

# LAPORAN DATA SCRAPING DAN PRE-PROCESSING DENGAN PYTHON

ARIF FRIMA ARI SUWADJI - 221011700443 - 03SIFE003

INPUT	OUTPUT
<pre># Proses Lemmatisasi import nltk from nltk.stem import WordNetLemmatizer  wnl = WordNetLemmatizer() list1 = list_desc[0].split(" ") # print(list1) for words in list1:     print(words + " --&gt; " + wnl.lemmatize(words))</pre>	<pre>In [6]: # Proses Lemmatisasi import nltk from nltk.stem import WordNetLemmatizer  In [11]: wnl = WordNetLemmatizer() list1 = list_desc[0].split(" ") # print(list1) for words in list1:     print(words + " --&gt; " + wnl.lemmatize(words))  Berkomunikasi --&gt; Berkomunikasi dengan --&gt; dengan teman --&gt; teman dan --&gt; dan keluarga --&gt; keluarga adalah --&gt; adalah hal --&gt; hal yang --&gt; yang menyenangkan. --&gt; menyenangkan. Di --&gt; Di era --&gt; era digital --&gt; digital seperti --&gt; seperti sekarang --&gt; sekarang ini, --&gt; ini, aplikasi --&gt; aplikasi chatting --&gt; chatting terbaik --&gt; terbaik menjadi --&gt; menjadi salah --&gt; salah satu --&gt; satu cara --&gt; cara untuk --&gt; untuk menjalin --&gt; menjalin hubungan --&gt; hubungan sosial --&gt; sosial kita. --&gt; kita.</pre>
<pre># Proses Stemming from nltk.stem import PorterStemmer from nltk.tokenize import word_tokenize  ps = PorterStemmer() words = list1 for x in words:     print(x, " : ", ps.stem(x))</pre>	<pre>In [8]: # Proses Stemming from nltk.stem import PorterStemmer from nltk.tokenize import word_tokenize  In [12]: ps = PorterStemmer() words = list1 for x in words:     print(x, " : ", ps.stem(x))  Berkomunikasi : berkomunikasi dengan : dengan teman : teman dan : dan keluarga : keluarga adalah : adalah hal : hal yang : yang menyenangkan. : menyenangkan. Di : di era : era digital : digit seperti : seperti sekarang : sekarang ini, : ini, aplikasi : aplikasi chatting : chat terbaik : terbaik menjadi : menjadi salah : salah satu : satu cara : cara untuk : untuk menjalin : menjalin hubungan : hubungan sosial : sosial kita. : kita.</pre>