Введение

Основная идея в использовании NLP. Стоит посмотреть в сторону библиотеки nltk и алгоритма основанного на "мешке слов"