

(eyError KeyError KeyError Anda Harap jelaskan errornya: KeyError: 'tweet' Gemini Hypothesis The error is caused by the line data tweet = list(data['tweet']) in the last cell. The error message KeyError: 'tweet' indicates that the key 'tweet' is not present in the data DataFrame. Suggested Changes You can fix this error by using the correct key to access the 'tweet' column. Since the DataFrame contains only two columns, you can access the 'tweet' column using the index 1. 囗 data\_tweet = list(data[1]) Use code with caution Alternatively, you can use the iloc method to access the 'tweet' column by its index: data\_tweet = list(data.iloc[:, 1]) Use code with caution Here's the corrected code: 囗 # Labeling with Naive Bayes data\_tweet = list(data[1]) polaritas = 0

status = []

else:

total\_positif = total\_negatif = total\_netr

analysis = TextBlob(tweet, classifier=

if analysis.classify() == 'positif':

elif analysis.classify() == 'netral':

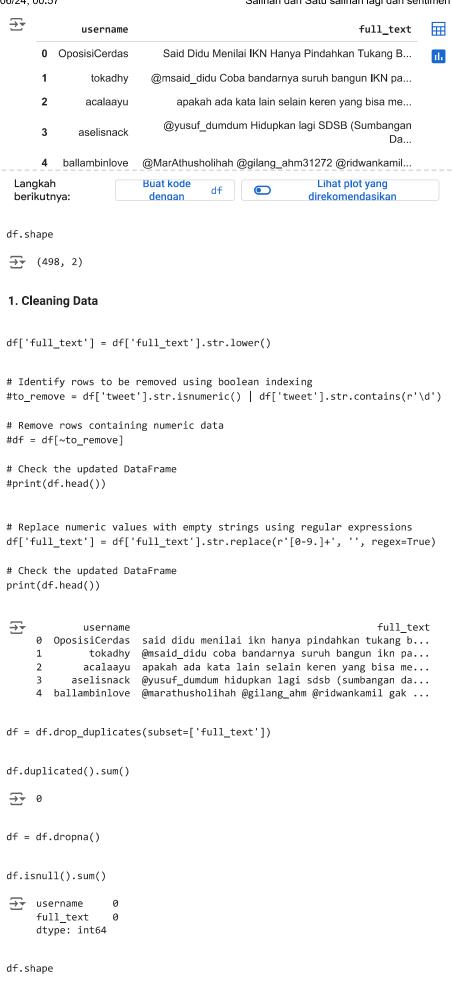
status.append(analysis.classify())

for i, tweet in enumerate(full\_text):

total positif += 1

total netral += 1

total\_negatif += 1



```
total += 1

print(f'\nHasil Analisis Data:\nPositif = print(f'Total Data : {total}')

Use code with caution

Sumber

1. aldi6889/data-dashboards
2.

Berikan penilaian untuk jawaban ini
```

```
→• (493, 2)
def clean_twitter_text(text):
    text = re.sub(r'@[A-Za-z0-9_]+', '', text) # hapus mention
    text = re.sub(r'#\w+', '', text) # hapus hastag
text = re.sub(r'RT[\s]+', '', text) # hapus Retweet
    text = re.sub(r'https?://\S+', '', text) # hapus url
    text = re.sub(r'[^A-Za-z0-9\s]', '', text) # hapus karakter non alpha
    text = re.sub(r'\s+', ' ', text).strip() # hapus spasi berlebih
    return text
df['full_text'] = df['full_text'].apply(clean_twitter_text)
df
\rightarrow
                                                                      full_text
                                                                                    扁
                    username
        0
               OposisiCerdas
                                  said didu menilai ikn hanya pindahkan tukang b...
        1
                      tokadhy
                                coba bandarnya suruh bangun ikn pasti lolosama...
        2
                     acalaayu
                                  apakah ada kata lain selain keren yang bisa me...
        3
                    aselisnack
                                hidupkan lagi sdsb sumbangan dana sosial berha...
        4
                ballambinlove
                                     gak salah nih data dari ppatk di kutip duit se...
       493
                posaktualcom
                                  pesimis pada ikn nusantara rocky gerung soroti...
       494
                  sidiq rajata
                               cctv pintar akan membantu ikn nusantara mengat...
       495
                ibun dabukke
                                    hut ri di ikn akan lebih spesial dengan lapang...
       496
                 MT93990025
                                    reatoran dibangun apa tak bakal rugikrn ikn it...
       497 AinurRo08452526 ikn bisa saja akan menjadi monumen proyek yg g...
     493 rows × 2 columns
 Langkah
                                                              Linat plot yang
 berikutnya:
                            dengan
                                                            direkomendasikan
2. Preprocessing Data
```

