

# Implementing Locality Sensitive Hashing from scratch

5 questions

---

1  
point

1. What is the document ID of Barack Obama's article?

35817

---

1  
point

2. Which bin contains Barack Obama's article? Enter its integer index.

50194

---

1  
point

3.  
Examine the bit representations of the bins containing Barack Obama and Joe Biden. In how many places do they agree?

- ☒ 16 out of 16 places (Barack Obama and Joe Biden fall into the same bin)
- ☐ 14 out of 16 places
- ☐ 12 out of 16 places
- ☐ 10 out of 16 places



8 out of 16 places

1  
point

4.

Refer to the section "Effect of nearby bin search". What was the smallest search radius that yielded the correct nearest neighbor for Obama, namely Joe Biden?

2

1  
point

5.

Suppose our goal was to produce 10 approximate nearest neighbors whose average distance from the query document is within 0.01 of the average for the true 10 nearest neighbors. For Barack Obama, the true 10 nearest neighbors are on average about 0.77. What was the smallest search radius for Barack Obama that produced an average distance of 0.78 or better?

7

Submit Quiz

