

LAPORAN PRAKTIKUM
DATA MAINING



Disusun Oleh :

Nama : Fahmi Adi Setiawan
NIM : 22230010
Mata Kuliah : Data Maining

Program Studi Sistem Informasi
Fakultas Sains dan Teknologi
Universitas Respati Yogyakarta
2025/2026

Kodingan dan hasil Running

```
[4]: !pip install Sastrawi

Requirement already satisfied: Sastrawi in c:\users\user\anaconda3\lib\site-packages (1.0.1)

[17]: dokumen1="Barangnya sesuai harapan, cakep dan mulus. Penjualnya juga baik dan ramah. Overall, puas banget belanja di sini."
#Case Folding
hasil_cf = str.lower(dokumen1)
print("Case Folding : "+hasil_cf);

#Tokenizing
from nltk.tokenize import RegexpTokenizer
tokenizer = RegexpTokenizer(r'\w+')
hasil_token = tokenizer.tokenize(hasil_cf)
print("Token:")
print(hasil_token)

# Stemming
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()
input_stemm = str(hasil_token) # Make sure 'hasil_token' is defined earlier in your code
hasil_stemm = stemmer.stem(input_stemm) # Use 'input_stemm' instead of 'input_stem'
print("Hasil Stemming:")
print(hasil_stemm)

# Filtering
#Tokenisasi ulang hasil stemming
token_stemm = tokenizer.tokenize(hasil_stemm)

# Ambil daftar stopwords
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
stop_factory = StopWordRemoverFactory()
stopwords = stop_factory.get_stop_words()

#Hapus stopwords dari hasil stemming
hasil_final = [word for word in token_stemm if word not in stopwords]
print("Hasil Setelah Stopword Removal:")
print(hasil_final)
```

```
Case Folding : barangnya sesuai harapan, cakep dan mulus. penjualnya juga baik dan ramah. overall, puas banget belanja di sini.
Token:
['barangnya', 'sesuai', 'harapan', 'cakep', 'dan', 'mulus', 'penjualnya', 'juga', 'baik', 'dan', 'ramah', 'overall', 'puas', 'banget', 'belanja', 'di', 'sini']
Hasil Stemming:
barang sesuai harap cakep dan mulus jual juga baik dan ramah overall puas banget belanja di sini
Hasil Setelah Stopword Removal:
['barang', 'sesuai', 'harap', 'cakep', 'mulus', 'jual', 'baik', 'ramah', 'overall', 'puas', 'banget', 'belanja', 'sini']
```

[]: