

1. Прочитать статью <http://www.scielo.org.mx/pdf/cys/v18n3/v18n3a7.pdf> и сделать следующее:

- Написать кратко (не более 300 слов (минимум 100), отчет в формате pdf) о различиях между cosine similarity и soft similarity. Привести примеры использования и написать собственный пример вычисления cosine similarity и soft similarity для произвольных векторов (не брать вектора из статьи! надо самим придумать координаты/размерность векторов)

a="Основная идея soft similarity заключается в том, что сходство между признаками известно из обычных источников типа академического словаря и не нуждается в изучении на основе данных. При отсутствии сходства между признаками, soft similarity равно стандартному сходству. Soft cosine measure по сравнению с обычным косинусом корректирует свои значения на дополнительную матрицу признаков полученных путём математических расчётов. Например, подсчитывается сколько раз то или иное слово появлялось в тексте. Основная идея различия soft similarity и soft cosine measure заключается в более глубокой проработке признаков модели. Если в случае с soft similarity берутся обычные признаки из открытых источников, то, в случае с soft cosine measure, новые качества признаков получаются путём проведения математических операций над признаками. На мой взгляд основным примером использования в данном случае является модель отвечающая за автоподбор слов при наборе текста на смартфонах или в поисковой строке браузера."

```
In [17]: 1 len(list(a.split(' ')))
```

```
Out[17]: 135
```

$a = (1, 0, 1, 1, 1)$ случайно добавили
 $b = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$ ненулевые связи между
 признаками. Остальные = 0

$$s_{1,2} = 0,2$$

$$s_{1-5} = 0,5$$

$$s_{3,4} = 0,3$$

$$s_{2-4} = 0,1$$

$$s_0 \text{ ft. cosine} = \frac{s_1 \quad s_5 \quad s_{3-4} \quad s_{1-5} \quad s_{2-4}}{1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 0,3 \cdot 1 \cdot 1 + 0,5 \cdot 1 \cdot 1 + 0,1 \cdot 1 \cdot 1}$$

$$\sqrt{0,3 \cdot 1 \cdot 1 + 0,5 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1} \cdot \sqrt{\quad}$$

$s_{3-4} \quad \cdot \quad s_{1-5} \quad s_1 \quad s_3 \quad s_4 \quad s_5$

$$\sqrt{s_{2-4} \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1} = \frac{2,9}{\sqrt{4,8} \cdot \sqrt{3,1}} = 0,75$$