- normalisasi
- stopword
- Tokenize
- Stemming

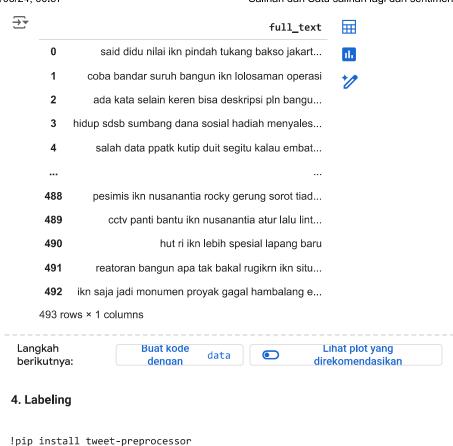
## #normalisasi

```
'anjg':'anjing', 'org2':'orang-orang', 'yack':'ya', 'kali':'sekali',
         'bnyak': 'banyak', 'rame2': 'ramai-ramai', 'aneh2': 'aneh-aneh', 'oenge
def normalisasi(str_text):
  for i in norm:
    str_text = str_text.replace(i,norm[i])
  return str_text
df['full_text'] = df['full_text'].apply(lambda x: normalisasi(x))
₹
                                                                              H
                   username
                                                                 full_text
       0
              OposisiCerdas
                               said didu menilai ikn hanya pindahkan tukang b...
       1
                    tokadhy
                              coba bandarnya suruh bangun ikn pasti lolosama...
       2
                   acalaayu
                               apakah ada kata lain selain keren yang bisa me...
       3
                  aselisnack
                             hidupkan lagi sdsb sumbangan dana sosial berha...
       4
               ballambinlove
                                   gak salah ini data dari ppatk di kutip duit se...
      493
                               pesimis pada ikn nusanantia rocky gerung sorot...
               posaktualcom
      494
                 sidiq_rajata
                             cctv pinanti akan membantu ikn nusanantia meng...
      495
               ibun_dabukke
                                 hut ri di ikn akan lebih spesial dengan lapang...
      496
                MT93990025
                                 reatoran dibangun apa tak bakal rugikrn ikn it...
                                   ikn bisa saja akan menjadi monumen proyak
      497 AinurRo08452526
                                                                    yang...
     493 rows × 2 columns
 Langkah
                         Buat kode
                                                         Lihat plot yang
                                             df
                                                       direkomendasikan
 berikutnya:
                          dengan
!pip install Sastrawi
     Collecting Sastrawi
       Downloading Sastrawi-1.0.1-py2.py3-none-any.whl (209 kB)
                                                      = 209.7/209.7 kB 3.8 MB/s (
     Installing collected packages: Sastrawi
     Successfully installed Sastrawi-1.0.1
#stopword
import Sastrawi
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemover
more_stop_words = ["yang", "yg", "gak", "ngisi", "udah", "d", "sih", "nya",
stop_words = StopWordRemoverFactory().get_stop_words()
stop_words.extend(more_stop_words)
new_array = ArrayDictionary(stop_words)
stop_words_remover_new = StopWordRemover(new_array)
def stopword(str_text):
  str_text = stop_words_remover_new.remove(str_text)
  return str_text
df['full_text'] = df['full_text'].apply(lambda x: stopword(x))
df.head()
```

```
₹
                                                            full_text
                                                                         丽
              username
                            said didu menilai ikn pindahkan tukang bakso j...
      OposisiCerdas
      1
                        coba bandarnya suruh bangun ikn lolosaman bero...
               tokadhy
      2
              acalaayu
                           ada kata selain keren bisa mendeskripsikan pln...
      3
             aselisnack hidupkan sdsb sumbangan dana sosial berhadiah ...
          ballambinlove
                            salah data ppatk kutip duit segitu kalau embat...
 Langkah
                         Buat kode
                                                         Lihat plot yang
 berikutnya:
                          dengan
                                                       direkomendasikan
#tokenize
tokenized = df['full text'].apply(lambda x:x.split())
tokenized
\overline{\Rightarrow}
     0
             [said, didu, menilai, ikn, pindahkan, tukang, ...
     1
             [coba, bandarnya, suruh, bangun, ikn, lolosama...
     2
             [ada, kata, selain, keren, bisa, mendeskripsik...
     3
             [hidupkan, sdsb, sumbangan, dana, sosial, berh...
             [salah, data, ppatk, kutip, duit, segitu, kala...
     493
             [pesimis, ikn, nusanantia, rocky, gerung, soro...
     494
             [cctv, pinanti, membantu, ikn, nusanantia, men...
     495
                [hut, ri, ikn, lebih, spesial, lapangan, baru]
     496
             [reatoran, dibangun, apa, tak, bakal, rugikrn,...
     497
             [ikn, saja, menjadi, monumen, proyak, gagal, h...
     Name: full_text, Length: 493, dtype: object
#stemming ( mengubah kata imbuhan menjadi kata dasar, contoh : dimana menjad
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
def stemming(text_cleaning):
  factory = StemmerFactory()
  stemmer = factory.create_stemmer()
  do = []
  for w in text_cleaning:
    dt = stemmer.stem(w)
    do.append(dt)
  d_{clean} = []
  d_clean = " ".join(do)
  print(d_clean)
  return d_clean
tokenized = tokenized.apply(stemming)
\overline{2}
```

```
progres bangun bandara ikn nusanantia terus jalan
     tabir ikn lihat video saribulih
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     progres bangun fasilitas umum ikn nunsatara terus jalan
     apa di banggain ikn cuma buat rumah jabat
     valid kalau jadi cadang just in case jakarta jatuh ibukota harus teta
     ikn proyak bancak jokowi kroni lihat endingnya akan tambah masalah r€
     kontruksi lapang upacara ikn ester joko anwar kaesang arafah ipar mau
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     nanti kawasan ikn nusanantia untuk kendara listrik electric vehicle (
     regime makin otoriter kritik dg olok baperan langsung lapor polisi a
     terima kasih motorola teknologi sangat butuh ikn nusanantia
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     buka merdeka sama buka thn bangun indonesia sentris buka ikn unt pera
     pentidak ada ev ikn nusanantia tak asal perintah juga pihak swasta
     pentidak ada ev ikn nusanantia tak asal perintah juga pihak swasta
     hitung emisi ikn nusanantia lebih cepat atas masalah polusi
     bukan cuma ikn sebut juga adil ekologiiklim ham tidak masuk sama seka
     cocok ikn hrs semua aspek trlibat amp libat sistematis cermat ukur sį
     lapang upacara ikn hampir siap luar biasa
     luar biasa motorola bawa teknologi canggih ikn nusanantia
     pantes perintah getol bikin tapera mau ngasih sanksi kalau kerja ga r
     upadate bangun ikn ester joko anwar kaesang arafah ipar maut congratu
     harus minta maaf kan dak sebut pulau sekalimantan dia celoteh ikn ko
     teknologi aman motorola sangat bantu ikn nusanantia
     ikn jadi simbol jaya jokowiinvestor antri daftar
     mantab om buleikn ibukota koruptor nepotisme mulai tingkat lurah cama
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     progres bangun bandara ikn nusanantia terus jalan ini landas pacu ru
     pesimis ikn nusanantia rocky gerung sorot tiada investor ambisi poli
     cctv panti bantu ikn nusanantia atur lalu lintas lebih baik
     hut ri ikn lebih spesial lapang baru
     reatoran bangun apa tak bakal rugikrn ikn situasi hutan paling bakal
     ikn saia iadi monumen provak gagal hambalang efek akan koruptor tang
tokenized.to_csv("/content/ikndatatokenize.csv", index=False)
data = pd.read_csv("/content/ikndatatokenize.csv", encoding = 'latin1')
data.head()
\rightarrow
                                           full_text
      0
             said didu nilai ikn pindah tukang bakso jakart...
      1
           coba bandar suruh bangun ikn lolosaman operasi
      2
            ada kata selain keren bisa deskripsi pln bangu...
      3 hidup sdsb sumbang dana sosial hadiah menyales...
      4
             salah data ppatk kutip duit segitu kalau embat...
 Langkah
                       Buat kode
                                                       Lihat plot yang
                                  data
                                            direkomendasikan
 berikutnya:
                        dengan
data.isnull().sum()
     full text
     dtype: int64
data = data.dropna()
data
```

