Оцінка важливості кожного слова в досліджуваному тексті проводиться з використанням методу дисперсійного оцінювання, який є оцінкою дискримінантної сили слів. Метод дисперсійного оцінювання дозволяє відділити із загальної множини широковживаних у тексті слів слова, що розташовані рівномірно й показав свою високу ефективність у попередніх дослідженнях.

Відповідно до існуючої математичної моделі [6], якщо деяке слово *A* в тексті, що складається з *N* слів,

*k*

позначене як

*An* , де індекс *k* – номер появи даного слова в тесті, а *n* – позиція даного слова в тексті, то

інтервал між послідовними появами слова при таких позначеннях буде величина

**

де на *m*-ій і *n*-ій позиціях в тексті знаходиться слово *A* , яке зустрілось *k*+1-ий і *k*-ий рази. Таким чином, дисперсійна оцінка розраховується за формулою



*Де *– середнє значення послідовності *, , ..., *

**– послідовності *, , ..., , k - кількість появ слова* A у тексті.

Вхідними даними блоку є лематизваний текстового контент визначеного фрагменту контенту цифрового документу навчального матеріалу, вихідними даними – впорядкована множина слів, кожному з яких співставлена оцінка його дисперсії, що позиціонується як оцінка важливості даного слова у досліджуваному фрагменті цифрового документу.

1 2 *k*

1 2 *k*