Определяем части речи.

пословная классификация

предл.

слово

класс
= часть речи
N - сущ   V - глагол   Adj, Adv   Pro
                    ↑прил ↑нар   ↑местоимение

класс зависит от соседних классов = поток

HMM — hidden markov model , скрытая марк. модель

триг
$p(x, y)$ = ф-ла
            классы
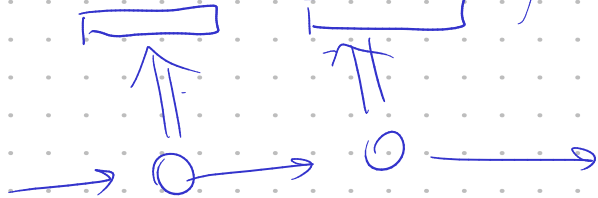
Conditional random field . $p(y|x)$ = ф-ла
                            $= p(x, y) / p(x)$

Для обучения

слово

признаки слово: ▭
окончание : 2 три буквы
форма : с большой буквы

⌈ см. больше, но для след слова
⌊ см. больше, но для пред слова

⌈ начало предл : true / false

Конец пред. ч tone /false

↓
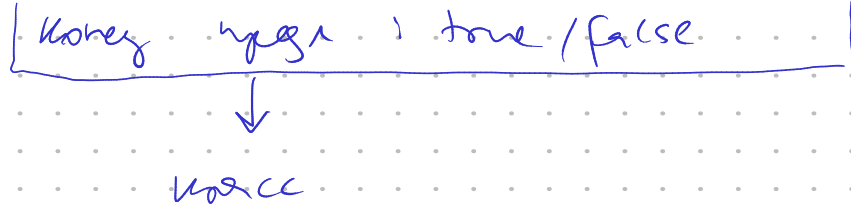
класс

Можно также решать задачу поиска имен. сущ.

NER

+ признаки о форме, част. речи,

Классы: ⬜ ⬜ ООО вектор ⬜

класс O_ther   O_ther   C$ CI   other

company start   company in
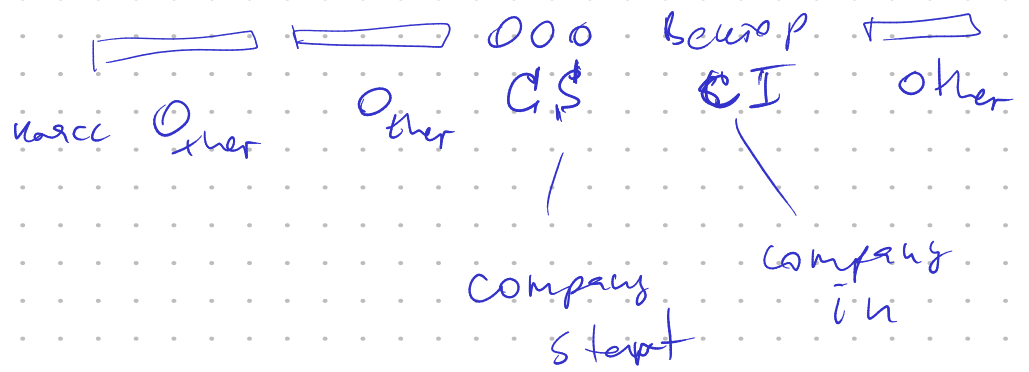
gazetter
+ газетир. слово ∈ списку названий

recall — полнота

сколько сущ определено как сущ
сколько настоящих сущ

precision — точность

сколько сущ оп. как сущ
сколько слов определен как сущ

$$f1 = \cfrac{2}{\frac{1}{r} + \frac{1}{p}}$$

$\sqrt{rp}$ - геом
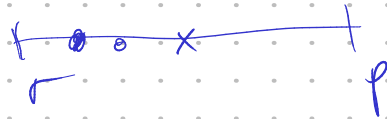
accuracy — точность = право ответов / всего ответов