

ARIF FRIMA ARI SUWADJI - 221011700443 - 03SIFE003

INPUT	OUTPUT
pwd	<pre>In [5]: pwd Out[5]: 'C:\\Users\\arifs.DESKTOP-EQJN52\\Documents\\unpam\\Semester3\\Pengantar_Big_Data\\KelasBigData'</pre>
import urllib import pandas as pd from bs4 import BeautifulSoup as bs with urllib.request.urlopen("https://jalantikus.com/tips/aplikasi-chatting/") as page: html = page.read() print(html)	<pre>In [1]: import urllib import pandas as pd from bs4 import BeautifulSoup as bs In [4]: with urllib.request.urlopen("https://jalantikus.com/tips/aplikasi-chatting/") as page: html = page.read() print(html) b<!doctype html><html lang="id"><head><meta charset="utf-8"><meta name="language" content="Indonesian"><meta name="viewport" content="width=device-width,initial-scale=1,maximum-scale=1,user-scalable=no,minimum-scale=1"><meta name="geo.region" content="ID-JK"><meta name="geo.placename" content="Jakarta"><meta name="google-site-verification" content="Cn_w4FueEW9fXZdrf3LbVU5mUilJ59PmsFC6jKeP0eI"><meta name="google-site-verification" content="wgw-epNGIIX_ys2V70Jy6ypAYnoZF1SHZ-gkw2YowBw"><meta property="fb:app_id" content="391633060859123"><meta property="fb:pages" content="387823967920385"><meta name="dailymotion-domain-verification" content="dmexhibfwg9x7fyuz"><meta name="dailymotion-domain-verification" content="dmehiqbesixystf2"><title data-rh="true">10 Aplikasi Chatting Gratis Terbaik dan Terbaru 2023 Selain WhatsApp JalanTikus</title><meta data-rh="true" name="description" content="Sudah menjadi alat yang sangat penting dalam menjaga hubungan sosial kita, kamu wajib tahu rekomendasi aplikasi chatting terbaik 2023 di sini."/><meta data-rh="true" name="robots" content="index,follow,nooodp,noyid"><meta data-rh="true" property="og:site_name" content="JalanTikus"><meta data-rh="true" property="og:url" content="https://jalantikus.com/tips/aplikasi-chatting/"><meta data-rh="true" property="og:type" content="article"><meta data-rh="true" property="og:title" content="Rekomendasi 10 Aplikasi Chatting Gratis Terbaik 2023, Ngobrol Digital Makin Seru!"><meta data-rh="true" property="og:description" content="Sudah menjadi alat yang sangat penting dalam menjaga hubungan sosial kita, kamu wajib tahu rekomendasi aplikasi chatting terbaik 2023 di sini."/><meta data-rh="true" property="og:image" content="https://assets.jalantikus.com/assets/cache/769/330/userfiles/2023/07/14/Aplikasi-chatting-terbaik-7eaf6.jpg"><meta data-rh="true" property="og:image:width" content="600"><meta data-rh="true" property="og:image:height" content="315"><meta data-rh="true" name="twitter:card" content="summary_large_image"><meta data-rh="true" name="twitter:creator" content="jalantikuscom"><meta data-rh="true" name="twitter:title" content="10 Aplikasi Chatting Gratis Terbaik dan Terbaru 2023 Selain WhatsApp JalanTikus"><meta data-rh="true" name="twitter:description" content="Sudah menjadi alat yang sangat penting dalam menjaga hubungan sosial kita, kamu wajib tahu rekomendasi aplikasi chatting terbaik 2023 di sini."/></head></pre>
soup = bs(html, 'lxml') print(type(soup)) print(soup.prettify())	<pre>In [5]: soup = bs(html, 'lxml') print(type(soup)) print(soup.prettify()) <class 'bs4.BeautifulSoup'> <!DOCTYPE html> <html lang="id"> <head> <meta charset="utf-8"> <meta content="Indonesian" name="language"> <meta content="width=device-width,initial-scale=1,maximum-scale=1,user-scalable=no,minimum-scale=1" name="viewport"> <meta content="ID-JK" name="geo.region"> <meta content="Jakarta" name="geo.placename"> <meta content="Cn_w4FueEW9fXZdrf3LbVU5mUilJ59PmsFC6jKeP0eI" name="google-site-verification"> <meta content="wgw-epNGIIX_ys2V70Jy6ypAYnoZF1SHZ-gkw2YowBw" name="google-site-verification"> <meta content="391633060859123" property="fb:app_id"> <meta content="387823967920385" property="fb:pages"> <meta content="dmexhibfwg9x7fyuz" name="dailymotion-domain-verification"> <meta content="dmehiqbesixystf2" name="dailymotion-domain-verification"> <title data-rh="true"> 10 Aplikasi Chatting Gratis Terbaik dan Terbaru 2023 Selain WhatsApp JalanTikus </title> <meta content="Sudah menjadi alat yang sangat penting dalam menjaga hubungan sosial kita, kamu wajib tahu rekomendasi aplikasi chatting terbaik 2023 di sini."/></pre>