!pip install textblob
!pip install wordcloud



```
!pip install nltk
    Collecting tweet-preprocessor
       Downloading tweet preprocessor-0.6.0-py3-none-any.whl (27 kB)
     Installing collected packages: tweet-preprocessor
    Successfully installed tweet-preprocessor-0.6.0
    Requirement already satisfied: textblob in /usr/local/lib/python3.10/di
     Requirement already satisfied: nltk>=3.1 in /usr/local/lib/python3.10/c
    Requirement already satisfied: click in /usr/local/lib/python3.10/dist-
    Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist
    Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/pythor
    Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-r
    Requirement already satisfied: wordcloud in /usr/local/lib/python3.10/c
    Requirement already satisfied: numpy>=1.6.1 in /usr/local/lib/python3.1
    Requirement already satisfied: pillow in /usr/local/lib/python3.10/dist
     Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/
    Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/pythc
    Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.1
    Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/pyth
    Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/pyth
    Requirement already satisfied: packaging>=20.0 in /usr/local/lib/pythor
    Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/pythc
    Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/r
    Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/di
```

Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-Requirement already satisfied: click in /usr/local/lib/python3.10/dist-Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-requirement satisfied: tqdm in /usr/local/lib/python3.

```
import preprocessor as p
from textblob import TextBlob
import nltk
from nltk.stem import PorterStemmer
from nltk.tokenize import word_tokenize
nltk.download('punkt')
→ [nltk_data] Downloading package punkt to /root/nltk_data...
     [nltk_data] Unzipping tokenizers/punkt.zip.
from textblob import TextBlob
data tweet = list(data['full text'])
polaritas = 0
status = []
total_positif = total_negatif = total_netral = total = 0
non_string_data = [] # Daftar untuk menyimpan data yang bukan string
for i, tweet in enumerate(data tweet):
    # Pastikan tweet adalah string
    if isinstance(tweet, str):
       # Corrected line:
        analysis = TextBlob(tweet)
       polaritas += analysis.polarity
        if analysis.sentiment.polarity > 0.0:
            total_positif += 1
            status.append('positif')
        elif analysis.sentiment.polarity == 0.0:
            total_netral += 1
            status.append('netral')
        else:
            total_negatif += 1
            status.append('negatif')
       total += 1
    else:
        # Jika tweet bukan string, tambahkan ke daftar non_string_data
        non_string_data.append((i, full_text))
print(f'hasil analisis data:\nPositif = {total_positif}\nNetral = {total_net
print(f'\nTotal Data: {total}')
# Tampilkan data yang bukan string
print("\nData yang bukan string:")
for index, value in non_string_data:
    print(f'Index {index}: {value}')
    hasil analisis data:
     Positif = 28
     Netral = 457
     Negatif = 8
     Total Data: 493
     Data yang bukan string:
data.isnull().sum()
    full_text
     dtype: int64
#tampilkan data
data['klasifikasi'] = status
```

