

MAKİNE ÖĞRENİMİ VE DERİN ÖĞRENME İLE DUYGU ANALİZİ

HAKKIMIZDA



ALEYNA KAHRAMAN
20060355



DİLARA BIYIKLI
20060334



FEYZANUR AYTEKİN
20060353



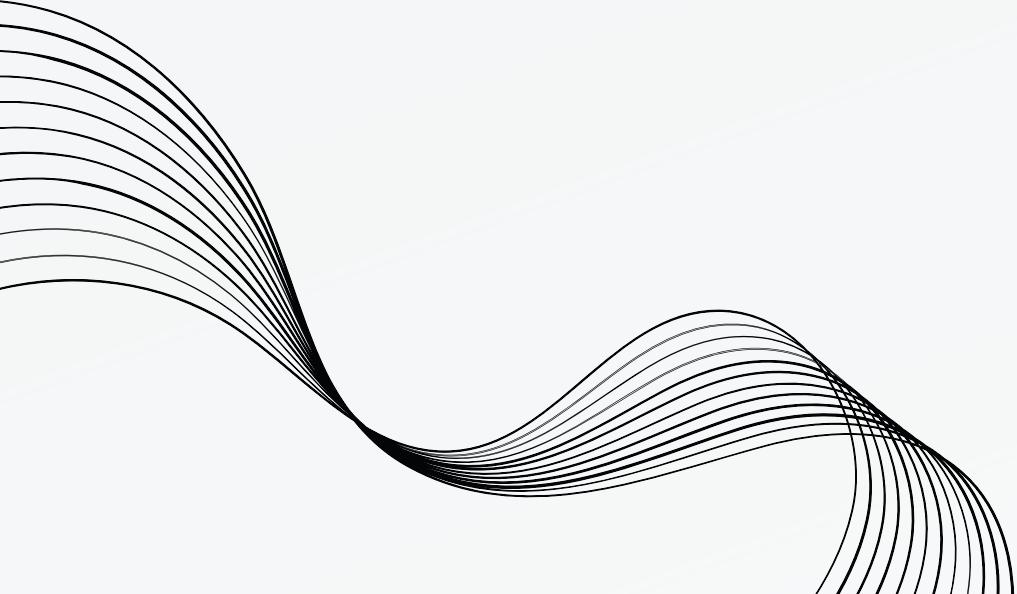
HASAN BASRİ DARGA
20060378

içerik

04	DUYGU ANALİZİ	09	METİN TEMSİL YÖNTEMLERİ
05	AMACIMIZ	10	ALGORİTMALAR
06	VERİSETLERİ	11	PERFORMANS METRİKLERİ
07	KÜTÜPHANELER	12	SONUÇLAR VE KARŞILAŞTIRMA
08	VERİ ÖN İŞLEME	17	KAZANIMLAR

