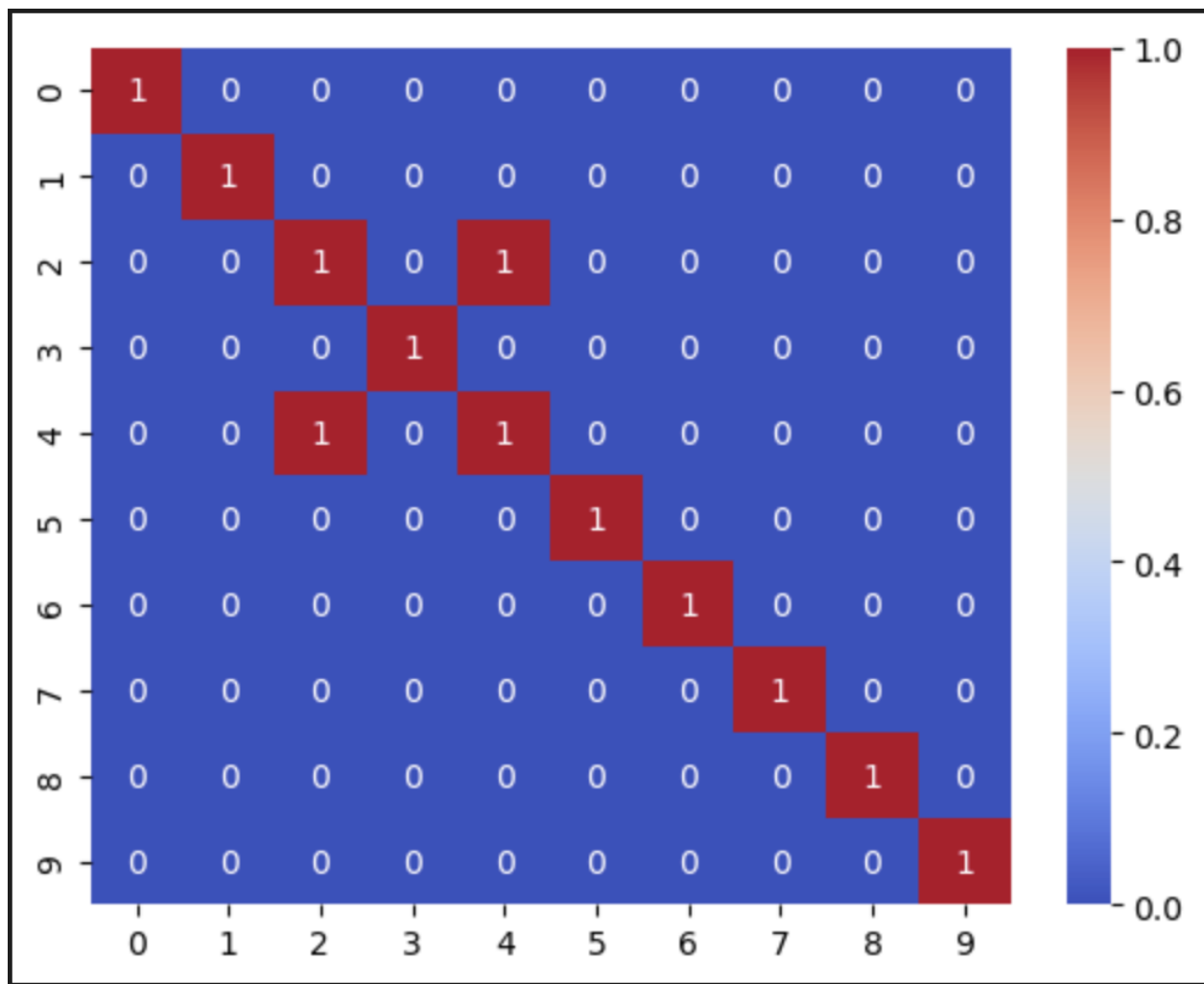
