

Stemming 是比較簡單的做法，通常稱為 Heuristic process，簡單來說就是制定規則來拆解單詞(rule-based)，像是看到 ing/ed 就去除。其中，又以 Porter stemmer 跟 Snowball stemmer 最為常見，前者是運用現有的規則移除常見單詞字尾，後者則是根據使用者對 Porter stemmer 提出的問題加以優化。

Lemmatization 是取出單詞的 Lemma，且 Lemmatization 需要有字典來尋找單詞的原型，除此之外 Lemmatization 也可以利用 POS tagging 的訊息來給出更正確的答案。

大部分的 lemmatization 演算法都是藉由維護大型詞典以及對輸入詞進行形態分析來達到目標，而 stemming 演算法則由簡單的字串切割、if else 規則、小型詞典所組成。基本上 Lemmatization 比起 Stemming 會是更好的選擇，因為它的精度較高，Stemming 唯一的優勢就是速度。