80.5491%	0,788	0,805	0,774	0.1 seconds
TF_trigram.arff	78.2997%	0,778	0,783	0,711	0.02 seconds
TF_unigram.arff	83.2947%	0,822	0,833	0,823	0.02 seconds
TFIDF_bigram.arff	76.4803%	0,76	0,765	0,762	0.02 seconds
TFIDF_trigram.arff	72.1469%	0,697	0,721	0,707	0.01 seconds
TFIDF_unigram.arff	80.6153%	0,806	0,806	0,806	0.03 seconds
TP_bigram.arff	80.0198%	0,781	0,8	0,783	0.01 seconds

Tablo 4'ün devamı

TP_trigram.arff	76.9765%	0,727	0,77	0,724	0.02 seconds
TP_unigram.arff	82.4677%	0,816	0,825	0,819	0.02 seconds
<b>meta.Bagging- functions.LibLINEAR</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	76.5465%	0,735	0,765	0,741	0.3 seconds
TF_trigram.arff	76.1164%	0,701	0,761	0,701	0.14 seconds
TF_unigram.arff	78.7297%	0,776	0,787	0,781	0.65 seconds
TFIDF_bigram.arff	73.0069%	0,719	0,73	0,724	1.35 seconds
TFIDF_trigram.arff	74.6279%	0,686	0,746	0,698	1.23 seconds
TFIDF_unigram.arff	74.1317%	0,751	0,741	0,746	1.56 seconds
TP_bigram.arff	75.2564%	0,726	0,753	0,735	0.22 seconds
TP_trigram.arff	75.5541%	0,697	0,756	0,703	0.17 seconds
TP_unigram.arff	77.2742%	0,768	0,773	0,77	0.82 seconds
<b>meta.Bagging- functions.Logistic</b>	<b>Correctly Classified Instance</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	70.3275%	0,711	0,703	0,707	472.89 seconds
TF_trigram.arff	74.0324%	0,692	0,74	0,705	248.9 seconds
TF_unigram.arff	67.5157%	0,705	0,675	0,688	208.56 seconds
TFIDF_bigram.arff	70.3936%	0,711	0,704	0,707	475.89 seconds
TFIDF_trigram.arff	74.1317%	0,693	0,741	0,706	246.83 seconds
TFIDF_unigram.arff	67.0857%	0,699	0,671	0,683	208.5 seconds
TP_bigram.arff	70.0298%	0,711	0,7	0,705	452.7 seconds
TP_trigram.arff	73.9332%	0,692	0,739	0,705	238.57 seconds
<b>TP_unigram.arff</b>	<b>66.9864%</b>	<b>0,698</b>	<b>0,67</b>	<b>0,682</b>	<b>204.44 seconds</b>
<b>meta.Bagging -lazy.IBk</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	73.2716%	0,687	0,733	0,701	0.05 seconds
TF_trigram.arff	75.7195%	0,686	0,757	0,69	0.01 seconds
TF_unigram.arff	70.6583%	0,668	0,707	0,683	0.04 seconds
TFIDF_bigram.arff	72.8746%	0,683	0,729	0,698	0.03 seconds
TFIDF_trigram.arff	75.7526%	0,687	0,758	0,691	0.02 seconds
TFIDF_unigram.arff	71.5845%	0,662	0,716	0,681	0.03 seconds
TP_bigram.arff	75.2564%	0,697	0,753	0,705	0.04 seconds
TP_trigram.arff	75.7856%	0,686	0,758	0,69	0.02 seconds
TP_unigram.arff	71.5845%	0,662	0,716	0,681	0.03 seconds
<b>meta.Bagging- trees.RandomForest</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	77.3073%	0,741	0,773	0,744	883.3 seconds
TF_trigram.arff	76.7450%	0,715	0,767	0,704	1085.66 second
TF_unigram.arff	81.4092%	0,811	0,814	0,777	353.69 seconds
TFIDF_bigram.arff	77.2081%	0,739	0,772	0,743	914 seconds
TFIDF_trigram.arff	76.6126%	0,711	0,766	0,702	1577.89 seconds
TFIDF_unigram.arff	81.5746%	0,816	0,816	0,777	334.85 seconds
TP_bigram.arff	77.4727%	0,743	0,775	0,746	899.55 seconds
TP_trigram.arff	76.7119%	0,714	0,767	0,703	1575.37 seconds
TP_unigram.arff	81.5084%	0,818	0,815	0,775	339.37 seconds