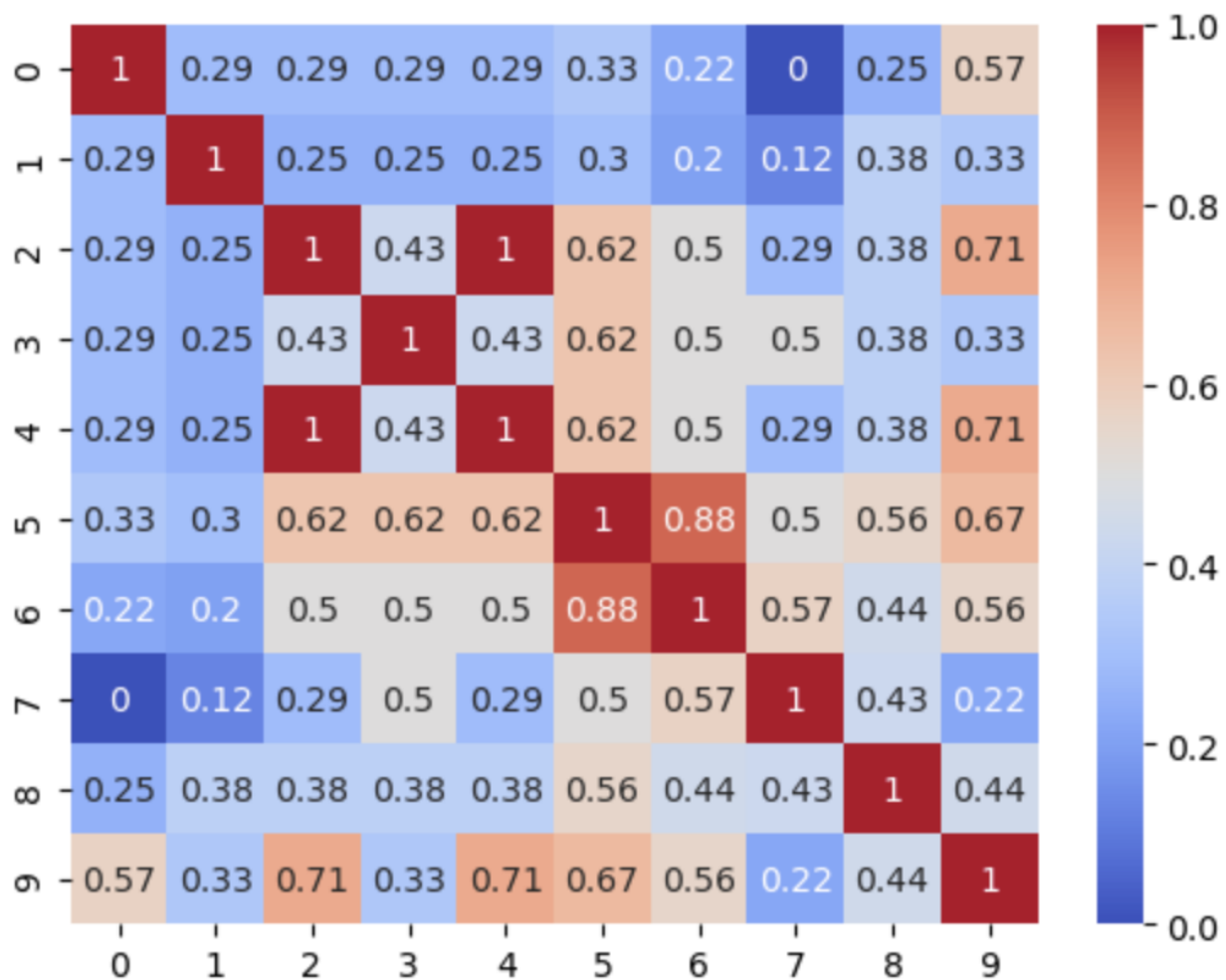


