

pba-modul2-id-text-classification

October 6, 2023

1 Modul Praktikum Pengolahan Bahasa Alami

1.1 Text Classification in Indonesian Language

Y. Sigit Purnomo W.P.

[Website](#)

Notebook ini merupakan materi untuk modul praktikum mata kuliah Pengolahan Bahasa Alami di Program Studi Informatika, Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta

Kebutuhan yang harus dipenuhi terkait modul praktikum ini adalah: 1. Python 2. NLTK 3. Scikit-learn 4. Gensim 5. Sastrawi

1.1.1 Inisialisasi

Langkah 1: Inisialisasi Library Langkah pertama yang harus dilakukan adalah mengimport library-library yang akan digunakan.

```
[1]: import numpy as np
import pandas as pd
import modSpellChecker as sc
from contractions import CONTRACTION_MAP
import re
import nltk
import string
from nltk.stem import WordNetLemmatizer
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
import gensim
from gensim.models import Word2Vec
from gensim import models
#from pattern.en import tag
from nltk.corpus import wordnet as wn
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
from gensim import corpora, models
import sklearn
#from normalization import normalize_corpus
```

```

from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.datasets import fetch_20newsgroups
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.naive_bayes import MultinomialNB

```

Langkah 2: Data Preparation Langkah kedua yang harus dilakukan adalah menyiapkan data yang akan digunakan untuk keperluan pengembangan model klasifikasi teks.

```

[2]: dataset = pd.read_csv('dataartikel.csv')
dataset

```

```

[2]:

```

		Judul \
0		Rupiah masih rawan bergejolak
1	SRC	sasar ritel tradisional jadi pilar ekonomi...
2		Kiat memaksimalkan Facebook untuk bisnis
3		Kiat menambah relasi bisnis
4		Prinsip dasar merintis bisnis
...		...
1304	Genoa	1 - 1 Sampdoria - Liga Serie A 2018/2019
1305	Villarreal	2 - 1 Real Betis - La Liga 2018/2019
1306	PSM	kembali ke puncak klasemen, Persib resmi t...
1307		Hasil imbang yang pahit bagi AC Milan
1308	Seruan	untuk Edy Rahmayadi di balik kegagalan ...

	Artikel	Jenis
0	Nilai tukar rupiah terhadap dolar Amerika Seri...	ekonomi
1	Perkembangan teknologi telah menyebabkan perub...	ekonomi
2	Media sosial kini bukan hanya menjadi alat unt...	ekonomi
3	Dalam menjalankan bisnis, menjalin relasi puny...	ekonomi
4	Merintis usaha memang bukan pekerjaan mudah. A...	ekonomi
...
1304	Genoa dan Sampdoria sama-sama gagal memetik po...	sport
1305	Villarreal menjaga langkah untuk mendekatkan d...	sport
1306	Persija Jakarta harus rela kembali ke peringka...	sport
1307	"Hasil imbang yang pahit." Begitu laman resmi ...	sport
1308	Timnas Indonesia mengakhiri babak grup Piala A...	sport

[1309 rows x 3 columns]

```

[3]: dataset.shape

```

```

[3]: (1309, 3)

```

```
[4]: df2 = dataset.groupby(['Jenis'])['Artikel'].count()
df2
```

```
[4]: Jenis
ekonomi          100
entertainment    134
gaya hidup       200
news             266
sains-tekno      202
sport            407
Name: Artikel, dtype: int64
```

```
[5]: label = dataset.iloc[:,2]
feature = dataset.iloc[:,1]
```

```
[6]: label[0:10]
```

```
[6]: 0    ekonomi
1    ekonomi
2    ekonomi
3    ekonomi
4    ekonomi
5    ekonomi
6    ekonomi
7    ekonomi
8    ekonomi
9    ekonomi
Name: Jenis, dtype: object
```

```
[7]: feature[0:10]
```

```
[7]: 0    Nilai tukar rupiah terhadap dolar Amerika Seri...
1    Perkembangan teknologi telah menyebabkan perub...
2    Media sosial kini bukan hanya menjadi alat unt...
3    Dalam menjalankan bisnis, menjalin relasi puny...
4    Merintis usaha memang bukan pekerjaan mudah. A...
5    Perdagangan di Bursa Efek Indonesia (BEI) ditu...
6    Indeks Harga Saham Gabungan (IHSG) terkoreksi ...
7    Proses ratifikasi tujuh perjanjian perdagangan...
8    Pemerintah mematangkan skema pembiayaan fasilt...
9    Anak muda zaman sekarang boleh jadi tidak asin...
Name: Artikel, dtype: object
```

Langkah 3: Normalisasi Data Langkah ketiga yang harus dilakukan adalah melakukan normalisasi data yang akan digunakan untuk keperluan pengembangan model klasifikasi teks.


```

def stemmer_text(text):
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    text = stemmer.stem(text)
    return text

# def lemmatize_text(text):
#     pos_tagged_text = pos_tag_text(text)
#     lemmatized_tokens = [wnl.lemmatize(word, pos_tag) if pos_tag
#                           else word for word,
#                           pos_tag in pos_tagged_text]
#     lemmatize_text=' '.join(lemmatized_tokens)
#     return lemmatize_text

```

```

[9]: def remove_special_characters(text):
    tokens = tokenize_text(text)
    pattern = re.compile('[{}]'.format(re.escape(string.punctuation)))
    filtered_tokens = filter(None, [pattern.sub('', token) for token in tokens])
    filtered_text = ' '.join(filtered_tokens)
    return filtered_text

factory = StopWordRemoverFactory()
stopword_list = factory.get_stop_words()

def remove_stopwords(text):
    tokens = tokenize_text(text)
    filtered_tokens = [token for token in tokens if token not in stopword_list]
    filtered_text = ' '.join(filtered_tokens)
    return filtered_text

def normalize_corpus(corpus, tokenize=False):
    normalized_corpus = []
    for text in corpus:
        text = expand_contractions(text, CONTRACTION_MAP)
        text = stemmer_text(text)
        text = remove_special_characters(text)
        text = repeatcharNormalize(text)
        text = remove_stopwords(text)

        normalized_corpus.append(text)
        if tokenize:
            text = tokenize_text(text)
            text = spellNormalize(text)
            normalized_corpus.append(text)
    return normalized_corpus

```

```
[10]: #training dataset
def prepare_datasets(corpus, labels, test_data_proportion=0.3):
    train_X, test_X, train_Y, test_Y = train_test_split(corpus, labels,
                                                         test_size=0.33,
                                                         random_state=42)
    return train_X, test_X, train_Y, test_Y

train_corpus, test_corpus, train_labels, test_labels = prepare_datasets(feature,
                                                                           label,
                                                                           test_data_proportion=0.3)
```

```
[11]: #memanggil fungsi normalisasi dataset yang telah ditraining
norm_train_corpus = normalize_corpus(train_corpus)
```

```
[12]: norm_train_corpus[0:10]
```

```
[12]: ['gonjanganjng soal siapa latih tim nasional indonesia depan jawab lalu rapat
komite eksekutif exco pssi senin 17 9 2018 malam federasi nama luis milla besut
skuat garuda putus sekaligus jawab spekulasi edar beberapa hari akhir kontrak
milla tak panjang pssi spekulasi sebut maklum kontrak akhir pascaasian games
milla kerap gagal penuh target misal ajang sea games 2017 latih usia 51 tahun
target raih emas nyata garuda u23 perunggu asi games patok masuk empat besar
hansamu yama pranata dkk finis 16 besar capai sebut tak masalah pssi banyak
suara minta milla tahan lihat main timnas baik tak capai target pssi milla bagai
latih timnas beri respons positif ucap anggota exco pssi refrizal goal indonesia
sama muka wakil ketua umum pssi joko driyono rapat malam turut jumlah poin
pertama soal putus angkat milla sang latih merespon baik tawar aju tiga sambung
joko bicara waktu datang milla milla tengah libur jalan latih lisensi uefa pro
kira datang indonesia 9 oktober hari laga uji coba lawan myanmar poin akhir kait
kontrak hitam atas putih sekretaris jenderal ratu tisha destria laku komunikasi
penting ucap joko kutip cnn indonesia hingga legalitas kontrak milla memang
jelas mantan main barcelona tiba indonesia rapat malam hotel sultan jakarta
pusat turut joko federasi bahas soal target piala aff 2018 jauh 11 edisi piala
aff sejak 1996 indonesia sekalipun keluar bagai juara edisi pssi target juara
muka milla status posisi orang dialog nyaman cara fisik muka temu muka kata joko
coach luis milla baru gabung 9 oktober ideal berinteraksi bicara program latih
detail target kontrak jauh joko memang mau bicara detail soal kontrak bocor
aspek tunjang macam gaji jumlah fasilitas bakal penuh pssi masa depan milla
jelas beberapa waktu lalu sebut sempat muka jumlah media masuk suara detik abar
pssi tunggak gaji milla lama tiga bulan total rp6 9 miliar pssi sebut ganti sewa
rumah milla lama tinggal bal sejak februari hingga agustus jalan pusat latih
sebut kemudian tengara jadi buntu negosiasi pssi milla akhir pekan kemarin
refrizal anggota exco pssi tak sangkal kabar sebut turut bayar gaji milla
sangkut langsung balik spanyol indonesia singkir asi games 2018 milla jadi latih
indonesia iring nasib bima sakti panjang minggu lalu bima sebut bagai salah satu
calon ganti milla tangan timnas widodo c putro hingga simon mcmenemy kini mantan
```

main timnas indonesia medio 1990 hingga 2000an jadi asisten milla posisi pegang kurun kuranglebih tahun akhir miguel gandia asisten latih milla kontrak periode latih kiper ubah milla datang tiga miguel latih kiper ganti eduardo perez coach bima bagai asisten latih ujar joko kerja sama millabimagandia benar tak buruk meski singkir babak 16 besar asi games sejati indonesia tampil cukup baik gabung grup a sama palestina hong kong indonesia hasil keluar bagai juara grup raih 9 poin hasil tiga kali menang sekali kalah indonesia singkir lawan uni emirates arab babak 16 besar lewat adu penalti milla tetap jadi latih timnas indonesia buat senang beberapa main timnas evan mas darmono kata milla paham main timnas coach milla tahu main ucap evan mas bolasport com h t kompas com menang 10 atas tuan rumah borneo fc buat persib kini unggul lima poin puncak klasemen liga 1 indonesia wakil indonesia undi berat babak awal china open ganda putra mayoritas wakil merahputih pool atas kevin marcus kejar capai juara semua turnamen super 1000 23 atlet cilik hasil raih beasiswa pb djarum 2018 sisih lebih 5 900 atlet partisipasi',

'kotak hitam baru temu kini semua utama keluarga korban tengah hasil lidi komite nasional selamat transportasi knkt kena jatuh pesawat lion air pklqp baru lepas landas jakarta tuju pangkalpinang senin 29 10 2017 agam spekulasi soal celaka muncul semua baru jawab knkt selesai lidi tahun sekarang ikut papar kembang lidi knkt kabar baru putar insiden jatuh pesawat boeing 737 max 8 baru usia tiga bulan air tanjung karawang kotak hitam black box knkt segera teliti isi rekam baru satu kotak hitam temu tahu flight data recorder fdr cockpit voice recorder cvr investigator knkt bambang iriawan sempat fdr kemudian koordinator investigator selamat udara knkt ony soerjo wibowo nyata fdr cvr badan nasional cari selamat bnpp basarnas kabar kumpar 2 11 serah seluruh temu barang pesawat sebut teliti knkt proses lidi dasar atur menteri pm no 74 tahun 2017 penuh konvensi chicago terbang sipil internasional annex 13 konvensi sebut sesuai annex 13 negara selidik celaka kasus knkt serah lapor awal preliminary report lidi sebut organisasi terbang sipil internasional international civil aviation organization icao waktu 30 hari sejak hari celaka jadi lapor kategori konfidensial rahasia buka publik gantung selidik wenang ony jelas lapor sebut upa data fakta celaka pesawat pklqp analisis jelas jadi ungkap data fakta kata katadata co id 1 11 data fakta bakal papar masuk kronologi jadi data terbang pesawat nama alamat operator pesawat kru tugas celaka jadi kru knkt selidik alam tingkat sehat nama nomor lisensi cantum lindung privasi sangkut lapor muat data badan meteorologi klimatologi geofisika bmkg kena kondisi alam mungkin rahasia ony tegas knkt putus buka lapor awal publik rakyat indonesia hak tahu data fakta tegas lapor akhir cepat mungkin waktu 12 bulan celaka jadi hasil lidi knkt dasar pm 74 2017 jadi rekomendasi selamat laku tindak cegah anggap perlu segera tingkat selamat terbang pihak wenang sangkut masuk negara icao apabila kait dokumen icao jelas wakil ketua knkt haryo satmiko wartakotalive com kapasitas siapa salah peristiwa sebut hasil lidi knkt kata haryo serah menteri hubung cum direktorat jenderal hubung udara lion air bagai operator terbang boeing bagai usaha buat pesawat polisi ri polri nyata selidik peristiwa sebut sisi nonteknis kepala divisi humas polri irjen pol setyo wasisto kutip kompas com 2 11 tugas bentuk tim bareskrimantisipasi teliti sabotase kata setyo tim sebut koordinasi knkt mulai kumpul data butuh hingga jumat 2 11 lapor antaranews com 65 kantong

jenazah datang rs polri kramat jati jakarta timur identifikasi tim disaster victim investigation dvi indonesia automatic fingerprint identification system inafis polri turut kepala rs polri kombes musyafak liputan6 com ambil sampel dna 238 potong tubuh 48 kantong jenazah jumlah baru satu jenazah hasil identifikasi jannatun cintya dewi 24 makam sidoarjo jawa timur kamis 1 11 anggota keluarga seluruh 189 korban serah data antemortem data diri korban mati baru 152 anggota keluarga ambil sampel dnanya cocok dna korban sedikit korban temu buat basarnas putus luas area cari cari tim sar gabung fokus radius 10 mil laut titik jatuh pesawat air tanjung karawang kini cari luas hingga laut kawasan bekas indramayu jawa barat humas basarnas bandung joshua banjarnahor tempo co tiga wni kerja bagai abk art tze kapal bendera singapura culik perompak laut dekat pointenoire kongo senin 29 10 2018 culik aju tuntutan menteri hubung budi karya sumadi evaluasi ulang ketat manajemen selamat maskapai terbang khusus biaya murah low cost carrier lcc formula upah jadi debat usaha nilai formula hitung naik beri sedang buruh nilai naik tak cukup tutup butuh',

'hiu prasejarah megalodon sedang hangat bincang benar hiu gigantik benar lewat trailer film the meg tayang megalodon fact vs fiction discovery channel nama megalodon santer dengar selisik fakta putar megalodon fakta pertama perlu tahu megalodon bukan nama benar ahli paleontologi klasifikasi predator prasejarah masuk genus carcharocles anggap masuk genus carcharodon nama ilmiah carcharocles megalodon carcharodon megalodon berapa besar benar hiu gigantik hingga sekarang laku ahli hiu sekalipun kira kerangka megalodon lengkap temu ragu muncul suatu estimasi ukur megalodon laku sisa tinggal gigi sebar kolom tulang belakang tubuh kotor batu buat sangat tantang iktiologis ahli perikanan john e randall pernah banding tinggi enamel hiu putih gigi megalodon kalau proporsi tubuh megalodon sama hiu putih ikan prasejarah kira milik panjang 13 meter ilmuwan clifford jeremiah lebar akar gigi megalodon hitung ukur megalodon dewasa capai 16 7 meter rumus ahli biologi michael gottfried tim simpul ukur megalodon capai hampir 16 meter berat 48 ton hitung ahli sebut ukur hiu kisar 20 hingga hampir 25 meter bagai banding spesies besar hidup hiu paus panjang hingga 12 1 meter estimasi ahli nyata megalodon lebih panjang jauh lebih berat jadi hiu besar bahkan mungkin ikan besar panjang masa megalodon makan paus fakta ungkap temu fosil tulang paus bekas luka sangat cocok ukur gerigi gigi megalodon banyak hiu megalodon terus lepas gigi fosil gigi temu tiap benua antartika analisis gigi mungkin ilmuwan spesies hidup hingga 2 6 juta tahun lalu ukur benar besar gigi megalodon besar pernah temu panjang capai 19cm banding gigi hiu putih panjang maksimal capai 7 6cm national geographic lapor hiu punya gigit paling kuat makhluk apa pernah hidup rahang besar bahkan taksir lebih kuat rahang tyrannosaurus rex ahli kira megalodon mangsa paus kecil masuk spesies piscobalaena nana mirip paus bungkuk paus balin lebih kecil panjang tak lima meter jelas punah megalodon suhu laut lebih dingin paus kecil ganti paus raksasa paus bungkuk paus biru memang megalodon mangsa penyu duyung dugong ikan asupan makin kurang akhir buat punah hipotesis punah predator super cukup agam dingin suhu samudra saing hiu paus baru peran orang yakin megalodon hidup laut hingga bahkan lapor tampak meski absah ragu live science nyata national oceanic and atmospheric administration percaya 95 persen laut 99 persen samudra jamah beberapa orang mungkin pikir megalodon meski laut sisa banyak rahasia ahli

biologi laut craig mcclain banyak alas ilmiah kuat pikir megalodon sembunyi laut kalau megalodon sekarang lihat gigi sebar mana hiu milik fosil 2 6 juta tahun lalu kata mcclain deep sea news konsensus teliti megalodon punah kisah megalodon terus hidup lewat novel karya fiksi 2013 discovery channel siar tayang judul megalodon the monster shark lives tayang tunjuk ilmuwan cari hiu panjang 20 meter panjang pantai afrika selatan dokumenter palsu nilai sesat bagaimana fakta megalodon memang pernah hidup laut bumi fosil dinosaurus 110 tahun sebut spesial awet cara sempurna ibarat lukis mona lisa tak cela yakin homo sapiens hidup 350 ribu tahun lalu jadi spesies manusia purba pertama cipta bahasa komunikasi nyata bangkai makhluk mirip monster loch ness temu buah pantai as meski banyak pakar duga bangkai hiu busuk',

'persib bandung milik peluang besar juara liga 1 indonesia musim 2018 sayang peluang ancang hilang ihwal pssi jatuh sanksi maung bandung atas tewas haringga sirla persib jamu persija jakarta minggu 23 9 2018 sanksi larang tim sepak bola bangga warga bandung sebut tanding stadion gelora bandung laut api gbla kandang persib sisa musim kompetisi tahu lewat putus komite disiplin komdis pssi ketuk palu senin 1 10 publikasi esok hari alias Selasa 2 10 empat sanksi tanggung pihak penting persib akibat kasus sebut komdis putus beri hukum persib upa sanksi tanding home luar pulau jawa kalimantan tonton akhir musim 2018 tengah musim kompetisi tahun 2019 tulis pssi laman resmi tanding luar pulau jawa lantas bobotoh juluk dukung persib tetap tonton hukum pssi suporter persib jangka waktu rupa sisa kompetisi 2018 hingga tengah kompetisi 2019 hukum lebih detail sila klik sini hukum sebut cukup rugi skuat persib pasal kini jadi tim paling unggul raih trofi liga 1 indonesia 2018 hingga pekan ke23 persib kini kokoh puncak klasemen emas 44 poin hasil 13 menang 5 imbang 5 kali kalah posisi cukup jauh kejar saing bawah psm makassar 38 poin bal united 37 poin persija 36 poin bila lihat kerja panjang 2018 capai apik sebut tak lepas bobotoh pasal persib tampil nyaris sempurna kala tanding kandang 11 laga lakon gbla pangeran biru nama persib mengantongi sembilan menang sekali imbang sekali kalah banding laga tandang hingga 12 tanding luar kandang jauh persib buku empat menang empat imbang empat kali kalah fakta sebut sulit rasa nafi peran gbla bobotoh laju impresif persib musim 11 laga sisa mana 6 main kandang tak keliru rasa bila sangat peluang bawa trofi liga kali tiga hukum pssi kini jadi batu sandungan mimpi sebut tambah beri sanksi serang andal ezechiel ndouassel top skor persib 14 gol sebut larang pssi main lima tanding depan anggap sikut sengaja dorong kepala main persija daftar main persib persija hukum baca berita lengkap <https://t.co/zy68vmgq5s> pic.twitter.com/qdxx5kz7jg karuan hukum federasi tersembut sambut emosi persib manajer persib bandung umuh muchtar misal umuh putus sebut sangat rugi tim penuh rekayasa putus adil benar putus sangat rekayasa ucap umuh wartawan h t goal indonesia kalau mau tuntutan pssi bubar adil sama ungkap roberto mario carlos gomez latih persib bandung latih asal argentina memang dengar langsung manajemen soal sanksi pssi resmi bila sanksi sebut resmi turut mario gomez hukum sebut lebih bahkan kata lebih baik persib main sisa musim kompetisi lebih mau main lebih baik usah main kata Kompas.com persib cara lembaga sangat kecewa putus komdis kini ajar isi putus komdis lanjut segera aju banding berat memori banding tulis manajemen laman resmi persib liga 1 indonesia gulir sembilan hari pssi setop Selasa 25 9 2018 jadwal sebut lebih cepat nyata menteri

pemuda olahraga imam nahrawi sebut kompetisi henti pekan usai temu komite eksekutif pssi umum mula kompetisi sepak bola atas indonesia nyata sekretaris jenderal pssi ratu tisha destria Selasa 2 10 pssi minta pt liga indonesia baru lib jalan liga 1 2018 hitung tanggal 5 oktober 2018 ucap tisha situs resmi pssi salah satu alas henti liga 1 kata ketua pssi selidik cara tuntas kasus tewas haringga cari solusi tak ulang masa datang edy nyata bakal rumus standard operating procedure sop baru lebih tegas benar tisha umum segera mula liga 1 sanksi persib dengar solusi sop ambil pssi cegah ulang kasus sama segera hasil bicara bakal hasil sop lebih keras redam henti giat rugi jadi terus ulang kata edy bahkan tahu pssi rembuk semua klub liga kelompok suporter asosiasi main cari akar masalah sepak bola indonesia rumus selesai tampak tulis panditfootball.com pssi rasa cukup hukum persib bobotoh benar tutup mata keras suporter kadung akut laten jadi tahun lama bagai kota suporter turun lapang suara protes kerja klub aksi justru buat klub ancam sanksi keras harap bonek penuh manajemen baya iwan legowo pecat putus komdis pssi untung bhayangkara peluang bal united tip psm persib kini tunggu sanksi',