LAPORAN DATA SCRAPING DAN PRE-PROCESSING DENGAN PYTHON

ARIF FRIMA ARI SUWADJI - 221011700443 - 03SIFE003

INPUT	OUTPUT
<pre> texts = soup.find_all("p") for item in texts: desc = item.get_text() print(desc) </pre>	<pre> In [27]: texts = soup.find_all("p") for item in texts: desc = item.get_text() print(desc) </pre> <p>Rabu, 20 Sep 2023, 18:10 WIB Berkomunikasi dengan teman dan keluarga adalah hal yang menyenangkan. Di era digital seperti sekarang ini, aplikasi chatting te rbaik menjadi salah satu cara untuk menjalin hubungan sosial kita. Aplikasi chatting menjadi salah satu pilihan komunikasi sehari-hari saat ini. WhatsApp adalah aplikasi chat yang populer dan ba nyak dipakai di Indonesia. Tapi, ada juga aplikasi chat terbaik lainnya yang bisa kamu coba. Di artikel ini, Jaka akan memberikan rekomendasi aplikasi chatting terbaik untuk tahun 2023. Aplikasi-aplikasi ini tidak hanya gratis, tetapi juga aman dan memiliki fitur-fitur menarik yang bisa membuat chattingmu lebih seru. WhatsApp adalah salah satu aplikasi pesan instan yang paling populer di dunia saat ini. Dengan lebih dari 2 miliar pengguna akt if bulanan, WhatsApp telah menjadi bagian tak terpisahkan dari kehidupan digital kita. Salah satu kelebihan utama WhatsApp adalah antarmuka pengguna yang sederhana dan mudah digunakan. Aplikasi ini dirancang dengan baik, sehingga pengguna baru pun dapat dengan cepat memahami cara menggunakannya. >>>Link Download WhatsApp APK<<< Telegram adalah aplikasi pesan instan yang diluncurkan pada tahun 2013 oleh Pavel Durov. Aplikasi ini dikembangkan dengan fokus pada kecepatan dan keamanan. Kamu dapat mengirim pesan teks, suara, gambar, video, dan berbagai jenis file lainnya. Salah satu keunggulan utama adalah kecepatan pengiriman pesan yang sangat cepat. Selain itu, Telegram juga menawarkan fitur pan ggilan suara dan video berkualitas tinggi, serta kapasitas penyimpanan yang besar untuk mengirim file dan dokumen. >>>Link Download Telegram APK<<< Facebook Messenger memungkinkanmu untuk dapat mengirim pesan teks secara langsung kepada teman-teman dengan mudah. Kamu dapat m embagikan berita terkini, mengatur pertemuan, atau hanya sekedar menyapa teman. Salah satu kelebihan utamanya adalah kemampuannya untuk memanfaatkan Facebook ID sebagai pengenalan, sehingga kamu tidak perlu be rtukar nomor telepon dengan orang lain untuk dapat mengirim pesan. >>>Link Download Facebook Messenger APK<<< Aplikasi LINE telah menjadi salah satu platform komunikasi paling populer di Indonesia. Dengan lebih dari 200 juta pengguna akt if bulanan, LINE menawarkan berbagai fitur menarik yang membuatnya menjadi aplikasi serba bisa. Menu-menu yang intuitif dan ikon yang jelas membuat penggunaan LINE menjadi sangat mudah, bahkan bagi mereka yang tidak terlalu</p>
