

```
import stanza
from collections import Counter
import re

# download and load Norwegian Stanza model
stanza.download("no")
nlp = stanza.Pipeline("no", processors="tokenize, pos, lemma")

# get text corpus
with open("elo_tesla.txt", "r", encoding="utf-8") as file:
    text = file.read()

# skip the punctuation and make text to lowercase
text_cleaned = re.sub(r"[^\w\s]", "", text.lower())

# Process text using Stanza
doc = nlp(text_cleaned)

# extract lemmas and original word forms
lemmas = []
word_forms = []

for sentence in doc.sentences:
    for word in sentence.words:
        lemmas.append(word.lemma)
        word_forms.append(word.text.lower())

# Counter for occurrences (lemma and word)
lemma_counts = Counter(lemmas)
word_counts = Counter(word_forms)

# 10 most frequent lemmas and word forms
top_lemmas = lemma_counts.most_common(10)
top_word_forms = word_counts.most_common(10)

# results
print("Top 10 Lemmas:")
for lemma, count in top_lemmas:
    print(f"{lemma}: {count}")

print("\nTop 10 Word Forms:")
for word, count in top_word_forms:
    print(f"{word}: {count}")
```