'sejak sutradara ngenest 2015 bagai film panjang perdana Ernest Prakasa kerap kolaborasi musisi kali gerak makin jauh musisi libat bukan jadi nyanyi soundtrack turut main Milly Mamet the movie mmtm milly peran Sissy Prescillia Mamet Dennis Adhiswara rupa tokoh film apa cinta aadc produksi Miles Films sama Starvision Ernest hadir tokoh sebut satu film ikat semesta aadc kali diri sendiri dunia film langkah sebut kerap istilah spinoff sempal film solo cerita tualang tokoh han solo tokoh film Star Wars contoh terap buah spinoff beda solo kisah tualang masa remaja han Ernest lebih pilih cerita lanjut Milly Mamet selang beberapa bulan jadi film aadc 2 2016 kisah meni milik orang anak mmtm potret lebih dekat hidup rumah tangga milly mamet segala problematikanya turut hadir tokoh konferensi pers kenal main Cassadina Kitchen Bakery Kuningan Jakarta selatan 27 8 2018 Ernest gaet tiga nyanyi sekaligus semarak film rencana tayang Desember 2018 tiga nyanyi sebut Melly Goeslaw bagai mamah itje Eva Celia Jojo Isyana Sarasvati Rika Melly 44 lama kenal bagai vokalis kelompok potret sama suami Anto Hoed kerap jadi tata musik komposer nyanyi lagu tema film original soundtrack alias OST Melly pernah jajal dunia akting lalu film kabayan jadi milyuner 2010 Brokenhearts 2012 mmtm mamah itje cerita orang selebritas instagram selebgram nyanyi ikut ajak gabung Eva Celia 25 sama Melly keponakan Mira Lesmana produser aadc bintang empat film akhir kali ambil film pendekar tongkat emas 2014 lalu pesan pendek beritagar id 28 8 Ernest kata Jojo peran Eva Tetangga rumah milly isyana 25 peran tokoh Rika lengkap trio nyanyi mmtm Solis Andal Sony Music Entertainment Indonesia label rekam tempat Ernest dulu jabat bagai assistant label manager jadi sekretaris James Peran Yoshi Sudarso Meira Anastasia bagai tandem Ernest tulis skenario sebut pilih tiga nyanyi tadi dasar sesuai cerita bikin dulu cerita baru pilih siapa kira karakter cocok peran ungkap Meira istri Ernest jadi asisten sutradara Anastasia Anjani Nabila Aulia beritagar id tiga nyanyi rangkap jabat bagai lantun soundtrack mmtm Lontar jawab spesifik Ernest kata tetap percaya nyanyi Sony Music Entertainment Indonesia nyanyi sebut Isyana tunggu jawab sang sutradara fokus laksana syuting langsung mulai 2 September 2018 nama main atas ikut syuting Julie Estelle Dinda Kanyadewi Surya Saputra Roy Marten tinggal peran komika muhaddkly Acho Awe bintang Emon

arafah rianti aci resti ardit erwandha sinambung aadc jalin turut hadir tampil dian sastrowardoyo bagai cinta adinia wirasti karmen titi kamal maura kata ernest porsini tampil geng cinta banyak ingat film rupa panggung milly mamet', 'percaya tim janji butuh waktu tumbuh nyata latih gennaro gattuso pascatimnya ac mil menang telak 41 atas sassuolo tanding enam liga serie a italia tanding markas sassuolo mapei stadium reggio emilia italia senin 1 10 2018 dini hari wib mil buka keran gol via tendang kaki kiri franck kessie menit ke39 fernandez saez suso castillejo azuaga lebar unggul mil menit ke50 60 suso kemudian tutup pesta gol mil menit empat tambah waktu babak satu gol balas sassuolo cetak filip djuricic menit 68 hasil mil hasil akhir laju kurang impresif kompetisi tinggi italia sebut pasal lawan empoli atalanta cagliari pasu kota mode italia tuai hasilimbang maksud gattuso soal butuh waktu tadi cara materi maupun capai klasemen liga italia musim tiga tim atas hitung kalah banding mil sial mil tetap tak mampu raih poin penuh justru hadap sassuolo notabene tengah posisi tiga tanding langsung mil mampu raih poin penuh tahu betul bagaimana derita lantas tunjuk kualitas tim butuh baik tampil ucap gattuso laman resmi mil menang jadi tanda bangkit mil rupa dukung i rossoneri harap bila lihat jalan tanding mil benar main cara solid lihat proses cipta gol semua gol buat il diavolo jadi lewat tendang luar kotak penalti mayoritas tak lalu lewat skema main tim aksi individu masingmasing main 3 ac mil scored 3 goals from outside the box in the same serie a match for the first time since october 2016 vs chievo aim sassuolomilandan gattuso aku skema gol luar kotak penalti bukan cara rencana laga arti cara memang rupa insting main bukan siap milik banyak peluang luar kotak penalti lawan empoli milik banyak peluang lewat cara sebut musim kata gattuso bekas main mil era 2000an meski turut gattuso menang sangat penting mil tiga poin mampu tumbuh percaya diri tim tim milik banyak bagus hasil bantu tingkat percaya diri segala kerja ucap gattuso lansir goal menang bukan angkat moral mil cara tim per individu main misal apa jadi suso tanding langsung serang asal spanyol lama paceklik gol 1 728 menit suso tak cetak gol catat opta turut suso menang sangat penting mil sassuolo main bagus mil milik lebih agresif determinasi tinggi ucap suso gagal leta tanding selalu percaya mampu menang mil mulai baik posisi klasemen enam tanding lalu emas sembilan poin tim kota inter mil sebut duduk posisi 10 klasemen sassuolo tanding tempat posisi tiga kini turun urut lima atas latih sassuolo roberto de zerbi tak terlalu kecewa bila kecewa turut zerbi gagal cetak gol babak pertama ucap selamat tim muda mudah tahan posisi klasemen kondisi tukas zebri sky sport h t newsmondo it kalah jadi ajar sassuolo terus mimpi coba takluk liga europa hilang lama beberapa tahun cardiff city 1 2 burnley epl 2018 2019 bologna 2 1 udinese liga serie a 2018 2019 villarreal 0 1 real valladolid la liga 2018 2019',

'kunjung bal akhir pekan lalu presiden joko jokowi widodo tiba umum program baru dana lurah mulai tahun depan perlu utama kota nama anggaran lurah kata jokowi kutip lapor inews jumat 19 10 2018 jokowi aku perintah jajar kerja cepat realisasi dana lurah masuk revisi atur perintah pp bagai payung hukum program program tiap lurah kota bakal dana segar bangun infrastruktur fasilitas lurah sambung ucap jokowi wakil menteri uang mardiasmo kata anggaran program masuk rancang anggaran belanja negara rapbn 2019 total anggaran rencana capai rp3 triliun ambil alokasi dana desa tahun 2019 jumlah rp73 triliun meski anggaran dana tetap

otomatis cair payung hukum dana desa program jalan lancar undang uu nomor 6 tahun 2014 desa mardiasmo tengah upaya opsi payung hukum bakal bahas direktorat jenderal imbang menteri uang pertama buat undang dana lurah sama dewan wakil rakyat dpr cara susun revisi pp cara yakin lebih cepat perlu lalu agam bahas sama dpr lebih kalau revisi pp coba pp sambung mardiasmo Kompas.com Senin 22 10 2018 cara garis besar program rupa dana desa jalan sejak duet Jokowi-Jusuf Kalla jalan perintah empat tahun silam program langsung tuai kritik lebih muncul tahun politik kubu oposisi wakil ketua Majelis Syura Partai Adil Sejahtera PKS Hidayat Nur Wahid nilai perintah buat program sejak lama timbul kesan politisasi sangat rawan kemudian disalahpahami bahas DPR kalau betul kadang kan wacana nyata enggak sebut Hidayat Lapor Liputan6.com cara prinsip dana desa dana lurah mirip sokong bangun infrastruktur fasilitas tingkat desa maupun lurah cara laksana program milik tiga beda satu cara terminologi uu nomor 6 tahun 2014 desa jelas desa satu masyarakat hukum milik batas wilayah milik wenang atur urusan penting masyarakat tempat dasar prakarsa masyarakat hak asal usul hak tradisional aku hormat sistem perintah negara rujuk uu nomor 17 tahun 2018 Camat Lurah perangkat Camat Camat Rupa wilayah daerah kabupaten kota pemimpin Camat wewenang asal Bupati Wali Kota pemimpin teritori pilih cara beda desa pemimpin kepala desa pilih langsung lurah pemimpin lurah rupa pegawai negeri sipil PNS tunjuk Bupati Wali Kota beda cakup luas wilayah desa lurah beda jelas lebih jauh menteri negeri Tjahjoe Kumoroto turut kutip detik.com cakup wilayah lurah lebih kecil banding desa alokasi anggaran tak besar desa meski lurah jadi milik jumlah penduduk lebih banyak desa soal infrastruktur desa lebih kompleks luas kendati tak rumit desa dana lurah tak rata alokasi dana lurah nilai penting tunjang perata bangun perintah kota nyata banyak lurah minim anggaran mampu tempat posisi bagai lurah suatu kota baik sangkut sarana prasarana maupun fasilitas umum sambung Tjahjoe tiga formulasi dana menteri uang Sri Mulyani Indrawati jelas dana lurah indikator beda desa desa perintah laku formulasi dasar jumlah penduduk miskin tingkat tinggal penduduk sayang Sri Mulyani papar lebih jauh formulasi dana lurah pakai pihak aku rumus sama jajar menteri negeri lepas bagaimana prokontra salur dana lurah perintah perlu salur dana daerah benar efektif fakta salur dana desa bilang efektif soal korupsi data menteri uang lansir serap dana desa sangat rendah September 2018 serap capai Rp37,9 triliun pagu anggaran besar Rp60 triliun dana desa jabat tahu dana benar apa bangun apa hingga bagaimana kelola ucap ekonom Universitas Indonesia UI Fithra Faisal CNN Indonesia fakta pilkada langsung serentak efektif bukan arti pilkada langsung lalu DPR jadi alternatif pilkada masa depan transfer dana pusat daerah dana desa Rp780 triliun padahal akuntabilitas kerja kepala daerah rendah apa jadi perintah pusat buat bagai terobos cepat bangun irama bangun daerah lambat',

'satu kotak hitam pesawat Boeing 737 MAX 8 maskapai Lion Air jatuh air tanjung karawang Jawa Barat Senin 29 10 2018 hasil temu Kamis 1 11 hari ping locator tangkap sinyal pancar kotak hitam instrumen penting pesawat benar warna oranye sebut temu Sertu Hendra Selam Batalyon Intai Amfibi Taifib Tim SAR TNI angkut laut rupa Tim SAR gabungan siar pers dinas terang angkat laut terima beritagar id 1 11 jelas Hendra temu kotak hitam duga isi data terbang flight data recorder FDR pukul 10 15 WIB 30 meter lokasi 500 meter titik jatuh pesawat kotak hitam bawa kapal riset Baruna Jaya milik Badan Kaji Terap Teknologi BPPT lanjut kotak

hitam serah komite nasional selamat transportasi knkt periksa lebih lanjut konferensi pers kantor menteri hubung Kamis 11 menteri hubung budi karya sumadi konfirmasi temu kotak hitam sebut jelas flight data voice recorder kata budi wartawan masuk elisa valenta beritagar id knkt kemudian konfirmasi kotak hitam temu fdr isi flight data recorder fdr cockpit voice recorder cvr temu ujar investigator komisi nasional selamat transportasi knkt bambang iriawan liputan6 com kotak hitam memang diri fdr rekam data lama terbang rekam suara cakap kokpit sebut cvr letak kotak pisah tempat posisi sama tubuh pesawat biasanya dekat ekor temu buka jalan tahu celaka timpa pesawat registrasi pklqp nomor terbang jt610 jakarta tuju pangkalpinang bangka belitung sebut knkt butuh waktu lama tahun periksa investigator knkt ony suryo wibowo jelas unduh data selesai jam proses analisis butuh waktu lama tahun knkt ujar ony tasya simatupang beritagar id 30 10 beri waktu satu tahun selesai lidi beri preliminary report lapor satu bulan preliminary report kondisi pilot bentuk serpih pilot lisensi bagai jelas ony tim disaster victim identification dvi polisi ri Rabu 31 10 2018 hasil identifikasi jenazah pertama korban jatuh pesawat masakapai lion air pklqp rute jakartapangkalpinang Senin 29 10 jenazah sebut identifikasi nama jannatun cintya dewi perempuan usia 24 tahun rupa karyawan menteri energi sumber daya mineral esdm Kamis 1 11 pagi kabar antaranews jenazah tiba rumah orang tua korban desa suruh sukodono sidoarjo Jawa Timur langsung makam anak tiga pasang bambang supriyadi surtiyem menteri esdm rasa hilang salah satu putri baik tempat kata endang sutisna wakil sekretariat jenderal menteri esdm hadir makam jannatun akrab sapa yayas terbang pesawat boeing 737 max 8 kode rute jt610 rangka tugas lama tiga hari pangkalpinang pulau bangka belitung sejak kantong jenazah pertama tiba rs polri baru jenazah jannatun hasil identifikasi utuh kondisi tubuh korban buat proses identifikasi makan waktu lama hingga Kamis 31 10 53 kantong jenazah tiba rumah sakit sebut jumlah kantong cermin jumlah korban temu peristiwa nahas jadi 189 orang pesawat sebut kapusdokes polri brigjen arthur tampi okezone com Kamis 31 10 nyata periksa 48 kantong jenazah risi potong tubuh korban jenazah jannatun kantong nomor register 00 lion tanjung priok 0010 436 x 2018 masuk kondisi relatif baik lengkap masuk jari tangan kapus inafis bareskrim polri brigjen hudi suryanto papar lalu ambil sidik jari korban periksa data tunggal rekam sidik jari kartu tanda duduk elektronik ektp lewat periksa lanjut hudi muncul identitas korban sidik jari ambil milik sama 13 titik data ektp turut ilmu daktiloskopi ilmu ajar sidik jari kalau temu sama 12 titik sidik jari banding identik meski satu jari manusia jelas hudi nah jenazah temu 13 titik sama cincin pasang jari tengah tangan kanan sidik jari cocok tim inafis laku klarifikasi lanjut banding data data antemortem data diri korban tinggal dunia serah keluarga tim inafis lalu lihat foto jenazah anggota keluarga utama lihat cincin tangan benar jannatun cintya dewi keluarga lihat foto takut kuat siap jadi lihat foto properti cincin jari tengah kanan jelas kepala bidang dvi mabes polri lisda cancer bbc indonesia jenazah utuh tim tanya keluarga segera ambil jenazah kubur tunggu utuh keluarga jannatun turut lisda kemudian putus segera bawa jenazah kubur sidoarjo ideal kata lisda proses pulang jenazah baik tunggu anggota tubuh lengkap laku asal pihak keluarga sabar tunggu ideal kan tahu lengkap kan tahu kapan lengkap dvi tak perlu identifikasi korban kelakaan pesawat teroris tewas sama bom perlu rekam cakap kokpit cvr kuat temu hasil

himpun lalu fdr rusa komponen indikator cepat udara sensor aoa celaka jt610
segera simpul hari ke11 operasi cari tim sar gabung ekor pesawat airbus a320
pkaxc qz8501 temu biasa kotak hitam pasang ketemu',

'dagang bursa efek indonesia bei tutup petang jumat 09 11 2018 koreksi turun 1
72 persen 102 65 poin level 5 874 15 panjang dagang saham hari ihsg sempat capai
5 936 03 poin titik tinggi rendah 5 859 09 poin total frekuensi dagang saham
hari capai 357 32 ribu kali nilai transaksi rp8 16 triliun banyak 116 emiten
alami kuat tahan gera ihsg 275 emiten lemah 224 emiten alami gera dewata
freightinternational tbk buku kuat harga saham paling signifikan dagang jumat 09
11 emiten sandi deal sebut pimpin kuat saham top gainer tutup kuat level rp254
69 33 per lembar sahamkota satu properti tbk catat lemah harga saham parah
dagang hari emiten sandi satu sebut duduk peringkat akhir top loser tutup lemah
22 50 persen level rp124 per lembar saham rp160 per lembar tutup h m sampoerna
tbk hmsh rupa saham paling banyak diperjualbelikan top mover total frekuensi
capai 35 50 ribu kali 10 sektor saham seluruh alami lemah lemah parah alami
sektor industri konsumsi capai 5 08 persen 123 88 poin ikut sektor manufaktur
besar 3 13 persen 47 50 poin lalu saham sektor tambang susut hingga 1 13 persen
21 20 poin adapun tiga sektor saham lemah paling tipis dagang susut besar 0 03
persen sektor industri dasar besar 0 44 persen sektor tani besar 0 28 persen
aksi jual saham investor asing raup dana rp2 64 triliun utama saham panin
financial tbk pnlf lipso karawaci tbk lpkr bumi resources tbk bumi investor
asing laku beli nila rp2 59 triliun khusus saham astra international tbk asi sri
rejeky isman tbk sril pakuwon jati tbk pwon wilayah mandailing natal sumatra
utara ancam hingga hari depan cuaca ekstrem posisi mata uang rupiah nilai sangat
rawan gejolak eksternal picu alir modal asing keluar salah sensor aoa kelabu
posisi hidung pesawat pilot terkadang tambah cepat tetap terbang padahal posisi
hidung pesawat tukik bawah',

'harga mahal bayar ac mil tahan imbang 11 tuan rumah real betis tanding empat
grup f liga europa jumat 9 11 2018 dini hari wib tiga main alami cedera ragu
tampil lawan juventus akhir pekan tiga main sebut franck kessie hak calhanoglu
mateo musacchio tiga lengkap derita mil lebih dulu hilang lucas biglia empat
bulan depan jadi kabar tak senang kubu i rossoneri pasal kini tengah alami tren
apik semua kompetisi empat tanding akhir tak pernah kalah satu imbang tiga
menang posisi tim asuh gennaro gattuso perlahan mulai naik klasemen liga serie a
italia bila pekan hingga sembilan luar posisi 10 besar kini mil empat besar
pekan ke11 masalah tren apik bakal temu sandungan akhir pekan jamu juventus
stadion san siro mil italia kondisi skuat compangcamping bukan perkara mudah
bahkan sulit mil lanjut tren positif tadi gattuso buat pusing kondisi tim besok
hitung berapa banyak benar main sisa ucap gattuso sky sport italia h t football
italia mantan main mil memang hitung ulang jumlah skuat bugar pasal bukan empat
main sebut atas tengara tak tampil waktu dekat gonzalo higuain jack bonaventura
mattia caldara davide calabria empat ragu tampil akhir pekan datang juventus
salah satu tim baik eropa italia duduk sini menang hadap main sepak bola cara
ucap gattuso tanding langsung stadion benito villamarin sevilla spanyol betis
buka unggul lewat tendang voli giovani lo ceslo menit ke12 untung mil milik suso
tendang bebas sisi kiri gawang betis menit ke62 hasil robek gawang pau lopez 13
suso has been directly involved in 13 goals 5 goals 8 assists in 14 games played