Tablo5. Temel algoritmaların RandomSubSpace sınıflandırıcı topluluğuna dayalı analizi

Algoritma	Veri seti için elde edilen sonuçlar				
<b>meta.RandomSubSpace-functions.Logistic</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	77.7373%	0,747	0,777	0,749	425.04 seconds
TF_trigram.arff	75.5872%	0,702	0,756	0,708	60.93 seconds
TF_unigram.arff	80.5491%	0,787	0,805	0,787	153.86 seconds
TFIDF_bigram.arff	77.5719%	0,744	0,776	0,746	426.64 seconds
TFIDF_trigram.arff	75.1571%	0,695	0,752	0,704	62.15 seconds
TFIDF_unigram.arff	80.8138%	0,791	0,808	0,79	130.94 seconds
TP_bigram.arff	77.6712%	0,747	0,777	0,75	399.04 seconds
TP_trigram.arff	75.4548%	0,699	0,755	0,706	54.91 seconds
TP_unigram.arff	80.0860%	0,782	0,801	0,784	138.53 seconds
<b>meta.RandomSubSpace-functions.LibLINEAR</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	78.8621%	0,762	0,789	0,751	0.36 seconds
TF_trigram.arff	76.2818%	0,7	0,763	0,695	0.3 seconds
TF_unigram.arff	81.4423%	0,799	0,814	0,794	0.78 seconds
TFIDF_bigram.arff	77.4727%	0,744	0,775	0,747	1.47 seconds
TFIDF_trigram.arff	76.1826%	0,706	0,762	0,705	1.11 seconds
TFIDF_unigram.arff	80.4499%	0,788	0,804	0,791	2.1 seconds
TP_bigram.arff	78.2997%	0,753	0,783	0,75	0.44 seconds
TP_trigram.arff	76.2818%	0,707	0,763	0,704	0.35 seconds
TP_unigram.arff	81.1446%	0,795	0,811	0,795	0.94 seconds
<b>meta.RandomSubSpace-RandomForest</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	78.6636%	0,758	0,787	0,749	531.26 seconds
TF_trigram.arff	76.9765%	0,72	0,77	0,697	1388 seconds
TF_unigram.arff	81.5415%	0,817	0,815	0,776	249.26 seconds
TFIDF_bigram.arff	78.6636%	0,758	0,787	0,746	544.84 seconds
TFIDF_trigram.arff	77.2412%	0,73	0,772	0,701	1400.58 seconds
TFIDF_unigram.arff	81.6077%	0,821	0,816	0,776	231.78 seconds
TP_bigram.arff	78.5643%	0,757	0,786	0,748	591.22 seconds
TP_trigram.arff	76.9765%	0,72	0,77	0,698	756.34 seconds
TP_unigram.arff	81.8062%	0,825	0,818	0,778	379.82 seconds
<b>meta.RandomSubSpace-NaiveBayesMultinomial</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	79.4244%	0,785	0,794	0,741	0.27 seconds
TF_trigram.arff	77.4065%	0,798	0,774	0,683	0.69 seconds
TF_unigram.arff	82.8316%	0,819	0,828	0,807	0.66 seconds
TFIDF_bigram.arff	78.3989%	0,77	0,784	0,775	0.41 seconds
TFIDF_trigram.arff	74.4955%	0,706	0,745	0,717	0.32 seconds
TFIDF_unigram.arff	81.0453%	0,804	0,81	0,807	0.6 seconds
TP_bigram.arff	80.2514%	0,787	0,803	0,766	0.41 seconds
TP_trigram.arff	77.9027%	0,773	0,779	0,701	0.32 seconds
TP_unigram.arff	82.7985%	0,816	0,828	0,814	0.59 seconds
<b>meta.RandomSubSpace-LazyIBK</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	76.7450%	0,72	0,767	0,716	0.27 seconds

Tablo 5'in devamı

TF_trigram.arff	76.9434%	0,72	0,769	0,7	0.33 seconds
TF_unigram.arff	77.2742%	0,73	0,773	0,712	0.58 seconds
TFIDF_bigram.arff	76.9765%	0,725	0,77	0,718	0.42 seconds
TFIDF_trigram.arff	76.8773%	0,717	0,769	0,699	0.35 seconds
TFIDF_unigram.arff	77.4727%	0,763	0,775	0,69	0.59 seconds
TP_bigram.arff	77.3404%	0,732	0,773	0,714	0.38 seconds
TP_trigram.arff	76.7780%	0,714	0,768	0,699	0.34 seconds
TP_unigram.arff	77.5389%	0,764	0,775	0,692	0.59 seconds

