**Subject: Information Retrieval**

Quiz: TF.IDF and Query Evaluation

Date:

Student name:

Student ID:

Chúng ta có 10 tài liệu như sau:

|  |  |
| --- | --- |
| **No.** | **Contents** |
| 1. | AABAAABBBAACD |
| 2. | ACDAAABBBCCCDDD |
| 3. | CCDDDCDDDDDCCC |
| 4. | ABCDEABCDEABCDE |
| 5. | ABCCCCCCCCCCBC |
| 6. | DEADEADDDDDE |
| 7. | BEBBBBEEECBBBC |
| 8. | ABCAAAABBE |
| 9. | AEEEEABCEEA |
| 10. | ABCBCBCBCDDD |

Chúng ta sẽ có một số câu query như sau:

|  |  |  |
| --- | --- | --- |
| **No** | **Query** | **Purpose** |
| 1. | A | Tài liệu có A |
| 2. | B | Tài liệu nói về B (ít nhất 30%) |
| 3. | A-B | Tài liệu có A và B |
| 4. | A-B | Tài liệu nói về A và B: ít nhất 3A hoặc 3B và |A-B|<=2 |

Ground Truth – Golden Set

1. {1,2,4,5,6,8,9,10}
2. {}