**Doğal Dil İşleme Final Projesi Raporu**

## Giriş

Bu rapor, doğal dil işleme final projesinin detaylarını içermektedir. Proje, vize ödevinde kullanılan 10.000 cümlelik önişlenmiş derlemin üzerinde Word2Vec ve TF-IDF gibi geleneksel yöntemlerle benzerlik analizi yapmayı hedeflemektedir.

## 1. Veri Hazırlığı

Vize ödevinde kullanılan derlem, önceki ödevde önişleme adımlarından geçirilmiştir. Önişlenmiş derlemin içinde geçen kelimeler ve cümleler üzerinde analiz yapmak için kullanılmıştır.

## 2. En Fazla Geçen 20 Kelime

Hazırlanan önişlenmiş derlemin içinde en fazla geçen 20 kelime belirlenmiştir. Bu kelimelerin sayısıyla birlikte aşağıdaki gibidir:

('movi', 29396)

('spider', 19745)

('man', 18710)

('film', 17349)

('peter', 12835)

('like', 11412)

('one', 10441)

('charact', 9054)

('good', 8734)

('spiderman', 8369)

('comic', 7558)

('great', 7411)

('stori', 7258)

('parker', 7109)

('scene', 6847)

('time', 6753)

('realli', 6504)

('first', 6463)

('see', 6424)

('make', 6062)

## 3. En Sık Geçen 20 Kelimenin Benzerleri

Her biri en sık geçen 20 kelimenin her biri için en benzer 5 farklı kelime tespit edilmiştir. Bu benzerlikler aşağıda listelenmiştir:

Benzer kelimeler for 'movi':

marvel 0.9976573586463928

superhero 0.9972503781318665

film 0.9950191974639893

year 0.9949743747711182

mcu 0.9936113357543945

Benzer kelimeler for 'spider':

noir 0.9913691282272339

multipl 0.9912149906158447

endear 0.9863166213035583

spicer 0.9843331575393677

hey 0.9842255711555481

Benzer kelimeler for 'man':

noir 0.9959309697151184

multipl 0.9881957769393921

spicer 0.98763507604599

robot 0.9866400361061096

ham 0.9865864515304565

Benzer kelimeler for 'film':

worthi 0.9983352422714233

one 0.9980840086936951

definit 0.9976164102554321

sick 0.997494637966156

date 0.9973233938217163

Benzer kelimeler for 'peter':

parker 0.9972549676895142

b 0.9965397715568542

york 0.9930593371391296

popular 0.9928998947143555

dimens 0.9919701814651489

Benzer kelimeler for 'like':

art 0.9986287355422974

uniqu 0.9985082149505615

captur 0.9983929991722107

watch 0.9983537793159485

revolutionari 0.9983094334602356

Benzer kelimeler for 'one':

possibl 0.9990041255950928

might 0.9989032745361328

definit 0.9984905123710632

consid 0.9983500242233276

worthi 0.9983270168304443

Benzer kelimeler for 'charact':

fit 0.9984706044197083

teen 0.9982566833496094

suddenli 0.9980126619338989

chang 0.9980049133300781

young 0.997838020324707

Benzer kelimeler for 'good':

music 0.9988294839859009

set 0.9988190531730652

perform 0.9988093376159668

talent 0.9987026453018188

constant 0.9985908269882202

Benzer kelimeler for 'spiderman':

digit 0.999093234539032

win 0.9990524053573608

post 0.9989985823631287

cartoon 0.998995840549469

fantast 0.9989145398139954

Benzer kelimeler for 'comic':

explos 0.9901590943336487

hardcor 0.983702540397644

tweak 0.9748416543006897

read 0.9747305512428284

ya 0.9741829037666321

Benzer kelimeler for 'great':

come 0.9976795315742493

perfect 0.9973742961883545

age 0.997233510017395

move 0.9970101118087769

like 0.9969801306724548

Benzer kelimeler for 'stori':

abl 0.9982413649559021

similar 0.9979584813117981

tell 0.9979427456855774

boy 0.9972786903381348

arc 0.9970746636390686

Benzer kelimeler for 'parker':

peter 0.9972550272941589

b 0.9944245219230652

charismat 0.9912707209587097

jane 0.9909078478813171

dimens 0.9892410039901733

Benzer kelimeler for 'scene':

bit 0.9994513988494873

flick 0.9992252588272095

current 0.9991979598999023

rate 0.9991514682769775

surpris 0.9991271495819092

Benzer kelimeler for 'time':

review 0.9990223050117493

imdb 0.9989420771598816

deserv 0.9988682866096497

amaz 0.998721182346344

perfect 0.9987172484397888

Benzer kelimeler for 'realli':

surpris 0.998808741569519

moment 0.9987602829933167

atmospher 0.998736560344696

simpli 0.9985832571983337

flick 0.9985635876655579

Benzer kelimeler for 'first':