Tablo6. Vote sınıflandırıcı topluluğuna dayalı analizler

Algoritma	Veri seti için elde edilen sonuçlar				
meta.Vote AVG	Correctly Classified Instances	Precision	Recall	F-Measure	Time taken to build model:
TF_bigram.arff	78.7297%	0,759	0,787	0,749	105.46 seconds
TF_trigram.arff	76.8111%	0,716	0,768	0,7	86 seconds
TF_unigram.arff	81.7400%	0,803	0,817	0,797	43.74 seconds
TFIDF_bigram.arff	76.6457%	0,743	0,766	0,75	47.93 seconds
TFIDF_trigram.arff	74.9587%	0,696	0,75	0,706	188.09 seconds
TFIDF_unigram.arff	79.5237%	0,78	0,795	0,784	17.09 seconds
TP_bigram.arff	78.8621%	0,761	0,789	0,754	93.01 seconds
TP_trigram.arff	76.5134%	0,71	0,765	0,704	102 seconds
TP_unigram.arff	81.6408%	0,801	0,816	0,798	38.57 seconds
<b>meta.Vote PROD</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	73.2054%	0,789	0,806	0,755	45.31 seconds
TF_trigram.arff	73.6024	0,762	0,782	0,69	74.18 second
TF_unigram.arff	74.4625%	0,823	0,836	0,817	57.11 seconds
TFIDF_bigram.arff	69.2689%	0,742	0,77	0,75	44.56 seconds
TFIDF_trigram.arff	72.6100%	0,689	0,752	0,701	296.1 seconds
TFIDF_unigram.arff	67.7804%	0,772	0,785	0,777	17.5 seconds
TP_bigram.arff	71.3199%	0,777	0,802	0,759	128.82 seconds
TP_trigram.arff	72.5438%	0,763	0,784	0,695	75.2 seconds
TP_unigram.arff	72.6761%	0,763	0,784	0,695	40.42 seconds
<b>meta.Vote MAJORITY</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	79.1598%	0,767	0,792	0,756	72.41 seconds
TF_trigram.arff	77.0096%	0,722	0,77	0,704	79.05 seconds
TF_unigram.arff	82.4347%	0,813	0,824	0,802	44.61 seconds
TFIDF_bigram.arff	76.9434%	0,744	0,769	0,75	75.48 seconds
TFIDF_trigram.arff	74.9587%	0,689	0,75	0,699	124.51 seconds
TFIDF_unigram.arff	80.5491%	0,788	0,805	0,788	27.98 seconds
TP_bigram.arff	78.7297%	0,76	0,787	0,755	74.97 seconds
TP_trigram.arff	76.4142%	0,711	0,764	0,708	132.91 seconds
TP_unigram.arff	82.0046%	0,807	0,82	0,798	29.42 seconds
<b>meta.Vote MIN</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>

Tablo 6'nın devamı

TF_bigram.arff	73.2385%	0,789	0,806	0,755	135.3 seconds
TF_trigram.arff	73.6355%	0,763	0,782	0,69	230.42 seconds
TF_unigram.arff	74.4625%	0,823	0,836	0,817	52.34 seconds
TFIDF_bigram.arff	69.3351%	0,742	0,77	0,749	129.83 seconds
TFIDF_trigram.arff	72.6100%	0,689	0,752	0,701	227.61 seconds
TFIDF_unigram.arff	67.8465%	0,772	0,784	0,777	49.94 seconds
TP_bigram.arff	71.3199%	0,777	0,802	0,759	71.51 seconds
TP_trigram.arff	72.5438%	0,763	0,784	0,695	129.84 seconds
TP_unigram.arff	72.6761%	0,809	0,824	0,807	29.3 seconds
<b>meta.Vote MAX</b>	<b>Correctly Classified Instances</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>Time taken to build model:</b>
TF_bigram.arff	76.8111	0,739	0,768	0,745	75.87 second
TF_trigram.arff	76.1495%	0,705	0,761	0,705	180.19 seconds
TF_unigram.arff	78.3328%	0,778	0,783	0,781	30.1 seconds
TFIDF_bigram.arff	72.3123%	0,721	0,723	0,722	105.75 seconds
TFIDF_trigram.arff	73.8339%	0,686	0,738	0,7	161.77 seconds
TFIDF_unigram.arff	72.5438%	0,753	0,725	0,736	32.15 seconds
TP_bigram.arff	75.1571%	0,73	0,752	0,737	103.44 seconds
TP_trigram.arff	75.5210%	0,7	0,755	0,707	124 seconds
TP_unigram.arff	76.3811%	0,766	0,764	0,765	29.55 seconds