<pre> list_rem = [] list_desc = [] texts = soup.find_all("p") for i, item in enumerate(texts): desc = item.get_text() if (i == 0 or i == 6 or i == 9 or i == 12 or i > 28): # print(i, desc); # print('tidak perlu dimasukkan ke list') list_rem.append(desc) else : list_desc.append(desc) print(list_desc) </pre>	<pre> In [38]: list_rem = [] list_desc = [] texts = soup.find_all("p") for i, item in enumerate(texts): desc = item.get_text() if (i == 0 or i == 6 or i == 9 or i == 12 or i > 28): # print(i, desc); # print('tidak perlu dimasukkan ke list') list_rem.append(desc) else : list_desc.append(desc) print(list_desc) </pre> <p>['Berkomunikasi dengan teman dan keluarga adalah hal yang menyenangkan. Di era digital seperti sekarang ini, aplikasi chatting terbaik menjadi salah satu cara untuk menjalin hubungan sosial kita.', 'Aplikasi chatting menjadi salah satu pilihan komunikasi sehari-hari saat ini. WhatsApp adalah aplikasi chat yang populer dan banyak dipakai di Indonesia. Tapi, ada juga aplikasi chat terbaik lainnya yang bisa kamu coba.', 'Di artikel ini, Jaka akan memberikan rekomendasi aplikasi chatting terbaik untuk tahun 2023. Aplikasi-aplikasi ini tidak hanya gratis, tetapi juga aman dan memiliki fitur-fitur menarik yang bisa membuat chattingmu lebih seru.', 'WhatsApp adalah salah satu aplikasi pesan instan yang paling populer di dunia saat ini. Dengan lebih dari 2 mili ar pengguna aktif bulanan, WhatsApp telah menjadi bagian tak terpisahkan dari kehidupan digital kita.', 'Salah satu kelebihan u tama WhatsApp adalah antarmuka pengguna yang sederhana dan mudah digunakan. Aplikasi ini dirancang dengan baik, sehingga penggu na baru pun dapat dengan cepat memahami cara menggunakannya.', 'Telegram adalah aplikasi pesan instan yang diluncurkan pada tah un 2013 oleh Pavel Durov. Aplikasi ini dikembangkan dengan fokus pada kecepatan dan keamanan. Kamu dapat mengirim pesan teks, s uara, gambar, video, dan berbagai jenis file lainnya.', 'Salah satu keunggulan utama adalah kecepatan pengiriman pesan yang san gat cepat. Selain itu, Telegram juga menawarkan fitur panggilan suara dan video berkualitas tinggi, serta kapasitas penyimpanan vane besar untuk mengirim file dan dokumen.', 'Facebook Messenger memungkinkanmu untuk dapat mengirim pesan teks secara langsung</p>

