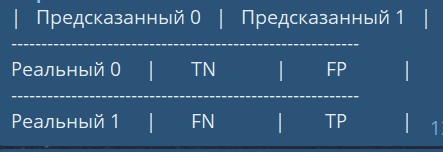
Знать!

1. мы просматриваем разреженную матрицу, в которой только ненулевые элементы отображаются в выводе. Разреженные матрицы используются для хранения больших матриц, в которых большинство элементов равны нулю. Это позволяет экономить память и ускоряет вычисления.  
   
2. Ошибка первого рода (False Positive, FP):  
   Происходит, когда модель неправильно классифицирует отрицательный пример как положительный.  
   Например, если модель классификации фишинговых писем классифицирует обычное письмо как фишинговое, это ошибка первого рода.  
   Также называется ложно положительным результатом.
3. Ошибка второго рода (False Negative, FN):  
   Происходит, когда модель неправильно классифицирует положительный пример как отрицательный.  
   Например, если модель классификации фишинговых писем классифицирует фишинговое письмо как обычное, это ошибка второго рода.  
   Также называется ложно отрицательным результатом.
4. Векторизатор TfidfVectorizer токенизирует на отдельные слова, удаляет английские стоп-слова «stop\_words='english'», приводит все символы к нижнему регистру «lowercase=True», создает словарь уникальных слов из набора данных, подсчитывает частоту каждого слова в каждом сообщении (Вычисления TF), определяет важность слов по всему набору данных, снижая вес часто встречающихся слов (Вычисления IDF), для каждого слова в каждом документе умножается частота слова TF на его обратную частоту IDF (Вычисление TF-IDF) и преобразует к числовому формату матрицы, где строки представляют сообщения, а столбцы — уникальные слова (Матрица TF-IDF)
5.   
   **TN (True Negative, Истинно отрицательные)**: количество правильных предсказаний, когда реальный класс 0 (обычные письма) и предсказанный класс также 0.  
   **FP (False Positive, Ложно положительные)**: количество неправильных предсказаний, когда реальный класс 0, но предсказанный класс 1 (фальшивое срабатывание, обычное письмо определено как фишинговое).  
   **FN (False Negative, Ложно отрицательные)**: количество неправильных предсказаний, когда реальный класс 1 (фишинговые письма), но предсказанный класс 0 (неудача в обнаружении фишинга).  
   **TP (True Positive, Истинно положительные)**: количество правильных предсказаний, когда реальный класс 1 и предсказанный класс также 1.
6. Наивный Байесовский классификатор называется "наивным" потому, что он предполагает, что все признаки в данных независимы друг от друга. Это предположение редко выполняется в реальности, но значительно упрощает вычисления.