Makine öğrenmesi teknikleri ile hızlı ve güvenilir şekilde tahminlemeler yapılabilmektedir. Makine öğrenmesine ait birçok algoritma mevcut olup, problemin kaynağına ve veri sayısına göre hangi algoritmanın kullanılacağına karar verilmektedir. Farklı algoritmalar kullandığımız verisetine göre özgüllük, duyarlılık ve F-ölçütü değerleri üretmektedir. Kullandığımız verisetine göre bu değerler göz önünde bulundurularak farklı algoritmalar arasından en iyi uyum sağlayan algoritma tercih edilerek sonuç iyileştirilir ve zaman maliyeti azaltılır.

#### 4. Sonuç

İş tatmini faktörlerini belirlemeye ve analiz etmeye yönelik olarak çalışanların çevrimiçi değerlendirmelerinin makine öğrenmesine dayalı analizini elde etmek amacıyla gerçekleştirilen çalışmalar hakkında bilgi vermektedir. Çalışanların şirketleri hakkında yaptıkları yorumlardan 1 ve 2 puan olanlar negatif, 3,4 ve 5 puanlar pozitif olarak değerlendirilmiş olup toplamda 18 firmadaki 2321 pozitif yorum ve 702 negatif yorum ele alınarak analizleri gerçekleştirilmiştir.

Metin belgesindeki özniteliklerin belirlenmesinde terim varlığı, terim sıklığı gibi temel temsil yöntemleri ve 1-gram, 2-gram, 3-gram modelleri dikkate alınarak incelenen veri seti için farklı temsil yöntemlerinin doğru sınıflandırma başarımları, F-ölçütü, duyarlılık, hassasiyet ölçütleri 5 temel sınıflandırıcı aracılığıyla karşılaştırmalı olarak değerlendirilmiştir.

Yapılan bu analizler sonucunda en yüksek başarı oranı Tablo 3'de görüleceği üzere 83.6586% başarı oranı ile meta.Stacking algoritmasında TF\_unigram.arff komfigürasyonu ile elde edilmiş olup deney sonucunun alınması 541.96 saniye sürmüştür. Bu deneyden elde edilen duyarlılık değeri 0,826; hassasiyet değeri 0,837 ve F-ölçütü değeri, 0,823 olup Terim sıklığı (Term

Frequency) konfigürasyonu ve 1-gram modeli ile en yüksek başarıyı vermiştir. Analizler sonucunda en düşük başarı oranı ise Tablo 4'te görülen 66.9864% başarı oranı ile meta.Bagging - functions.Logistic algoritmasından elde edilerek deneyin yapılması 204.44 saniye sürmüştür. Analiz sonucu elde edilen bu en düşük başarı oranı TP\_unigram.arff konfigürasyonu ile yani terim varlığı (Term Present) ve 1-gram modeli kullanılarak elde edilmiş olup sırasıyla duyarlılık;hassasiyet ve F-ölçütü değerleri 0,698; 0,67 ve 0,682 olarak bulunmuştur.

#### Kaynakça

- [1] Akşit Aşık, Nuran (2010), “Çalışanların İş Doyumunu Etkileyen Bireysel Ve Örgütsel Faktörler İle Sonuçlarına İlişkin Kavramsal Değerlendirme”, Türk İdare Dergisi, 467(Haziran), 31-51.
- [2] Luthans, F., Baack, D. ve Taylor, L. (1992). *Organizational Commitment: Analysis of Antecedents*, Human Relations.
- [3] S. Strohmeier, F. Piazza, Domain driven data mining in human resource management: a review of current research, *Expert Systems with Applications* 40 (7) (2013) 2410–2420.
- [4] Yawalkar, V. V., (2019), “Study of Artificial Intelligence and its role in Human Resource Management”, *International Journal of Research and Analytical Reviews (IJRAR)*, pp.20-24).
- [5] Sebastiani, Fabrizio. "Machine learning in automated text categorization." *ACM computing surveys (CSUR)* 34.1 (2002): 1-47.
- [6] U. Özkaya, L. Seyfi. (2021), “Yere Nüfuz Eden Radar B Tarama Görüntülerinin Az Parametreye Sahip Konvolüsyonel Sinir Ağı İle Değerlendirilmesi” *Geomatik*, 6(2), 84-92.