so far in 2018 19 season fixed betismilanharga mahal tanggung mil bukan cedera main sempat lolos fase grup f liga europa kini tahan peringkat nilai tujuh poin bawah betis delapan poin olympiakos kuntit belakang mil nilai sama jadi kabar buruk mil pasal jalan laga tandang markas olympiakos laga pamungkas grup f tengah desember bila tuai hasil kalah alamat buruk publik mil mungkin lolos babak grup tak mungkir mil rupa klub kaya sejarah prestasi benua eropa satu dekade akhir kondisi tak terlalu baik terseokseok liga domestik kurang apik eropa musim salah satu bukti bila laju mil jadi tanda tanya jumlah klub mapan eropa mentas liga europa beberapa satu tiket fase 32 besar liga kasta benua eropa sebut hingga tanding empat tujuh tim mampu injak kaki babak gugur arsenal chelsea eintracht frankfurt dinamo zagreb lazio bayer leverkusen fc zurich results 64 goals scored 7 teams through to the round of 32 arsenal chelsea frankfurt dinamo zagreb lazio leverkusen zurich uel pic twitter com anpxchblpdmayoritas klub lang babak lanjut raih hasil positif dini hari tadi chelsea kalah tuan rumah bate borisov skor 10 lazio tunduk tamu marseille skor 21 arsenal main imbang 00 tamu sporting lisbon zurich derita kalah 01 leverkusen hasil lolos musabab peringkat tiga grup a ludogorets razgards hasil kumpul poin empat tanding andai tang tanding sisa poin akhir ludogorets tak mampu lewat koleksi zurich peringkat poin sembilan adapun menang chelsea atas bukan langkah babak gugur sang cetak gol olivier giroud puasa sejak mei 2018 794 olivier girouds goal for chelsea was his first in 794 minutes of action for the club since netting against liverpool in may relief batcbe europaleague puas tampil giroud tahu tiap serang cetak gol bagus percaya diri ucap latih chelsea maurizio sarri bbc langkah arsenal 32 besar kabar buruk striker danny welbeck tepi waktu cukup lama alami cedera engkel kanan menit 25 mungkin patah sedang rumah sakit soal kabar tunggu turut cedera serius ucap juru taktik arsenal unai emery anthony ginting maju empat final fuzhou china open 2018 hadap shi yuqi lawan kalah final regu putra asi games 2018 indonesia tak pernah tang partai perdana tiga edisi piala aff akhir edisi piala aff indonesia selalu kalah singapura jonatan christie hadap anthony sinisuka ginting rebut tempat empat final fuzhou china open 2018']

```
[13]: norm_test_corpus = normalize_corpus(test_corpus)
```

```
[14]: norm_test_corpus[0:10]
```

```
[14]: ['mata rupa jendela jiwa manfaat artificial intelligence ai cerdas buat baca pribadi turut buah teliti baru terbit jurnal frontiers in human neuroscience teliti kembang buah teknologi baru manfaat cerdas buat deteksi gera mata uak pribadi orang kerja kata mampu revolusi bagaimana manusia komunikasi mesin teknologi depan sistem ponsel pintar paham prediksi perilaku potensi tawar layanan lebih personal teknologi sebut robot tugas tani orang lanjut usia mobil kemudi sendiri gim video interaktif teliti university of south australia kerja sama univeristy of stuttgart max planck institute for informatics jerman flinders university australia selatan rancang algoritma mampu ajar mesin kemampuan komputer tindak cara eksplisit diprogram menggunakan data kumpul 42 flinders university algoritma ajar mesin rekam gera mata laku tugas hari headset lacak
```

mata buat sensormotoric instruments kena headset lama puluh menit kemudian teliti periksa cara silang temu kuesioner biasa indikasi ciri pribadi dominan lalu teliti periksa subjek lama laku giat beda misal jalan toko banding jalan laboratorium periksa prediksi algoritma buat skenario dasar ciri pribadi big five buka hati nurani ekstraversi ramah neurotisisme cerdas buat mampu identifikasi empat butir neurotisisme extraversion ramah hati nurani hasil teliti hasil tebak semua ciri pribadi buka dasar hasil teliti simpul gera apa wakil pribadi bawah salah satu kontribusi utama kerja tunjuk pertama kali tingkat neurotisisme ekstraversi ramah teliti rasa tahu perseptual tipe pribadi individu prediksi gera mata rekam lama laku tugas hari tulis teliti studi perlu perhati cerdas buat memang mampu prediksi jenis pribadi mampu laku tingkat akurasi sangat tinggi sempurna lebih lanjut teliti pikir teknologi macam cara dramatis tingkat interaksi mesin beri mampu asisten virtual cara baca suasana hati pribadi tobias loetscher university of south australia libat teliti kata buah unggah situs universitas teliti tunjuk mungkin masa depan robot benar paham emosi orang hadap orang selalu cari layan tingkat personal kata loetscher robot komputer sadar cara sosial adaptasi isyarat nonverbal teliti beri sempat kembang robot komputer jadi lebih alami lebih baik tafsir sinyal sosial manusia urai olivia carter orang ahli saraf university of melbourne kata new scientist meski temu tarik beri tanda masalah privasi baru teknologi lacak mata informasi sama rekam mata frekuensi bicara pungkash turut sebut mudah rekam tahu orang sebut teliti klaim cara cepat murah mudah akurat buah studi baru tunjuk cerdas kait tingkat konektivitas antar beberapa area otak kurang interaksi studi tunjuk giat lamun rupa tanda otak efisien laku agam tugas tinggal informasi',

'coldplay rilis buah dokumenter judul a head full of dreams lalu platform video manasuka amazon prime video ingat 20 tahun diri band sebut jadi buah band besar kata chris martin rekam video 1996 kutip rolling stone guy will jon chris jangan kalian lupa besar besar ucap sang vokalis kemudian jadi nyata martin jonny buckland gitar will champion drum guy berryman bas jadi coldplay sebagai salah satu kelompok musik paling populer pengaruh dunia a head full of dreams cerita lebih kena band poprock asal inggris band anak kuliah manggung kafe lokal london hingga jadi musisi kelas dunia hasil ratus juta dolar lansir hollywood reporter 12 10 2018 dokumenter sedia cara eksklusif amazon prime video inggris amerika serikat australia selandia baru 16 november 14 november film putar 2 000 bioskop seluruh dunia a head full of dreams sutradara mat whitecross sineas usia 41 kenal coldplay sejak band sebut diri kuliah london whitecross bahkan pernah rekam musik coldplay empat orang anggota dengar musik pertama kali kamar asrama rupa hubung baik terus jaga whitecross garap beberapa videoklip coldplay paradise a sky full of stars adventure of a lifetime sebagai sutradara dokumenter whitecross punya banyak alam pernah film jalan oasis band rock asal inggris oasis supersonic lelaki profesi sebagai sunting adegan garap dokumenter the shock doctrine 2009 moving to mars 2009 dokudrama the road to Guantanamo 2006 judul a head full of dreams ambil titel album studio tujuh coldplay rilis 2015 album single pertama adventure of a lifetime ikut tur dunia hasil banyak uang aksi panggung coldplay 114 konser tur a head full of dreams tonton 5 juta lebih gemar sebar 31 negara cara total tur tujuh coldplay kumpul 523 juta as rp7 9 triliun jadi tur laris tiga dunia 360 degree u2 a bigger bang the rolling stones tingkat

sangat besar banding rekor tur empat coldplay viva la vida gelar 2008 hingga 2010 tur hasil 209 4 juta lansir allmusic empat anggota coldplay pertama kali kumpul 1996 kuliah university college london jalan kafe pub cukup panjang hingga dapur rekam parlophone records tarik rekrut coldplay april 1999 tahun ikut album parachutes rilis coldplay mulai jadi bincang banding band besar inggris radiohead oasis musik lebih santai lirik melodi ngepop musik coldplay seberang amerika serikat bahkan seluruh dunia 2002 coldplay rilis album a rush of blood to the head seluruh dunia lagu jago the scientist album buat coldplay laku tur global pertama lalu mula dominasi atas bandband besar seluruh dunia lalu banyak sama luar biasa tetap sama bagai buah kelompok kata martin cuplik film sebut',

'nakita id sarita abdul mukti faisal harris resmi cerai rabu 14 11 kemarin adil agama jakarta selatan humas adil agama jakarta selatan jarkasih perkara perdata nomor 2620 pdt g 2018 pa js cerai gugat gugat aju saudara sarita lalu faisal harris putus adil agama jakarta selatan kemarin rabu 14 11 2018 jelas jarkasih sejak awal aju cerai sarita memang minta banyak syarat faisal mulai hak asuh anak harta gono gin sarita cabut semua tuntutan sebut rasa cepat selesai perkara cerai perkara kan kalau lihat awal banyak tuntutan hak asuh anak harta kemudian nafkah harta sama kemudian memang iring jalan perkara cabut semua asesor pokok perkara ucap jarkasih kutip grid id sarita sama sekali kaya peser mantan suami sekarang faisal aku sesal pilih perempuan jennifer dunn putus serah semua harta sarita ungkap acara sarita elza syarief jumat 16 11 kemarin bikin draftnya atur semua harta faisal jatuh semua sarita haris rasa salah milik perempuan tutur elza semua harta serah sarita kira apa kaya faisal sebut bagai usaha tajir lansir laman tribun wow faisal harris rupa orang usaha bisnis properti mobil mewah sempat jabat bagai vice president ferrari owners club indonesia foci 2015 hingga 2016 foci rupa komunitas milik mobil mewah merek ferrari sedang sepakat faisal sarita ibu empat anak aku mobil rumah mewah mobil sport mclaren mobil jeep sayang mobil anak kecil rumah pondok indah memang rumah anak beliau putus keluar rumah ujar sarita lansir laman bangka pos sedang sisi harta faisal harris sarita punya harta kaya sendiri lansir tribun seleb sempat terjun dunia hiburan tanah air bagai model video klip lagu nyanyi inka christie kini sarita tengah tekun kerja baru bagai orang selebgram terima banyak endorsement bak selebritis papan atas tanah air sarita bahkan lebar sayap menjagi vlogger bayang beberapa kaya milik meski tak nafkah mantan suami bahkan resmi cerai sarita anak lebih dulu milik rumah mewah bagai mobil mewah',

