

```

1 import java.util.HashMap;
2
3 public class DocumentCollection {
4     private HashMap<String, Document> documents = new HashMap<>();
5
6     public DocumentCollection(String[] paths) {
7         try {
8             for (String path : paths) {
9                 documents.put(path, new Document(path));
10            }
11        } catch (Exception e) {
12            System.out.println("Error Reading file!: " + e.getMessage());
13        }
14    }
15
16    /**
17     * calculates tf_idf
18     * @param term: term w to calculate tf_idf
19     * @param document: document d in collection to calculate tf_idf
20     * @return value for tf_idf
21     */
22    public double get_tf_idf(String term, String document) {
23        double tf = documents.get(document).get_tf(term);
24        try {
25            double idf = get_idf(term);
26            return calculate_tf_idf(tf, idf);
27        } catch (Exception e) {
28            System.out.println("Term does not occur: " + e.getMessage());
29            return 0;
30        }
31    }
32
33    public static double calculate_tf_idf(double tf, double idf) {
34        return tf * idf;
35    }
36
37    /**
38     * returns idf
39     *
40     * @param term: given term i to use to calculate idf
41     * @return value for idf
42     * @throws Exception
43     */
44    public double get_idf(String term) throws Exception {
45        int n_i = 0;
46        for (Document document : documents.values()) {
47            if (document.doesTermAccur(term)) {
48                n_i += 1;
49            }
50        }
51        return calculate_idf(documents.size(), n_i);
52    }
53
54    /**
55     * Method to calculate inverse document frequency
56     * @param N: Total number of documents
57     * @param n_i: Number of occurrences of a term
58     * @return value for idf
59     */
60    public static double calculate_idf(double N, double n_i) {
61        return log2(N / n_i);
62    }
63
64    private static double log2(double x) {
65        return (Math.log(x) / Math.log(2));
66    }
67 }

```