LAPORAN DATA SCRAPING DAN PRE-PROCESSING DENGAN PYTHON

ARIF FRIMA ARI SUWADJI - 221011700443 - 03SIFE003

INPUT	OUTPUT
DATA PRE-PROCESSING-Tokenizing	
<pre>df['data'] = df['data'].apply(lambda x: " .join(x.lower() for x in x.split())) df['data']</pre>	<pre>In [41]: df['data'] = df['data'].apply(lambda x: " ".join(x.lower() for x in x.split())) In [42]: df['data'] Out[42]: 0 berkomunikasi dengan teman dan keluarga adalah... 1 aplikasi chatting menjadi salah satu pilihan k... 2 di artikel ini, jika akan memberikan rekomen... 3 whatsapp adalah salah satu aplikasi pesan inst... 4 salah satu kelebihan utama whatsapp adalah ant... 5 telegram adalah aplikasi pesan instan yang dil... 6 salah satu keunggulan utama adalah kecepatan p... 7 facebook messenger memungkinkanmu untuk dapat ... 8 salah satu kelebihan utamanya adalah kemampuan... 9 aplikasi line telah menjadi salah satu platfor... 10 menu-menu yang intuitif dan ikon yang jelas me... 11 wechat telah menjadi salah satu aplikasi pesan... 12 fitur utama wechat adalah pesan teks dan suara... 13 skype telah menjadi salah satu aplikasi komuni... 14 selain panggilan suara dan video, skype juga m... 15 viber adalah aplikasi chat free yang memungkin... 16 selain panggilan suara dan video, viber juga m... 17 discord telah menjadi salah satu platform komu... 18 pengguna dapat membuat server sendiri, membuat... 19 signal adalah aplikasi chatting online yang sa... 20 signal menempatkan privasi pengguna sebagai pr... 21 google chat memiliki berbagai fitur yang memun... 22 google chat memungkinkan pengguna untuk membua... 23 itulah, daftar 10 aplikasi chatting terbaik unt... 24 setiap aplikasi memiliki kelebihan dan kekurangan... Name: data, dtype: object</pre>
<pre>import nltk def proses(kalimat) : kalimat = kalimat.lower() kalimat = nltk.tokenize.word_tokenize(kalimat) return kalimat hasil = df['data'] hasil = hasil.apply(proses) hasil</pre>	<pre>In [44]: import nltk In [45]: def proses(kalimat) : kalimat = kalimat.lower() kalimat = nltk.tokenize.word_tokenize(kalimat) return kalimat In [46]: hasil = df['data'] hasil = hasil.apply(proses) Out[46]: 0 [berkomunikasi, dengan, teman, dan, keluarga, ... 1 [aplikasi, chatting, menjadi, salah, satu, pil... 2 [di, artikel, ini, ,, jika, akan, memberikan, ... 3 [whatsapp, adalah, salah, satu, aplikasi, pesa... 4 [salah, satu, kelebihan, utama, whatsapp, adal... 5 [telegram, adalah, aplikasi, pesan, instan, ya... 6 [salah, satu, keunggulan, utama, adalah, kecep... 7 [facebook, messenger, memungkinkanmu, untuk, d... 8 [salah, satu, kelebihan, utamanya, adalah, kem... 9 [aplikasi, line, telah, menjadi, salah, satu, ... 10 [menu-menu, yang, intuitif, dan, ikon, yang, j... 11 [wechat, telah, menjadi, salah, satu, aplikasi... 12 [fitur, utama, wechat, adalah, pesan, teks, da... 13 [skype, telah, menjadi, salah, satu, aplikasi,... 14 [selain, panggilan, suara, dan, video, ,, skyp... 15 [viber, adalah, aplikasi, chat, free, yang, me... 16 [selain, panggilan, suara, dan, video, ,, vibe... 17 [discord, telah, menjadi, salah, satu, platfor... 18 [pengguna, dapat, membuat, server, sendiri, ,,... 19 [signal, adalah, aplikasi, chatting, online, y... 20 [signal, menempatkan, privasi, pengguna, sebag... 21 [google, chat, memiliki, berbagai, fitur, yang... 22 [google, chat, memungkinkan, pengguna, untuk, ... 23 [itulah, daftar, 10, aplikasi, chatting, terba... 24 [setiap, aplikasi, memiliki, kelebihan, dan, k... Name: data, dtype: object</pre>