'akhir lancong makin gari banyak kalang banyak alternatif plesiran luar negeri lancong olah makin beri mudah tempat libur tarik unjung tempat baru pernah kunjung jadi solusi jenuh landa jadi momen senang libur perlu rencana cara matang pilih lokasi siap akomodasi hingga hitung bujet milik rupa penting tak lewat perlu hitung risiko mungkin jadi lancong libur senang balik jadi sumber masalah ikut beberapa risiko lancong jadi siapa ubah batal jadwal terbang masalah paling sering alami lancong ubah batal jadwal terbang dadak sangat rugi lancong utama milik waktu cuti batas rencana jalan buat cara matang jadi beranta bagasi rusak hilangselain ubah batal jadwal terbang cara dadak masalah sering alami lancong bagasi rusak hilang risiko macam jadi siapa rencana libur jadi ganggu barang bawa penting pakai obat lengkap butuh beli keluar tambah dokumen jalan hilangbagi pergi tempat baru kenal cukup jauh wajib ekstra hati jaga baran

bawa harga misal dokumen jalan tak duga curi copet lalai buat hilang dokumen penting libur jadi peristiwa sebut jalan nyaman repot buat lapor hilang urus dokumen jalan pulang negara kota asal celaka sakit lama perjalanan musibah jadi kapan tak sedang libur celaka sakit lama jalan jadi salah satu risiko buruk hadap lancong sebut jadi libur ganggu momen libur senang ubah jadi senang kondisi fisik bugar sakit akibat celaka hindar risiko sebut atas meyasatinya laku siap matang pergi salah satu penting siap lindung diri asuransi jalan beli asuransi jalan happytrip happyone id santun lambat jalan ganti biaya jalan bayar muka bila jadi batal jalan luar kuasa happytrip beri lindung atas hilang rusa bagasi masuk barang pakai bawa lama laku jalan sakit alami celaka jalan ganti biaya medis santun celaka perlu yuk cari tahu lebih lanjut happytrip lalu taut jangan langsung panik rasa nyeri area dada payudara jadi bukan ganggu serius lengkap ruang ac olah jadi mandatori tingkat nyaman rumah lebih matahari luar ruang asa sangat sengat vaksin baru cegah tbc temu beri harap baru pasien orang',

'musik nyata dampak manusia buah studi baru ungkap fakta laboratorium apa grup musik ac dc lagu tahun 1980 meski ikon musik metal sebut lantun rock and roll as not no noise pollution nyata musik rok milik efek bahaya sistem lingkung turut greg kot rock and roll acu gaya musik populer amerika serikat tahun 1950an lalu kembang tengah 1960an buat lebih kenal cara luas musik rok meski istilah rock and roll terus lanjut gaya rock and roll tahap awal instrumen piano saksofon biasa instrumen utama instrumen umum ganti tambah gitar era tengah hingga akhir 1950an rock and roll klasik biasa main satu gitar listrik satu lead satu ritme double bass string bass tengah 1950an gitar bas listrik set drum laboratorium mississippi state university msu as laku uji coba sebut ac dc hypothesis ahli satwa liar kemudian temu musik ac dc guns n roses metallica pengaruh bagaimana serangga serang hama hambat tumbuh tanam alir musik milik pengaruh bahkan main tingkat volume sama asisten profesor department of biology msu brandon barton tim risi mahasiswa pengaruh bagai alir musik suara ekosistem baton kata inspirasi laku teliti dengar salah satu grup musik favorit dengar ac dc paham lagu pikir rock and roll ain t noise pollution rupa buah hipotesis uji kata barton hipotesis alternatif rock and roll rupa polusi suara uji teori tim sebut merngumpulkan jumlah kepik bagaimana musik pengaruh konsumsi kutu daun hama serang tanam kedelai alir rok country folk tim teliti tambah suara khas lingkung kota suara mesin bor beton klakson mobil tingkat bising 95 hingga 100 desibel tara bising hasil mesin potong rumput hasil musik country bukti pengaruh perilaku predasi hama kepik musik rok raung kepik makan tengah jumlah kutu daun bagai mangsa lama periode sama tanam kedelai papar musik 25 persen lebih kecil tanam papar segera temu satu hari tengah hari 15 jam lebih subjek uji coba papar musik rok ac dc suara bising kota benar kurang jumlah kutu daun konsumsi jumlah besar kata barton tempat tanam kedelai bilik sendiri tempat tanam ruang kutu daun laku eksperimen sebut tanam kutu daun ikan dampak atas musik kata segera kepik masuk beda colok lihat predator makan hampir semua kutu daun lingkung musik rok suara bising kota predator konsumsi kutu daun jauh lebih sedikit satu sempat papar musik ac dc kepik sang predator kendali kutu daun pikir akhir 180 kutu per tanam rata kata barton banyak kutu daun satu tanam cukup benar rusak tanam jadi lihat kurang biomassa tanam ac dc cara langsung kurang produksi

kedelai tekan bukti barton kata percaya beda beberapa atribut musik buat lebih sulit kepik temu kutu daun getar bas timbul respons antipemangsa suara simulasi predator ganggu lingkung jadi henti makan mungkin habis lebih banyak waktu mangsa musik buat lebih sulit tangkap kutu kata barton kata tingkat predasi tingkat sama volume pelan barton kata teliti pertama tunjuk efek polusi suara seluruh ekosistem ganggu ubah satu sistem satu spesies pengaruh tiap spesies potensi pengaruh kepik khusus spesies harlequin mangsa utama kutu daun ketidakmampuan mangsa hama bunuh tanam tani semprot pestisida kimia jaga tanam tetap sehat rugi sehat manusia studi publikasi jurnal ecology and evolution dedikasi gitaris ac dc malcolm young tinggal makalah selesai studi baru tim teliti internasional ikan wawas baru atas neenk moyang duduk asia tenggara akibat buru bangsa romawi kuno kini spesies paus hilang daerah mediterania bangun percaya dorong kus sehat twitter kunci akun deteksi ubah dadak perilaku akun sebut',

'tuan rumah real madrid buat kejut tim tamu levante lanjut la liga spanyol tanding estadio santiago bernab u sabtu 20 10 2018 tuan rumah real madrid serah tim tamu levante skor 12 meski kalah real madrid benar menang kuasa bola levante 57 persen banding 43 persen real madrid cipta peluang lebih banyak levante cetak gol pertama lalu jos luis morales menit 6 kemudian roger mart tambah unggul lewat titik putih levante menit 13 akhir real madrid pecah buntu menit 72 lalu marcelo jadi gol tutup tanding kali skor 12 tanda kalah real madrid cara untun la liga adapun levante menang tiga posisi tujuh klasemen susun main real madrid thibaut courtois lvaro odrizola g bale 46 rapha l varane sergio ramos marcelo luka modri casemiro isco k benzema 60 lucas v zquez mariano d az marco asensio ceballos 60 cadang keylor navas nacho fern ndez sergio reguil n ton kroos ceballos gareth bale karim benzema levante oier olaz bal erick cabaco sergio postigo chema 75 r ber pier jason ruben rochina c doukour 67 enis bardhi jos campa a to o jos luis morales roger e boateng 79 cadang aitor pedro l pez chema cheick doukour moises simon raphael dwamena emmanuel boateng kartu e cabaco 53 jason 89 marcelo 90 4 catat redaksi teks gambar artikel publikasi cara otomatis program komputer ubah angka jadi narasi lapor otomatis hasil tanding robotorial akses kanal arena khusus topik hasil tanding tuan rumah real madrid buat kejut tim tamu levante lanjut la liga spanyol tanding estadio santiago bernab u sabtu 20 10 2018 tuan rumah real madrid serah tim tamu levante skor 12 meski kalah real madrid benar menang kuasa bola levante 57 persen banding 43 persen real madrid cipta peluang lebih banyak levante cetak gol pertama lalu jos luis morales menit 6 kemudian roger mart tambah unggul lewat titik putih levante menit 13 akhir real madrid pecah buntu menit 72 lalu marcelo jadi gol tutup tanding kali skor 12 tanda kalah real madrid cara untun la liga adapun levante menang tiga posisi tujuh klasemen susun main real madrid thibaut courtois lvaro odrizola g bale 46 rapha l varane sergio ramos marcelo luka modri casemiro isco k benzema 60 lucas v zquez mariano d az marco asensio ceballos 60 cadang keylor navas nacho fern ndez sergio reguil n ton kroos ceballos gareth bale karim benzema levante oier olaz bal erick cabaco sergio postigo chema 75 r ber pier jason ruben rochina c doukour 67 enis bardhi jos campa a to o jos luis morales roger e boateng 79 cadang aitor pedro l pez chema cheick doukour moises simon raphael dwamena emmanuel boateng kartu e cabaco 53 jason 89 marcelo 90 4 catat

redaksi teks gambar artikel publikasi cara otomatis program komputer ubah angka jadi narasi lapor otomatis hasil tanding robotorial akses kanal arena khusus topik hasil tanding valencia 1 1 legan s la liga 2018 2019 barcelona 4 2 sevilla la liga 2018 2019 udinese 0 3 napoli liga serie a 2018 2019',

'ekor harimau sumatera panthera tigris sumatrae temu jebak lorong ruco pasar camat pulau burung kabupaten indragiri hilir riau rabu 14 11 2018 pihak wenang upaya evakuasi informasi harimau kawasan pasar camat pulau burung tim rescue langsung gerak lokasi kata kepala bbksda riau suharyono kompas com Kamis 15 11 tim rescue diri tim masingmasing satu tim rengat kabupaten indragiri hulu satu tim pekanbaru tim rengat bawa obat bius kerah tim isi tugas medis bawa obat bius perlu evakuasi tim pertama diri lima orang tim jumlah lebih bawa bius kata suharyono kutip liputan6 turut suharyono jadi korban jiwa meski si kucing raksasa belang sebut masuk pasar pusat ekonomi utama masyarakat pulau burung cukup ramai rabu 15 11 pagi tim rescue pasang jaring kayu dinding sela ruco harimau nahas sebut lari diri camat pulau burung m yusuf kata harimau sebut tembak bius tembak bius tugas bbksda riau tim upaya evakuasi harimau lorong ruco kata yusuf kepala bidang tata usaha bbksda riau hartono go riau 15 11 kata laku sosialisasi ingat warga tetap waspada libat pihak aman tak warga dekat jadi korban jangan terlalu dekat hewan buas sebut penasaran mau lihat dekat bukan hewan jinak ujar hingga tahu mana harimau sebut asal apa sana suharyono duga harimau muncul rawa luas empat hektare dekat pasar tahu harimau mangsa beberapa ekor ternak warga dusun teluk nibung pulau burung beberapa waktu lau jelas beberapa waktu lalu lapor muncul ekor harimau lokasi dekat temu jejak beda induk individu kata suharyono lanjut turut suharyono tim laku orientasi lapang tindak perlu langkah ikut laku evakuasi masuk cek sehat harimau tadi laku observasi rehabilitasi stress hs perlu waktu ucap tribunnews 15 11 rehabilitasi selesai bbksda cari tempat huni ikut sesuai habitat harimau sebut beberapa lokasi sesuai habitat mungkin besar harimau datar rendah rawa cari alternatif sesuai habitat asli kata riau warga desa kute keramil camat linge isaq kabupaten aceh tengah aceh kejut hadir kawan harimau sumatera rabu 14 11 warga tahu temu ekor kerbau kondisi tubuh cabik sore hari dasar saksi sahrin warga tempat ekor kerbau milik makan harimau rabu dini hari satu induk satu anak induk sempat sembelih sedang anak kerbau habis makan warga lihat bekas tapak harimau capai tiga ekor kata sahrin harap pihak wenang segera tangan tak konflik lanjut makan korban kepala bksda aceh sapto aji prabowo benar lapor hadir harimau sumatera kampung warga aceh tengah sejak rabu malam tim bksda resort 5 wilayah aceh tengah laku monitor desa sebut laku patroli tenang warga jelas sapto aji orang pawang harimau laku patroli hutan desa kute keramil pantau harimau sebut sapto aji duga harimau masuk mukim duduk tempat hidup rusak akibat bala liar desak aparat tegak hukum segera proses hukum laku ilegal logging tekan potensi konflik satwa liar manusia turut lapor baru harimau sumatera aceh kira sisa 200 ekor langsung sapto aji diskusi satwa liar kantor aliansi jurnalis independen aji banda aceh rabu 14 11 turut kondisi picu kasus buru aktivitas ambah bala liar kisar 200 ekor 67 ekor sebar kawasan ulu masen kawasan leuser 130 ekor kira sebut sesuai hasil monitoring 2013 2015 kata tempo co 15 11 kumpul cara seluruh jumlah populasi harimau sumatera 600 ekor hingga indonesia 600 ekor kata ahli ekologi satwa liar lanskap world wide fund for nature wwf indonesia sunarto kabar bisnis 30 7 turut