data **₹** full\_text klasifikasi 0 said didu nilai ikn pindah tukang bakso jakart... netral 1 coba bandar suruh bangun ikn lolosaman operasi netral 2 ada kata selain keren bisa deskripsi pln bangu... netral 3 hidup sdsb sumbang dana sosial hadiah menyales... netral 4 salah data ppatk kutip duit segitu kalau embat... netral ... ... 488 pesimis ikn nusanantia rocky gerung sorot tiad... netral 489 cctv panti bantu ikn nusanantia atur lalu lint... netral 490 hut ri ikn lebih spesial lapang baru netral 491 reatoran bangun apa tak bakal rugikrn ikn situ... netral 492 ikn saja jadi monumen proyak gagal hambalang e... netral 493 rows × 2 columns Langkah Buat kode Lihat plot yang data berikutnya: dengan direkomendasikan

## 5. Visualisasi

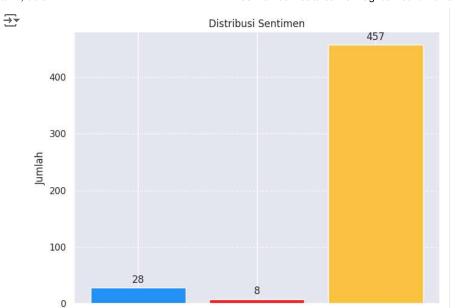
```
# Corrected code
from wordcloud import WordCloud, STOPWORDS
def plot_cloud(wordcloud):
 plt.figure(figsize=(10,8))
 plt.imshow(wordcloud, interpolation='bilinear')
 plt.axis('off')
 plt.show()
all_words = ' '.join([tweets for tweets in data['full_text']])
wordcloud = WordCloud(
    width = 3000,
    height = 2000,
    random_state = 3,
    background_color='black',
    colormap="Blues_r",
    collocations = False,
    stopwords=STOPWORDS
).generate(all_words)
plot_cloud(wordcloud)
```





```
#visualisasi Sentimen
import seaborn as sns
sns.set_theme()
labels = ['Positif', 'Negatif', 'Netral']
counts = [total_positif, total_negatif, total_netral]
def show_bar_chart(labels,counts, title):
 fig, ax = plt.subplots(figsize=(8,6))
 bars = ax.bar(labels,counts,color=['#2394f7','#f72323','#fac343'])
 for bar, count in zip(bars, counts):
      height = bar.get_height()
      ax.annotate(f'{count}', xy=(bar.get_x() + bar.get_width()/2, height),
                  xytext=(0,3),
                  textcoords='offset points',
                  ha='center',va='bottom')
  ax.grid(axis='y', linestyle='--',alpha=0.7)
  ax.set_xlabel('Sentimen')
  ax.set_ylabel('Jumlah')
 ax.set_title(title)
 plt.show()
show_bar_chart(labels,counts,'Distribusi Sentimen')
```