exampl 0.9989839196205139

hope 0.9988575577735901

make 0.9988414645195007

would 0.9988242983818054

storytel 0.998748242855072

Benzer kelimeler for 'see':

say 0.9991855621337891

go 0.999081015586853

differ 0.9986361861228943

releas 0.998613178730011

someon 0.9984040260314941

Benzer kelimeler for 'make':

golden 0.9992586374282837

storytel 0.9991883039474487

game 0.9991649389266968

cours 0.9991142749786377

alway 0.9990787506103516

## 4. En Benzer Cümle Analizi

Proje kapsamında önişlenmiş derlemin içinden 5 farklı cümle seçilmiştir. Her bir seçilen cümle için, vize ödevinde verilen sıra numarasıyla en benzer 3 farklı cümle tespit edilmiştir. Bu benzer cümleler aşağıdaki gibidir:

Seçili Cümle: film phenomen mani reason first importantli visual impress film seen day anim style alon mystic wonder qualiti anim stop howev ad mani littl fun detail made feel like watch sort blend comic movi addit gorgeou visual natur spider man impress humor meta refer refer cringey spider man scene post credit meme refer person favorit ton littl detail sure never truli appreci never consum comic book imagin film even special marvel comic fan gripe spider man fail keep attent act longer would benefit greatli cut minut said best movi seen month hope review reach someon go see urg reconsid

9402: like movi much one best marvel movi movi person like action adventur fantanc special effect movi nice remak older version spider man chang wonder graphic act actor movi great stori also nice set also nice movi wast time money enjoy movi much great movi marvel also famili type film movi seen whole famili tri watch movi everi thing movi perfect action stori like anim movi - Benzerlik Skoru: 0.999925971031189

5444: probabl worst movi ever seen saw greatest movi ever must seen total differ film tobey maguir annoy first make want bash face shovel whole movi kirsten dunst ok best play normal superhero love interest alfr molina act unmemor best two good perform film j k simmon j jonah jameson willem dafo norman osbourn film combin second movi remind soap opera rather comic book focu place annoy peter parker pin mari jane instead real reason go see move spider man go comic book movi expect see someth par godfath want entertain movi fall way short mark - Benzerlik Skoru: 0.9999006986618042

813: love enjoy watch movi daughter great film - Benzerlik Skoru: 0.9998863339424133

Seçili Cümle: cannot stress enough good film critic actual got thing right visual stun great storylin featur comedi tragedi

En Benzer 3 Cümle:

333: well absolut brilliant go expect much boy exceed beyond watch long think concept would work anim first admit stupidli wrong visual film stun music done use perfect overal everyth film fantast genuin kick watch sooner write review tend focu neg complain good film honestli anyth bad list good thing day b liter everyth pleas watch - Benzerlik Skoru: 0.9999052882194519

1793: went averag expect truli blown away good movi filler scene everyth flow made perfect sens except pace saw art magnific look forward work everyon see film - Benzerlik Skoru: 0.9998571276664734

1660: honest put one day tv poor amaz film well worth watch watch till credit - Benzerlik Skoru: 0.9998531937599182

Seçili Cümle: lush astound stylis artwork meticul blend screen print style graphic edgi colour super funni sweet w almost hammi deadpool type humour fantast movi realli enjoy visual treat sure

En Benzer 3 Cümle:

963: lush astound stylis artwork meticul blend screen print style graphic edgi colour super funni sweet w almost hammi deadpool type humour fantast movi realli enjoy visual treat sure - Benzerlik Skoru: 1.0000001192092896

1722: visual audio masterpiec everi singl scene piec liter artwork soundtrack rival anyth find opinion nice stori lot heart humor enjoy hour sure - Benzerlik Skoru: 0.9999434947967529

1294: everi movi come truli make impact joaquin perform scenographi brillianc grotesqu haunt cringi hard watch time mesmer blink eye watch tragic serious funni moment emot rollercoast sometim multipl emot pop time - Benzerlik Skoru: 0.9999223351478577

1155: went see thought children would like forc sit tortur attempt humour afterward found kid like either laugh bit overal bore daughter word watch alreadi humour overli forc point like slap face action film bit rush describ clumsi stori okay bit rush jump one thing anoth one minut noth happen next minut everyth happen bare transit main thing like style anim use complet honesti suck massiv overal bitter disappoint recommend anybodi wast time money want watch wait till free - Benzerlik Skoru: 0.9999170899391174

## 5. Sonuç

Proje kapsamında gerçekleştirilen analizler sonucunda, önişlenmiş derlemin içindeki kelimeler ve cümleler üzerinde başarılı bir benzerlik analizi yapılmıştır. Elde edilen sonuçlar raporun bu bölümünde detaylı olarak açıklanmıştır.