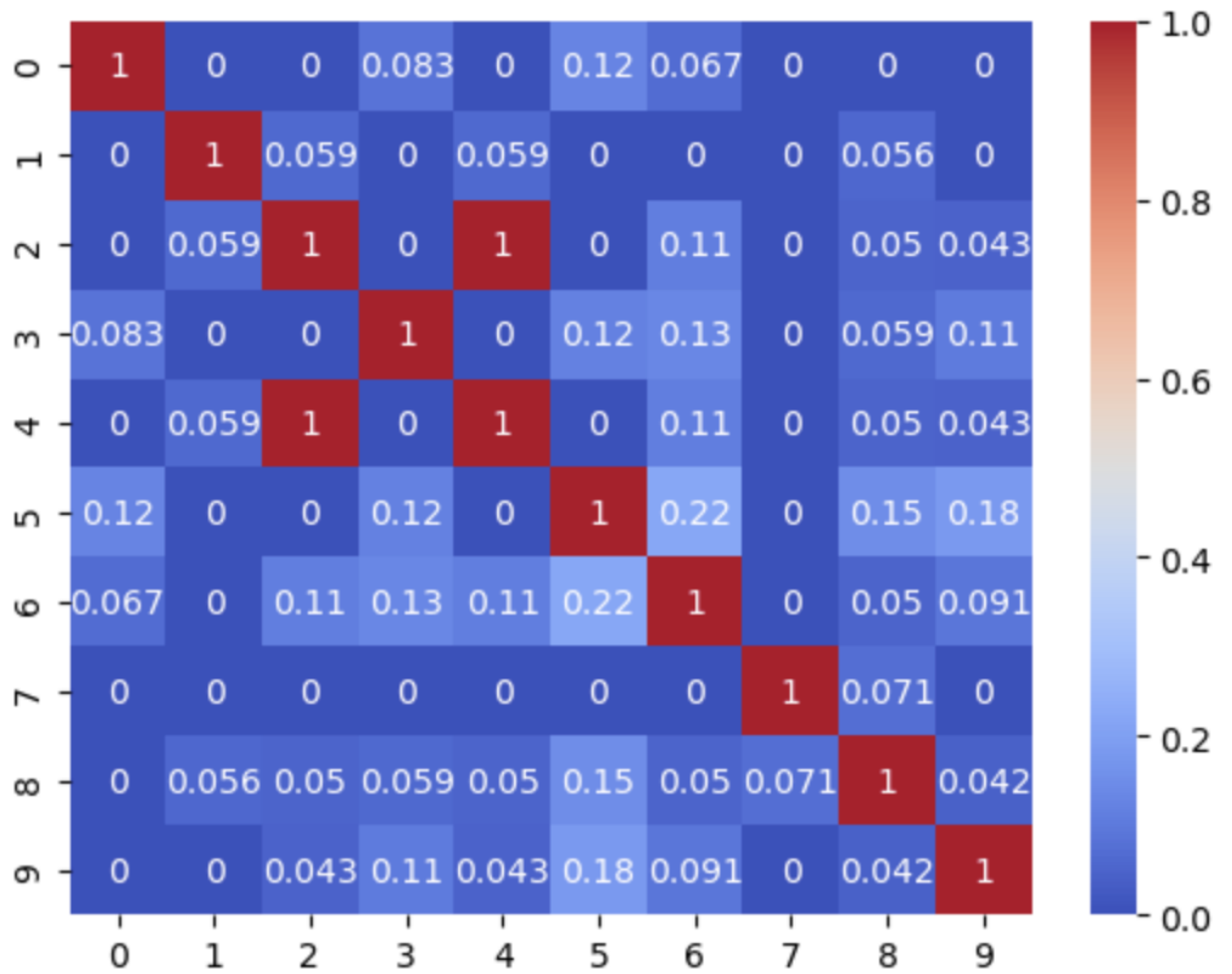
1. Jaccard