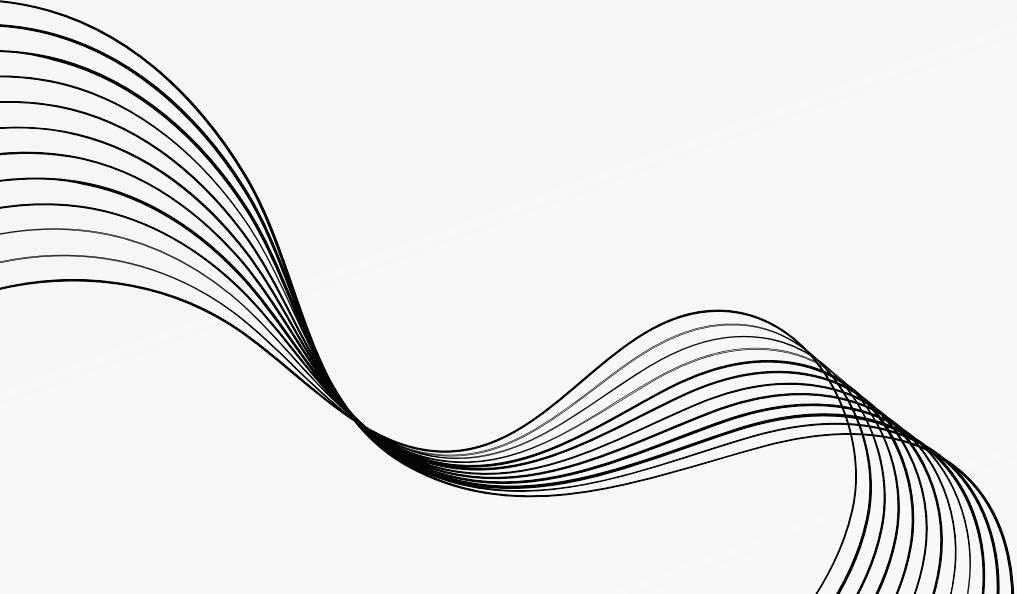
DUYGU ANALİZİ

- Sosyal Medya İzleme ve Yorum Analizi
- Ürün İncelemeleri ve Pazar Araştırmaları
- Müşteri Hizmetleri ve Memnuniyet Analizi
- Finansal Piyasa Tahmini
- Politika ve Kamuoyu Araştırmaları
- Sanat ve Kültürde Uygulamalar