data wwf global olah lokadata beritagar id indonesia negara tiga populasi harimau tinggi 371 ekor india nomor satu populasi harimau 2 226 ekor ikut rusia 433 ekor dagang kulit harimau jadi hingga 2018 temu kasus dagang kulit harimau masuk sidang konflik manusia harimau panjang tahun 20152018 catat sembilan kasus sebar kabupaten aceh selatan bireuen singkil subulusalam aceh besar gayo lues aceh tengah aceh jaya bener riah jadi temu unik arti ketidakstabilan populasi alam liar kawin silang antar spesies mudah beri perban lalu teknologi baru masyarakat kini tak perlu sambang bank bahkan anjung tunai mandiri atm transaksi jaring makin banyak main pubg bakal rilis versi sony ps4 7 desember 2018',

'kasus bunuh jurnalis arab saudi jamal khashoggi terus muncul agam isu spekulasi arab saudi aku khashoggi bunuh muncul isu khashoggi bunuh tahu saudi senjata kimia konflik yaman isu pakai senjata kimia arab saudi muncul lewat media asal inggris the sunday express tulis cenderung sumir konfirmasi kemudian kutip agam media masuk kantor berita turki anadolu agency tulis kutip kisah orang teman anonim khashoggi sumber cerita temu khashoggi pek tuju gedung konsulat saudi istanbul turki 2 oktober 2018 khashoggi lihat murung bahkan khawatir sumber tanya khawatir koresponden media amerika the washington post memberitahu kabar saudi senjata kimia yaman the express berita arab saudi kabar peluru isi fosfor putih produksi amerika serikat hadap warga sipil pasu yaman september lalu arab saudi pimpin intervensi hadap yaman sejak 2015 arab saudi sekutu kerah kuat militer henti upaya kudeta hadap pemerintah yaman laku militan hussein badreddin alhouthi arab saudi dukung kubu anggap bagai pemerintah sah bawah pimpin abdrabbuh mansour had khashoggi kata sumber maksud dokumen bukti kabar senjata kimia yaman pakai senjata kimia jadi sorot dunia besar 193 negara dunia buat konvensi larang pakai senjata kimia krisis suriah arab saudi pernah tuding presiden suriah bashar al assad senjata kimia balik yaman arab saudi dukung kubu guling presiden assad tak sekutu iran kini senjata kimia singgung mati khashoggi meski semua terang benar polisi turki arab saudi selidik arab saudi lalu media resmi pemerintah umum hasil lidi awal jumat 19 10 2018 lalu temu khashoggi orang temu konsulat libat tengkar kelahi tim selidik periksa 18 warga saudi duga libat 15 kabar datang istanbul temu khashoggi konsulat saudi copot beberapa jabat senior masuk saud alqahtani nasihat muhammad bin salman mayor jenderal ahmed alasiri orang jabat senior kantor intelijen tangkap alasiri benar duga kasus bunuh jerat pangeran muhammad bin salman populer sebut mbs lama kasus bunuh khashoggi kerap kait kuasa arab saudi menteri luar negeri arab saudi adel al jubeir tegas laku balik bunuh khashoggi hukum arab saudi meski investigasi butuh waktu lama masalah jadi histeria ujar menteri luar negeri adel aljubeir kutip voaindonesia ap kira orang lanjut salah arab saudi lidi selesai laku kasus bunuh khashoggi konsulat arab saudi istanbul memang jadi sorot dunia aksi solidaritas tuntutan kasus kasus bunuh khashoggi langsung bagai negara masuk indonesia presiden joko widodo pernah nyata kasus khashoggi tangan cara transparan seksama jokowi perhati hadap kasus khashoggi temu adel aljubeir istana bogor senin 22 10 2018 lalu total dana ganti rugi santun keluarga korban tinggal dunia akibat celaka pesawat lion air jt610 kira lebih rp3 miliar per orang broker joki cpns minta upah rp125 juta140 juta tahu joki blok umur hidup larang ikut tes cpns benua afrika lama memang terlalu lirik usaha indonesia bagai ladang bisnis banding kawasan eropa maupun as',

'raih sukses jadi impi besar orang nasihat kerja keras wujud sukses mungkin biasa dengar nyata cukup lifehack lansir kerja keras semata bawa pengaruh besar jalan tuju sukses balik jumlah faktor butuh raih sukses misal koneksi aku miliarder asal inggris richard branson turut sukses bisnis capai bantu orang tulis forbes jadi acu raih sukses perlu ajar tenang teliti tunjuk pikir kondisi prima stres tenang turut dr travis bradberry ceo talentsmart mampu kelola emosi tetap tenang bawah tekan milik pengaruh langsung kerja bukan omong kosong belaka talentsmart laku teliti hadap lebih satu juta orang hasil temu 90 persen orang sukses terampil kelola emosi stres tetap tenang kendali jadi bagaimana tetap tenang tekan gampang mulai laku 5 10 biasa bawah lansir lifehack tidur cukupanda milik cukup waktu tidur kualitas tidur nyenyak tidur sangat penting tubuh cipta pikir positif per hari tidur kurang 67 jam sehat produktivitas ganggu selalu pikir positifstres perlu larut stres kamus orang sukses kata serah kondisi apa selalu usaha sikap positif hadap masa sulit tetap efektif rencana baik masa depan hindar kafeinanda gemar seruput kopi saran mulai banyak minum air putih ganti kopi air putih teliti tunjuk minum kopi rangsang otak rasa lebih ekstra tenang cenderung buat mudah cemas khawatir tertawa bercandajika lebih banyak tertawa bantu lepas hormon timbul rasa tenang emosi lebih kontrol mulai sisip humor obrol selalu punya rencanaikuti cara orang sukses selalu milik rencana ingat rencana buat perlu sesuai mampu diri sendiri memang perlu tahu tren bukan arti ikut arus punya rencana bantu tetap organisir fokus capai tuju lebih mudah percaya diri buka jalan capai puncak hasil banyak orang sadar tumpuk picu stres kewalahan laku sederhana hidup tak perlu rumit habis waktu orang positif nyata penting sehat bahagia jaga sehat olahraga jalan gaya hidup sehat memang penting orang positif bantu lebih sehat cara fisik mental',

'tutup seluruh gerai 7eleven urai jumlah masalah main bisnis retail tanah air meski tak laku semua main apa alami 7eleven patut jadi ajar hadir gerai waralaba akrab sebut sevel memang sempat jadi lokasi nongkrong andal anak muda jakarta bagaimana sejak buka 2008 model bisnis gabung toko kelontong convenience store restoran cepat saji buka 24 jam sedia tempat duduk jaring internet gratis modal kacang kulit satu botol minum siapa bebas duduk obrol mulai matahari terbit hingga benam balik model bisnis jadi bumerang 7eleven ketua waralaba lisensi indonesia wali levita supit liputan6 com kata bangkrut 7eleven jadi lantaran konsumen lebih banyak habis waktu nongkrong belanja kalau lihat anak muda nongkrongnya lebih banyak belanja target usaha capai ujar levita senin 3 7 2017 akhir model bisnis 7eleven tiru main lawson hadir 2011 family mart hadir 2013 padahal baru hadir banyak sebut tenar 7eleven mampu saing main waralaba circlek lembaga peringkat kredit internasional fitch ratings pakat model bisnis aplikasi 7eleven indonesia solid bisnis jadi tak ukur model bisnis buat usaha gelontor biaya operasional besar associate director fitch ratings indonesia olly prayudi siar pers sebut profil bisnis cara signifikan beda minimarket alfamart indomaret beri tekan lebih besar belanja bahan makan jaring ritel tadi jauh lebih luas banding 7eleven bagai catat sejak buka hingga tutup per 30 juni 2017 gerai 7eleven jakarta kota besar rencana proyeksi ekspansi gagal wujud model bisnis 7eleven indonesia bilang beda banding negara singapura australia misal 7eleven sekadar minimarket umum harga sama malaysia beberapa gerai 7eleven mulai terap konsep sama indonesia kursi meja sedia tak banyak gerai jakarta meski olly pakat

faktor regulasi terap indonesia jadi sandungan besar bisnis gerai kelola pt modern sevel indonesia anak pt modern international tbk salah satu regulasi kurang untung larang jual minum alkohol minimarket lalu atur menteri dagang permendag no 06 mdag per 1 2015 kendali awas hadap peredaran jual minum alkohol rerata tahun 30 60 gerai 7eleven baru buka jakarta buat jumlah gerai 7eleven terus tambah tahun 2011 50an gerai tahun 2012 jumlah tambah hampir kali lipat tahun 2014 jumlah gerai 7eleven jakarta capai 190 tahun banyak 40 gerai baru buka jual bersih naik 24 5 persen jadi rp971 7 miliar tahun rp778 3 miliar tahun sebut bagai puncak jaya 7eleven tahun ikut jual turun jumlah gerai tahun 2015 total jual bersih turun jadi rp886 84 miliar tahun 7eleven pertama kali laku tutup 20 gerai menyebut model bisnis sebut soal regulasi tak jelas kendati gerai tutup matahari tetap laku ekspansi buka tiga gerai lokasi beda meski jarang rilis ponsel murah pasar oppo indonesia jalan cukup baik hasil survei nyata tunjuk harga bukan faktor utama beli']

Langkah 4: Ekstraksi Fitur Langkah keempat yang harus dilakukan adalah melakukan ekstraksi fitur terhadap hasil normalisasi data yang akan digunakan untuk keperluan pengembangan model klasifikasi teks.

