Задача: выделить фрагменты в тексте и протегировать. Метод классификации токенов (sequence labeling). $x_i \to y_i = \{0,1\}^T$. Если $y_{i,j} = 1$, то i-й токен входит в фрагмент с тегом j. Обучаем модель минимизацией бинарной кросс-энтропии:

$$\mathcal{L}(y, \hat{y}) = \sum_{i,j} logloss(y_{i,j}, \hat{y}_{i,j}) \to \min$$

$$logloss(y, \hat{y}) = -ylog(\hat{y}) - (1-y)log(1-\hat{y})$$