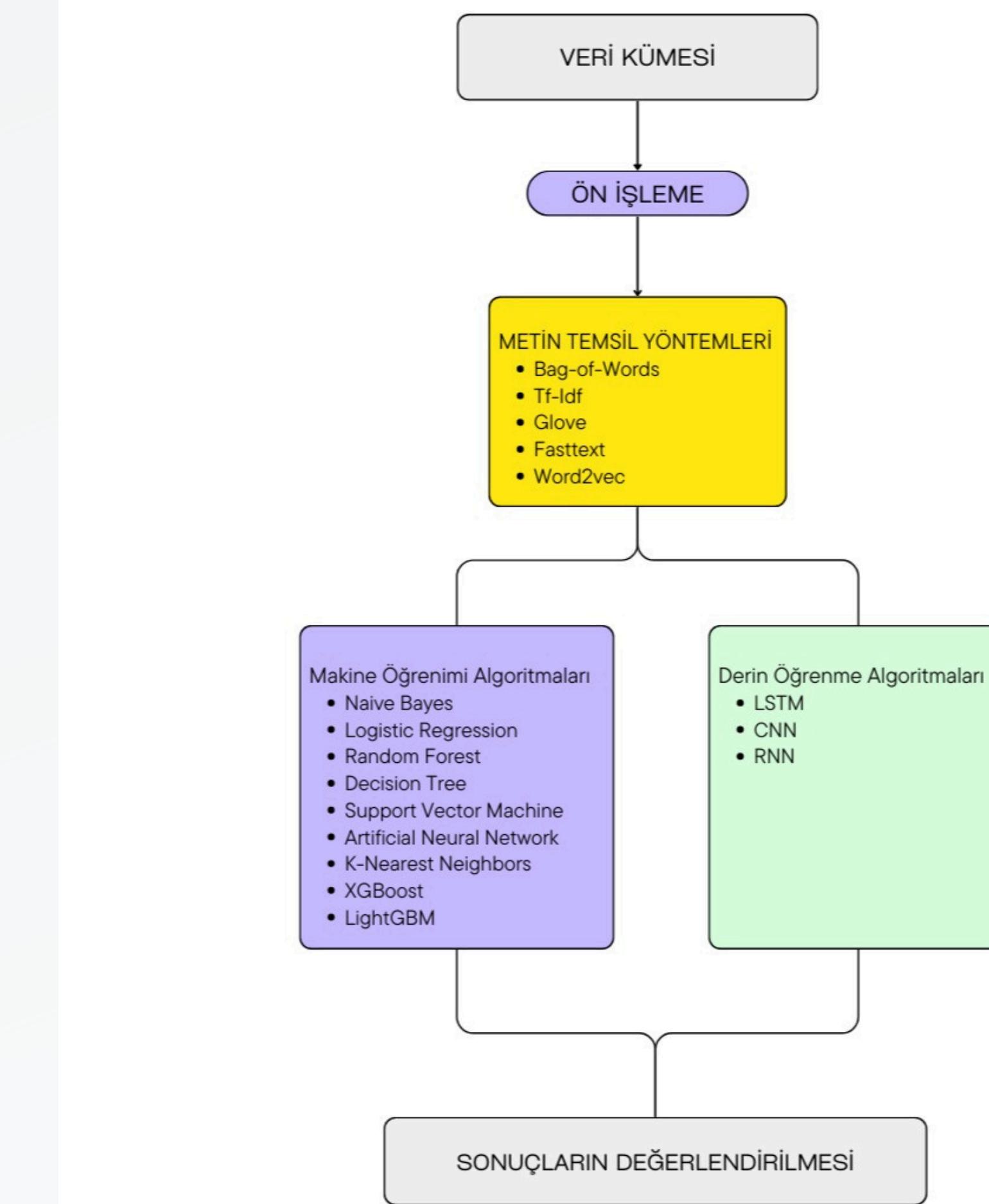


AMACIMIZ

- Geniş Kapsamlı Literatür Araştırması
- Farklı Veri Setlerinin Kullanılması
- İki ve Üç Sınıflı Sınıflandırma
- Çeşitli Metin Temsil Yöntemlerinin Kullanılması
- Çeşitli Makine Öğrenimi ve Derin Öğrenme Algoritmalarının Kullanılması



ÇALIŞMANIN YAPISI



VERİSETLERİ



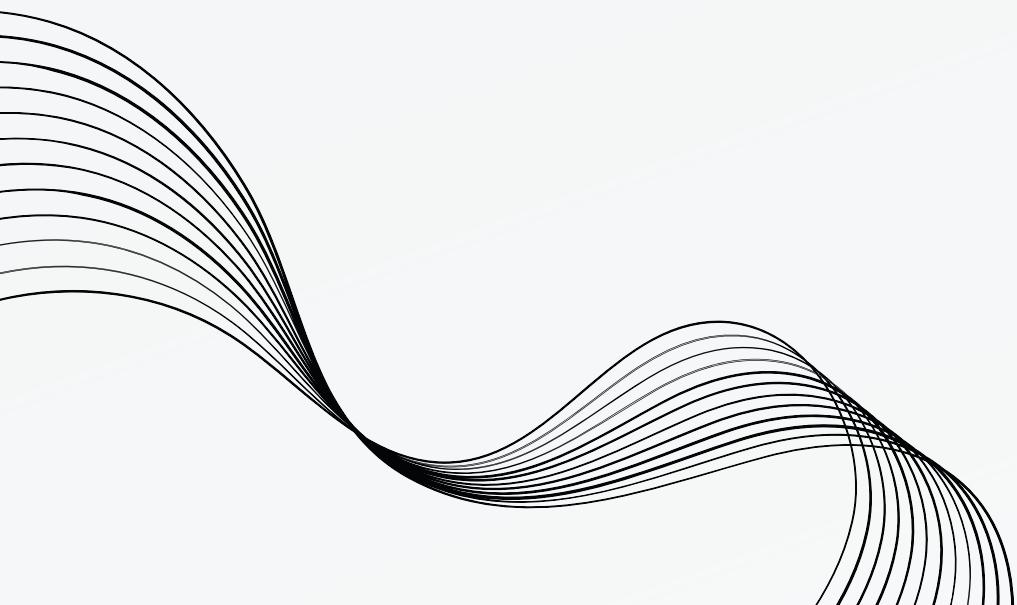
- Amazon veri seti olumlu, olumsuz ve nötr veriler içeren 10.000 veriden oluşmaktadır.

- IMDb veri seti olumlu ve olumsuz veriler içeren 20.000 veriden oluşmaktadır.