LAPORAN DATA SCRAPING DAN PRE-PROCESSING DENGAN PYTHON

ARIF FRIMA ARI SUWADJI - 221011700443 - 03SIFE003

INPUT	OUTPUT
DATA PRE-PROCESSING-Filtering	
from string import punctuation list_desc	<pre>In [47]: from string import punctuation</pre> <pre>In [52]: list_desc</pre> <pre>Out[52]: ['Berkomunikasi dengan teman dan keluarga adalah hal yang menyenangkan. Di era digital seperti sekarang ini, aplikasi chatting terbaik menjadi salah satu cara untuk menjalin hubungan sosial kita.', 'Aplikasi chatting menjadi salah satu pilihan komunikasi sehari-hari saat ini. WhatsApp adalah aplikasi chat yang populer dan banyak dipakai di Indonesia. Tapi, ada juga aplikasi chat terbaik lainnya yang bisa kamu coba.', 'Di artikel ini, Jaka akan memberikan rekomendasi aplikasi chatting terbaik untuk tahun 2023. Aplikasi-aplikasi ini tidak hanya gratis, tetapi juga aman dan memiliki fitur-fitur menarik yang bisa membuat chattingmu lebih seru.', 'WhatsApp adalah salah satu aplikasi pesan instan yang paling populer di dunia saat ini. Dengan lebih dari 2 miliar pengguna aktif bulanan, WhatsApp telah menjadi bagian tak terpisahkan dari kehidupan digital kita.', 'Salah satu kelebihan utama WhatsApp adalah antarmuka pengguna yang sederhana dan mudah digunakan. Aplikasi ini dirancang dengan</pre>
stopwords = set(nltk.corpus.stopwords.words('indonesian')) def proses(kalimat): kalimat = kalimat.lower() kalimat = nltk.tokenize.word_tokenize(kalimat) kalimat = [kata for kata in kalimat if kata not in stopwords] return kalimat hasil = df['data'] hasil = hasil.apply(proses) hasil	<pre>In [59]: stopwords = set(nltk.corpus.stopwords.words('indonesian')) def proses(kalimat): kalimat = kalimat.lower() kalimat = nltk.tokenize.word_tokenize(kalimat) kalimat = [kata for kata in kalimat if kata not in stopwords] return kalimat</pre> <pre>In [60]: hasil = df['data'] hasil = hasil.apply(proses) hasil</pre> <pre>Out[60]: 0 [berkomunikasi, teman, keluarga, menyenangkan,... 1 [aplikasi, chatting, salah, pilihan, komunikas... 2 [artikel, ,, jaka, rekomendasi, aplikasi, chat... 3 [whatsapp, salah, aplikasi, pesan, instan, pop... 4 [salah, kelebihan, utama, whatsapp, antarmuka,... 5 [telegram, aplikasi, pesan, instan, diluncurka... 6 [salah, keunggulan, utama, kecepatan, pengirim... 7 [facebook, messenger, memungkinkanmu, mengirim... 8 [salah, kelebihan, utamanya, kemampuannya, mem... 9 [aplikasi, line, salah, platform, komunikasi, ... 10 [menu-menu, intuitif, ikon, penggunaan, line, ... 11 [wechat, salah, aplikasi, pesan, instan, popul... 12 [fitur, utama, wechat, pesan, teks, suara, pen... 13 [skype, salah, aplikasi, komunikasi, populer, ... 14 [panggilan, suara, video, ,, skype, menyediaka... 15 [viber, aplikasi, chat, free, pengguna, panggi... 16 [panggilan, suara, video, ,, viber, menyediaka... 17 [discord, salah, platform, komunikasi, terpopu... 18 [pengguna, server, ,, saluran, suara, teks, ,... 19 [signal, aplikasi, chatting, online, populer, ... 20 [signal, menempatkan, privasi, pengguna, prior... 21 [google, chat, memiliki, fitur, kolaborasi, ti... 22 [google, chat, pengguna, bergabung, ruang, obr... 23 [daftar, 10, aplikasi, chatting, terbaik, 2023... 24 [aplikasi, memiliki, kelebihan, kekurangan, ,... Name: data, dtype: object</pre>