Netral



Negatif

Sentimen

## 6. Klasifikasi Sentimen Menggunaakan Naive Bayes Clasifier

Positif

#dataset = data.drop(['tweet'], axis=1, inplace=False)
dataset=data
dataset = [tuple(x) for x in dataset.to\_records(index=False)]

dataset

 $\overline{2}$ 

```
('menteri investasi bahlil sebut soal investor asing ikn apa kata
     dulu sekarang',
       'netral'),
      ('presiden republik indonesia joko widodo laku leta batu pertama
     groundbreaking bangun pusat ekosistem transisi energi layan digital
     pertama besar tanah air pln hub jantung ibu kota negara ikn
     nusanantia rabu',
       'netral'),
      ('ngebolehin siapa tiap masa tiap bijak pro kontra g jamin indo
     kaya bijak ikn akan yang kontra',
       'netral'),
      ('teknologi panti akan tingkat nyaman hidup ikn nusanantia',
     'netral'),
      ('erick thohir ajar bangun pusat uang ikn uea', 'netral'),
      ('proyak ikn nusanantia cipta lapang kerja signifikan indonesia',
     'netral'),
      ('bulan mau pindah ikn agustus percaya', 'netral'),
      ('jubir oikn pasti video wna sebut ikn kota koruptor hoaks kaesang
     ipar maut arafah cerai login',
       'netral'),
      ('electrizen pindah ibu kota negara ikn bukan dar diri gedung
     perintah merupakaan transformasi indonesia wujud adab lebih maju',
       'netral'),
      ('kalau kbanyakan investor asing tar kacung yaman anak abah teriak
     lagiikn cuma aseng cukong dllhadehhhh',
       'netral'),
      ('investor asing ikn masuk klaster dua kaesang ipar maut arafah
     ester cerai',
#train data
import random
set positif = []
set negatif = []
set_netral = []
for n in dataset:
  if(n[1] == 'Positif'):
    set_positif.append(n)
  elif(n[1] == 'Negatif'):
    set_negatif.append(n)
    set_netral.append(n)
set_positif = random.sample(set_positif, k=int(len(set_positif)/2))
set_negatif = random.sample(set_negatif, k=int(len(set_negatif)/2))
set netral = random.sample(set netral, k=int(len(set netral)/2))
train = set positif + set negatif + set netral
train set = []
for n in train:
 train_set.append(n)
from textblob.classifiers import NaiveBayesClassifier
cl = NaiveBayesClassifier(train_set)
print("Akuran Test:", cl.accuracy(dataset))
Akuran Test: 0.9655172413793104
```

```
# Labeling with Naive Bayes
data_tweet = list(data['full_text'])
polaritas = 0
status = []
total_positif = total_negatif = total_netral = total = 0
for i, tweet in enumerate(data_tweet):
    analysis = TextBlob(tweet, classifier=cl)
    if analysis.classify() == 'positif':
      total_positif += 1
    elif analysis.classify() == 'netral':
      total_netral += 1
    else:
      total_negatif += 1
    status.append(analysis.classify())
    total += 1
print(f'\nHasil Analisis Data:\nPositif = {total_positif}\nNetral = {total_
print(f'Total Data : {total}')
\rightarrow
     Hasil Analisis Data:
     Positif = 15
     Netral = 478
     Negatif = 0
     Total Data: 493
status = pd.DataFrame({'Klasifikasi Bayes:': status})
data['klasifikasi_bayes'] = status
#visualisasi Sentimen
import seaborn as sns
sns.set_theme()
labels = ['Positif', 'Negatif', 'Netral']
counts = [total_positif, total_negatif, total_netral]
def show_bar_chart(labels,counts, title):
 fig, ax = plt.subplots(figsize=(8,6))
 bars = ax.bar(labels,counts,color=['#2394f7','#f72323','#fac343'])
  for bar, count in zip(bars, counts):
      height = bar.get_height()
      ax.annotate(f'{count}', xy=(bar.get_x() + bar.get_width()/2, height),
                  xytext=(0,3),
                  textcoords='offset points',
                  ha='center',va='bottom')
  ax.grid(axis='y', linestyle='--',alpha=0.7)
  ax.set xlabel('Sentimen')
  ax.set_ylabel('Jumlah')
  ax.set_title(title)
 plt.show()
show_bar_chart(labels,counts,'Distribusi Sentimen Bayes')
```