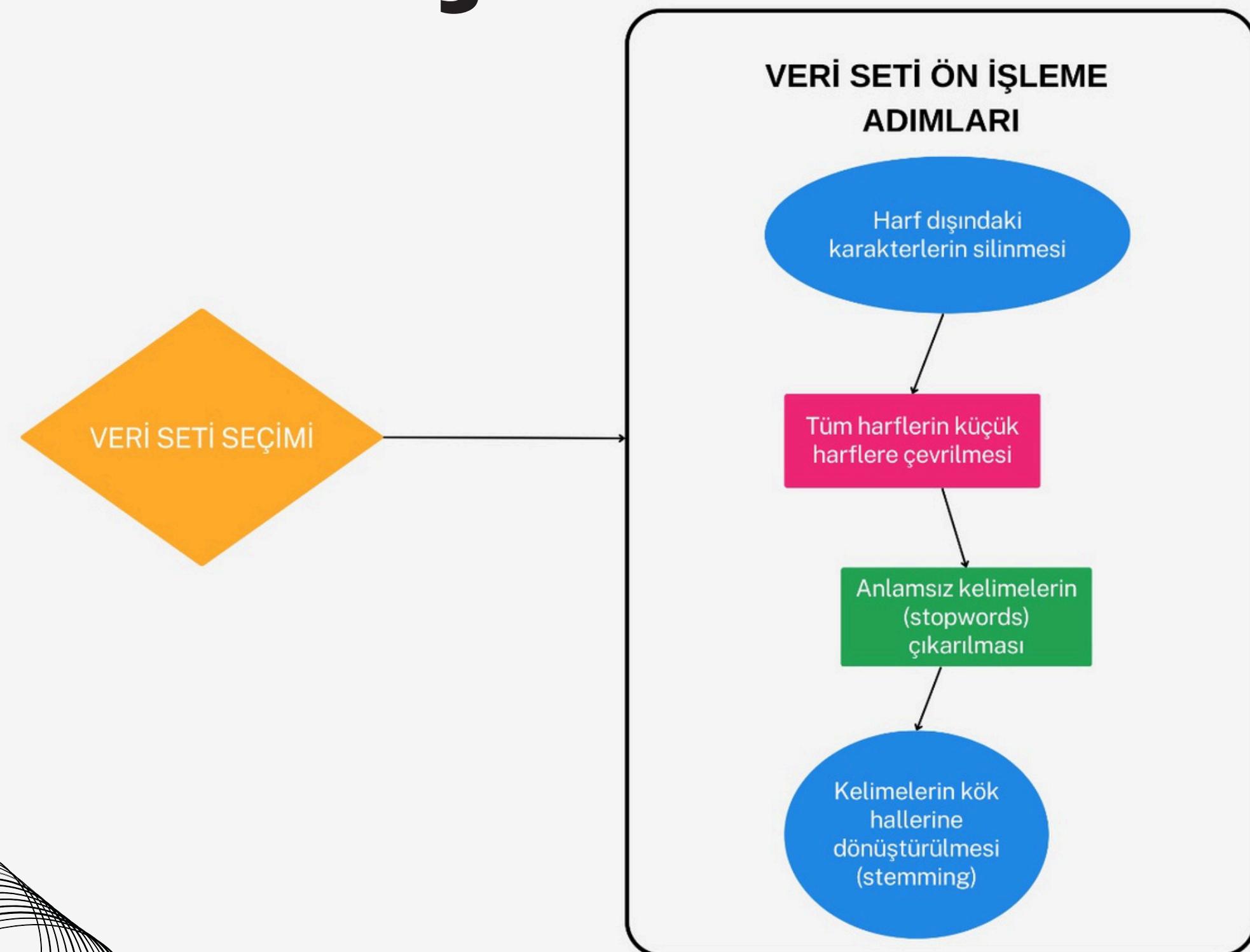


KÜTÜPHANELER

- Numpy
- Pandas
- Scikit-Learn
- Tensorflow
- Gensim

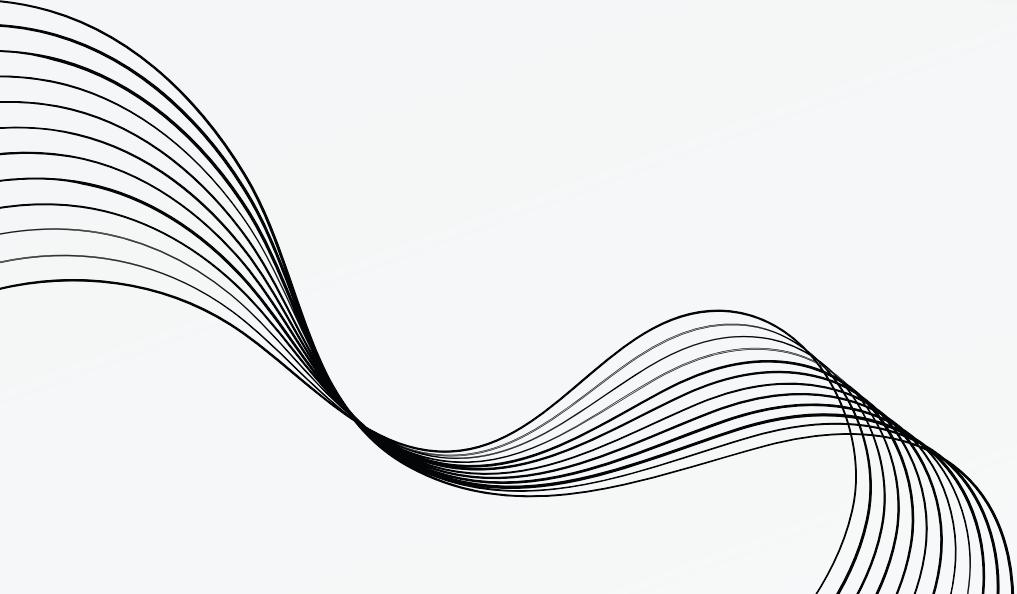


VERİ ÖN İŞLEME



METİN TEMSİL YÖNTEMLERİ

- Bag-of-Words
- Tf-Idf
- Glove
- Fasttext
- Word2vec



ALGORİTMALAR

Derin Öğrenme

LSTM

CNN

RNN

Makine Öğrenimi

Naive Bayes

Logistic Regression

Random Forest

Decision Tree

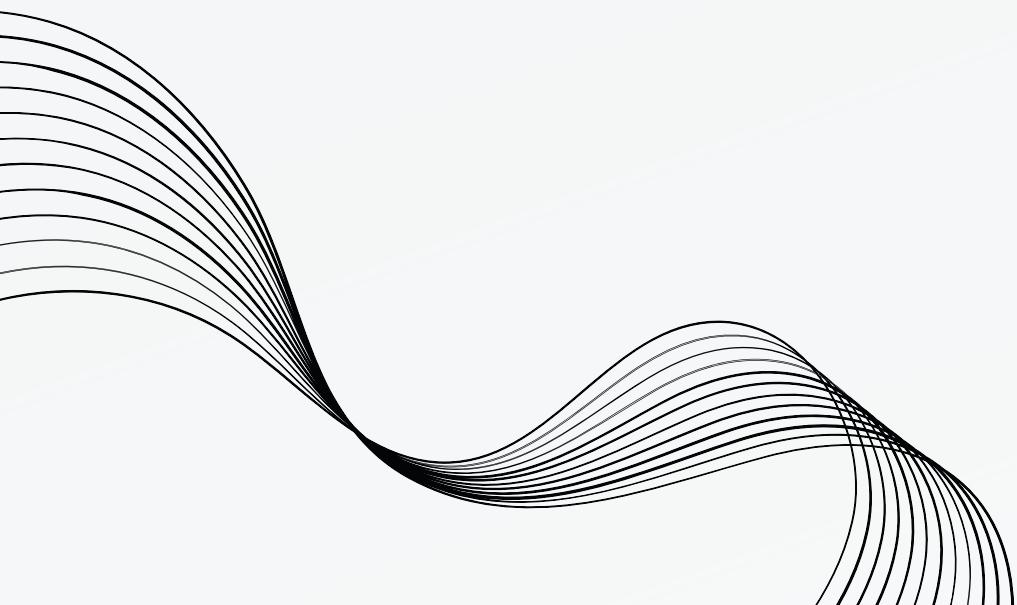
Support Vector Machine

Artificial Neural Network

K-Nearest Neighbors

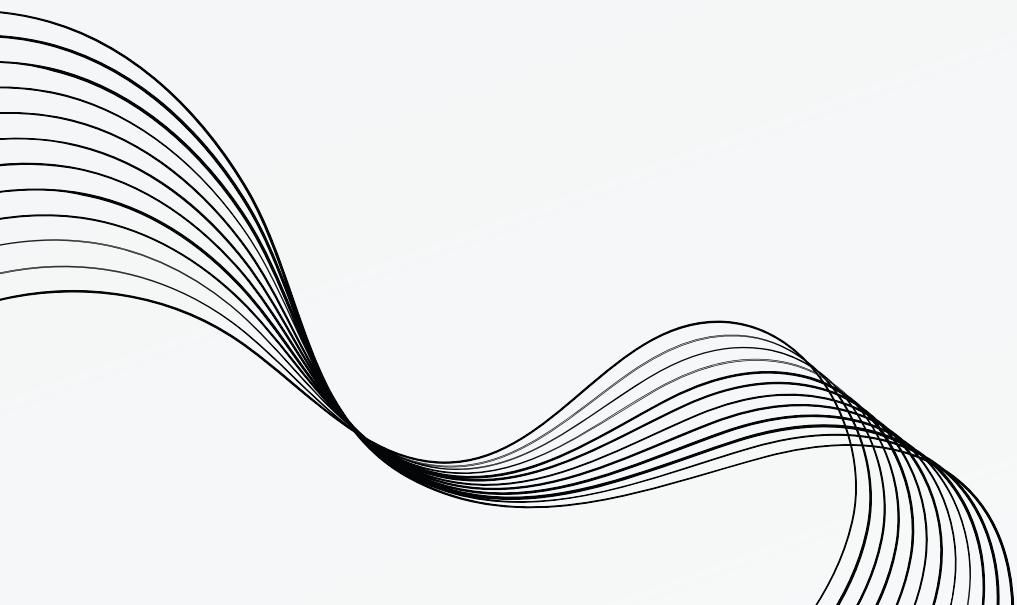
XGBoost

LightGBM



PERFORMANS METRİKLERİ

- Doğruluk
- Duyarlılık
- Hassasiyet
- F1-Skor



AMAZON VERİ SETİ DERİN ÖĞRENME SONUÇLARI

Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
TF-IDF	CNN	88.29	0.90	0.88	0.83
	RNN	88.00	0.83	0.88	0.83
	LSTM	88.29	0.90	0.88	0.83
WORD2VEC	CNN	88.10	0.79	0.88	0.83
	RNN	88.15	0.78	0.88	0.83
	LSTM	88.00	0.81	0.88	0.84
GLOVE	CNN	88.29	0.90	0.88	0.83
	RNN	88.29	0.90	0.88	0.83
	LSTM	88.24	0.80	0.78	0.83
FASTTEXT	CNN	88.29	0.90	0.88	0.83
	RNN	88.29	0.87	0.88	0.83
	LSTM	88.30	0.90	0.88	0.83
BINARY	CNN	88.05	0.90	0.88	0.83
	RNN	88.29	0.90	0.88	0.83
	LSTM	88.29	0.78	0.88	0.83

AMAZON SETİ MAKİNE ÖĞRENME SONUÇLARI

Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
TF-IDF	BNB	84.05	0.79	0.84	0.81
	LR	88.34	0.82	0.88	0.83
	RF	88.24	0.78	0.88	0.83
	DT	83.32	0.82	0.83	0.83
	SVM	88.29	0.84	0.88	0.83
	ANN	87.03	0.84	0.87	0.85
	KNN	87.51	0.79	0.88	0.83
	XGBOOST	88.20	0.83	0.88	0.84
	LightGBM	88.29	0.84	0.88	0.84
WORD2VEC	BNB	88.10	0.81	0.88	0.83
	LR	88.05	0.81	0.88	0.83
	RF	88.34	0.87	0.88	0.83
	DT	80.30	0.82	0.80	0.81
	SVM	88.29	0.84	0.88	0.83
	ANN	83.71	0.83	0.84	0.83
	KNN	86.44	0.81	0.86	0.83
	XGBOOST	88.59	0.84	0.89	0.84
	LightGBM	88.39	0.83	0.88	0.84
GLOVE	BNB	84.73	0.79	0.85	0.82
	LR	87.95	0.80	0.88	0.83
	RF	88.20	0.83	0.88	0.83
	DT	77.96	0.80	0.78	0.79
	SVM	88.29	0.90	0.88	0.83
	ANN	87.56	0.81	0.88	0.83
	KNN	87.03	0.80	0.87	0.83
	XGBOOST	88.10	0.83	0.88	0.83
	LightGBM	88.20	0.83	0.88	0.83

Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
FASTTEXT	BNB	82.35	0.80	0.82	0.81
	LR	88.05	0.82	0.88	0.83
	RF	88.29	0.85	0.88	0.83
	DT	77.86	0.80	0.78	0.79
	SVM	88.29	0.90	0.88	0.83
	ANN	87.46	0.81	0.87	0.83
	KNN	86.15	0.80	0.86	0.83
	XGBOOST	88.20	0.80	0.88	0.83
	LightGBM	88.10	0.78	0.88	0.83
BINARY	BNB	83.95	0.79	0.84	0.81
	LR	86.78	0.84	0.87	0.85
	RF	88.29	0.78	0.88	0.83
	DT	83.08	0.82	0.84	0.83
	SVM	88.29	0.78	0.88	0.83
	ANN	86.25	0.83	0.86	0.84
	KNN	88.29	0.85	0.88	0.83
	XGBOOST	88.34	0.83	0.88	0.84
	LightGBM	88.34	0.84	0.88	0.85

IMDB VERİ SETİ DERİN ÖĞRENME SONUÇLARI

Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
TF-IDF	CNN	87.24	0.88	0.87	0.87
	RNN	51.80	0.75	0.49	0.32
	LSTM	51.20	0.75	0.51	0.35
WORD2VEC	CNN	84.77	0.85	0.85	0.85
	RNN	76.67	0.80	0.77	0.76
	LSTM	85.37	0.85	0.85	0.85
GLOVE	CNN	75.02	0.77	0.75	0.75
	RNN	69.57	0.71	0.70	0.69
	LSTM	80.54	0.82	0.81	0.80
FASTTEXT	CNN	85.07	0.86	0.85	0.85
	RNN	73.65	0.79	0.74	0.73
	LSTM	87.84	0.88	0.88	0.88
BINARY	CNN	78.50	0.79	0.79	0.78
	RNN	48.00	0.49	0.48	0.33
	LSTM	52.00	0.75	0.52	0.36

IMDB SETİ MAKİNE ÖĞRENME SONUÇLARI

Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
TF-IDF	BNB	85.65	0.86	0.86	0.86
	LR	88.37	0.88	0.88	0.88
	RF	84.05	0.84	0.84	0.84
	DT	70.17	0.70	0.70	0.70
	SVM	88.85	0.89	0.89	0.89
	ANN	87.45	0.87	0.87	0.87
	KNN	74.52	0.75	0.75	0.74
	XGBOOST	83.52	0.84	0.84	0.84
	LightGBM	84.70	0.85	0.85	0.85
WORD2VEC	BNB	51.25	0.59	0.51	0.35
	LR	86.65	0.87	0.87	0.87
	RF	83.02	0.83	0.83	0.83
	DT	71.97	0.72	0.72	0.72
	SVM	86.87	0.87	0.87	0.87
	ANN	84.47	0.85	0.85	0.85
	KNN	80.55	0.81	0.81	0.81
	XGBOOST	85.20	0.85	0.85	0.85
	LightGBM	84.77	0.85	0.85	0.85
GLOVE	BNB	67.00	0.67	0.67	0.67
	LR	75.75	0.76	0.76	0.76
	RF	73.15	0.73	0.73	0.73
	DT	61.65	0.62	0.62	0.62
	SVM	75.05	0.75	0.75	0.75
	ANN	72.60	0.73	0.73	0.73
	KNN	66.12	0.66	0.66	0.66
	XGBOOST	74.02	0.74	0.74	0.74
	LightGBM	73.77	0.74	0.74	0.74