[15]: *#Fungsi untuk mengekstraksi feature menggunakan TF-IDF Model*

```
def tfidf_transformer(bow_matrix):
    transformer = TfidfTransformer(norm='l2',
                                   smooth_idf=True,
                                   use_idf=True)
    tfidf_matrix = transformer.fit_transform(bow_matrix)
    return transformer, tfidf_matrix

def tfidf_extractor(corpus, ngram_range=(1,1)):
    vectorizer = TfidfVectorizer(min_df=1,
                                 norm='l2',
                                 smooth_idf=True,
                                 use_idf=True,
                                 ngram_range=ngram_range)
    features = vectorizer.fit_transform(corpus)
    return vectorizer, features
```

[16]: *#menggunakan model TF-IDF untuk mengekstraksi feature*

```
tfidf_vectorizer, tfidf_train_features = tfidf_extractor(norm_train_corpus)
tfidf_test_features = tfidf_vectorizer.transform(norm_test_corpus)
```

Langkah 5: Model Training Langkah kelima yang harus dilakukan adalah melakukan training model klasifikasi teks.

[17]:

```
#kami mengevaluasi model menggunakan fungsi train_predict_evaluate_model
#kami mengukur nilai accuracy,precision, recall, dan F1 score apakah nilainya
↳ sudah sesuai atau tidak. Jika nilainya mendekati 1 maka proses klasifikasi
↳ berhasil
```

```
from sklearn.linear_model import SGDClassifier

train_features=tfidf_train_features
train_labels=train_labels
test_features=tfidf_test_features
test_labels=test_labels
```

```
[18]: df_train = pd.DataFrame()
df_train['norm_train'] = norm_train_corpus
df_train['train_labels'] = train_labels
df_test = pd.DataFrame()
df_test['norm_test'] = norm_test_corpus
df_test['test_labels'] = test_labels
```

```
[19]: df_train.head()
```

```
[19]:
```

	norm_train	train_labels
0	gonjanganjing soal siapa latih tim nasional i...	ekonomi
1	kotak hitam baru temu kini semua utama keluarg...	ekonomi
2	hiu prasejarah megalodon sedang hangat bincang...	ekonomi
3	persib bandung milik peluang besar juara liga ...	NaN
4	sejak sutradara ngenest 2015 bagai film panjan...	ekonomi

```
[20]: df_test.head()
```

```
[20]:
```

	norm_test	test_labels
0	mata rupa jendela jiwa manfaat artificial inte...	NaN
1	coldplay rilis buah dokumenter judul a head fu...	NaN
2	nakita id sarita abdul mukti faisal harris res...	NaN
3	akhir lancong makin gari banyak kalang banyak ...	ekonomi
4	musik nyata dampak manusia buah studi baru ung...	NaN

```
[21]: train_features[0:1]
```

```
[21]: <1x23276 sparse matrix of type '<class 'numpy.float64'>'
      with 314 stored elements in Compressed Sparse Row format>
```

```
[22]: train_labels[0:1]
```

```
[22]: 376      sport
      Name: Jenis, dtype: object
```



```
[23]: from sklearn.svm import LinearSVC
      from sklearn.naive_bayes import MultinomialNB
```

```
[165]: clf = SGDClassifier(loss='hinge', max_iter=100)
      #clf = LinearSVC(penalty='l2', loss='squared_hinge', max_iter=100)
      #clf = MultinomialNB()
```

```
[166]: clf.fit(train_features, train_labels)
      predictions = clf.predict(test_features)

      print("Label test: \n" + format(test_labels[:10]))
      print("Prediction test: \n"+ format(predictions[:10]))
```

Label test:

```
1148      sains-tekno
1049      entertainment
982      entertainment
808      gaya hidup
1195      sains-tekno
240      sport
1118      sains-tekno
596      news
924      gaya hidup
65      ekonomi
```

Name: Jenis, dtype: object

Prediction test:

```
['sains-tekno' 'entertainment' 'news' 'gaya hidup' 'sains-tekno' 'sport'
 'sains-tekno' 'news' 'gaya hidup' 'ekonomi']
```

```
[167]: print("Accuraction: "+format(clf.score(test_features,test_labels)))
```

Accuraction: 0.9236111111111112

```
[168]: def get_metrics(true_labels, predicted_labels):
      print('Accuracy: ', np.round(metrics.accuracy_score(true_labels,
                                                             predicted_labels),2))

      print('Precision: ', np.round(metrics.precision_score(true_labels,
                                                             predicted_labels,
                                                             average='weighted'),2))

      print('Recall: ', np.round(metrics.recall_score(true_labels,
                                                         predicted_labels,
                                                         average='weighted'),2))

      print('F1 Score: ', np.round(metrics.f1_score(true_labels,
                                                       predicted_labels,
                                                       average='weighted'),2))
```

```
# mengevaluasi performa prediksi menggunakan model
```

```
get_metrics(true_labels=test_labels,predicted_labels=predictions)
```

Accuracy: 0.92

Precision: 0.93

Recall: 0.92

F1 Score: 0.92

```
[169]: test = ('Sebuah kapsul NASA mendarat di gurun Utah, AS, pada Minggu (24/9 waktu_
↪setempat. Kapsul tersebut berisi sampel dari asteroid Benny yang dikumpulkan_
↪oleh wahana antariksa OSIRIS-REx milik NASA pada Oktober 2020 lalu.' +
'SOSIRIS-REx telah menghabiskan waktu selama 7 tahun di luar angkasa untuk_
↪menunggu momen yang tepat mengambil potongan asteroid Bennu. Kini, sampel_
↪tersebut berhasil dibawa OSIRIS-REx ke Bumi, mendarat dengan selamat di_
↪Tempat Uji dan Pelatihan Angkatan Udara AS di Utah pada Minggu (24/4) pukul_
↪08.52 waktu setempat.')
```

```
[170]: test_eko = ('PT Bank Mandiri (Persero) Tbk (BMRI) melepas seluruh kepemilikan_
↪sahamnya di PT AXA Insurance Indonesia. Meski tak lagi punya saham, Bank_
↪Mandiri tidak akan mengakhiri kerja sama bancassurance dengan AXA Insurance_
↪Indonesia.' +
'Secara rinci, sisa 138.000 saham milik Bank Mandiri dilepas ke dua_
↪pemegang saham, yakni Anil Panjwani dan Manoj Ramkrashin Tolani._
↪Masing-masing mendapat 69.000 saham atau sekitar 10 persen dari modal_
↪ditempatkan dan disetor milik AXA Insurance Indonesia.')
```

```
[171]: test_or = ('PSSI telah mengumumkan skuad Timnas Indonesia yang akan berliga_
↪dalam Kualifikasi Piala Dunia 2026 melawan Brunei Darussalam. Dalam daftar_
↪25 pemain tersebut, terdapat satu nama yang absen, yaitu Ivar Jenner, pemain_
↪Jong FC Utrecht.'+
'Keberhasilan Timnas Indonesia melawan Brunei Darussalam dalam dua_
↪leg pertandingan ini akan membawa mereka melangkah ke babak kedua_
↪Kualifikasi Piala Dunia 2026. Tim Indonesia akan bersaing di Grup F babak_
↪kedua bersama dengan Irak, Vietnam, dan Filipina.')
```

```
[172]: test_enter = ('Setelah berkompetisi dalam Busan International Film Festival_
↪(BIFF) 2023, kini film bergenre detektif "24 Jam Bersama Gaspar" dirilis_
↪trailer resminya. Trailer tersebut sempat dikeluarkan oleh akun YouTube_
↪resmi Visinema @Visinema Pictures pada Selasa (26/9/2023) kemarin. Di dalam_
↪trailer tersebut, tokoh Gaspar diperankan oleh aktor Reza Rahadian.'+
'Melalui sinopsis resmi film 24 Jam Bersama Gaspar, tokoh bernama_
↪"Gaspar" merupakan seorang detektif partikelir yang akan dinyatakan mati_
↪dalam tenggat waktu 24 jam. Dalam kurun waktu itu, Gaspar berupaya keras_
↪untuk membalaskan dendam dan merampok sebuah toko perhiasan yang ia incar._
↪Keinginannya itu ditimbulkan berkat munculnya petunjuk mengenai hilangnya_
↪Kirana, seorang gadis yang menjadi teman masa kecil Gaspar.')
```

```
[173]: test_gaya = ('Berbagai modus pinjaman pribadi atau pinpri menjamur di media_
↳sosal. Tak sedikit masyarakat yang lantas memilih pinpri sebagai alternatif_
↳pembiayaan karena prosesnya cepat. Perencana Keuangan Nadia Isnuari Harsya_
↳mengatakan, kebanyakan orang yang terlilit pinjaman pribadi atau pinpri_
↳adalah mereka yang terjebak dengan gaya hidup.'+
      'Ia menekankan, seseorang yang telah berani mengambil pinjaman_
↳berarti harus mengembalikan pinjaman tersebut. Selain itu, masyarakat_
↳dinilai belum dapat meminjam secara pintar. Beberapa peminjam hanya_
↳memikirkan kesenangan sesaat saja tanpa berpikir soal konsekuensinya.')
```

```
[174]: test_gh = ('Kalangan Gen Z dan Milenial tumbuh di lingkungan dengan kemajuan_
↳teknologi dalam kehidupan sehari-hari. Hal ini pun turut melahirkan_
↳kebiasaan baru, khususnya dalam berbelanja dan memilih metode pembayaran_
↳tanpa uang tunai atau disebut gaya hidup cashless.'+
      'Selain itu, kenyamanan, kecepatan, dan keamanan yang ditawarkan pun_
↳menjadi faktor-faktor terbesar yang mendorong metode pembayaran digital._
↳Apalagi di era smartphone dan internet, kemudahan pembayaran digital_
↳memiliki daya tarik tersendiri, baik itu melalui dompet digital, QR, hingga_
↳kartu kredit contactless.','gaya-hidup')
```

```
[175]: line = tfidf_vectorizer.transform([test_gaya])
line.shape
```

```
[175]: (1, 23276)
```

```
[176]: line
```

```
[176]: <1x23276 sparse matrix of type '<class 'numpy.float64'>'
      with 22 stored elements in Compressed Sparse Row format>
```

```
[177]: test = clf.predict(line)
```

```
[178]: test
```

```
[178]: array(['gaya hidup'], dtype='<U13')
```

```
[ ]:
```