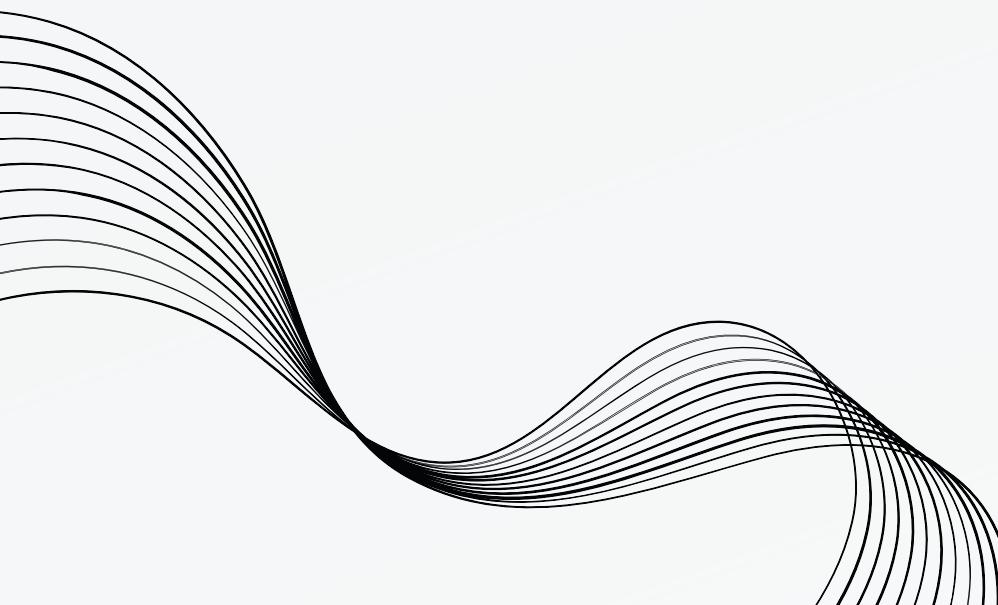
Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)	Precision	Recall	F1-Score
FASTTEXT	BNB	71.75	0.72	0.72	0.72
	LR	85.07	0.85	0.85	0.85
	RF	79.85	0.80	0.80	0.80
	DT	69.35	0.69	0.69	0.69
	SVM	85.22	0.85	0.85	0.85
	ANN	81.70	0.82	0.82	0.82
	KNN	77.05	0.77	0.77	0.77
	XGBOOST	82.82	0.83	0.83	0.83
	LightGBM	82.67	0.83	0.83	0.83
BINARY	BNB	85.72	0.86	0.86	0.86
	LR	87.12	0.87	0.87	0.87
	RF	84.22	0.84	0.84	0.84
	DT	70.70	0.71	0.71	0.71
	SVM	88.07	0.88	0.88	0.88
	ANN	86.15	0.86	0.86	0.86
	KNN	59.87	0.62	0.60	0.57
	XGBOOST	84.10	0.84	0.84	0.84
	LightGBM	85.07	0.85	0.85	0.85

KARŞILAŞTIRMA

ÇALIŞMA	Veri Kümesi	Metin Temsil Yöntemi	Sınıflandırma Algoritması	Accuracy(%)
Poornima ve Priya	Twitter	TF-IDF	LR	86.23
Pandya vd.	Twitter	WORD2VEC	LSTM	81.33
Rawat vd.	Twitter	TF-IDF	SVC	97
Joshi vd.	Twitter	TF-IDF	DT	75.36
Jagaeedesan vd.	Twitter	TF-IDF BAG-OF-WORDS	RF SVM	95 95
ÇALIŞMAMIZ	Amazon	WORD2VEC	XGBOOST	88.59
ÇALIŞMAMIZ	IMDb	TF-IDF	SVM	88.85

KAZANIMLAR

- Güncel Araştırmaları ve Literatürü İnceleme Becerisinin Kazanılması
- Veri Ön İşleme ve Hazırlama Becerilerinin Kazanılması
- Makine Öğrenimi ve Derin Öğrenme Algoritmaları Hakkında Bilgi Edinmek
- Analitik Düşünme ve Raporlama Becerilerin Kazanılması
- Programlama ve Teknik Becerilerin Gelişmesi
- Github Kullanımı



TEŞEKKÜRLER