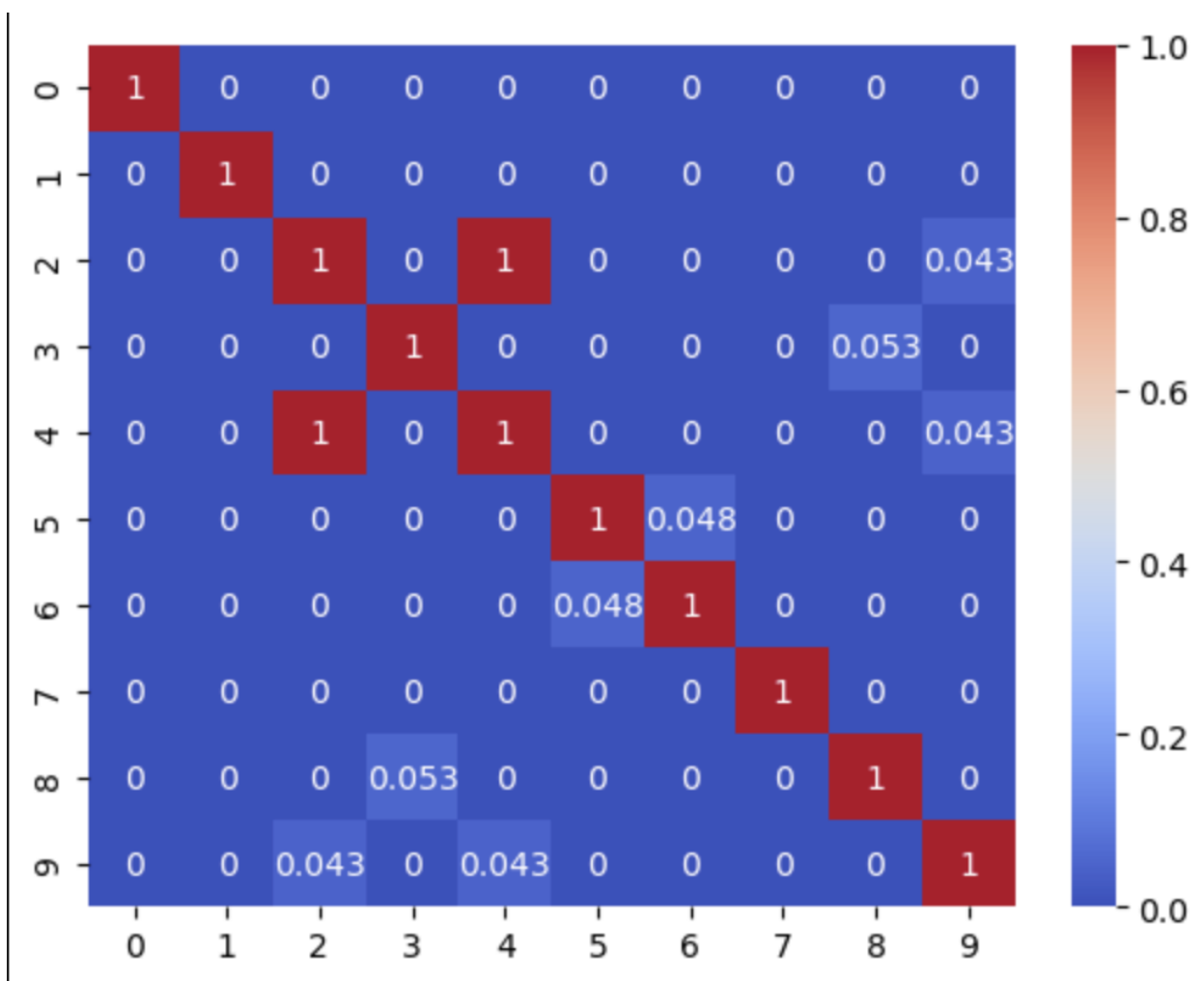
2. BF Jaccard



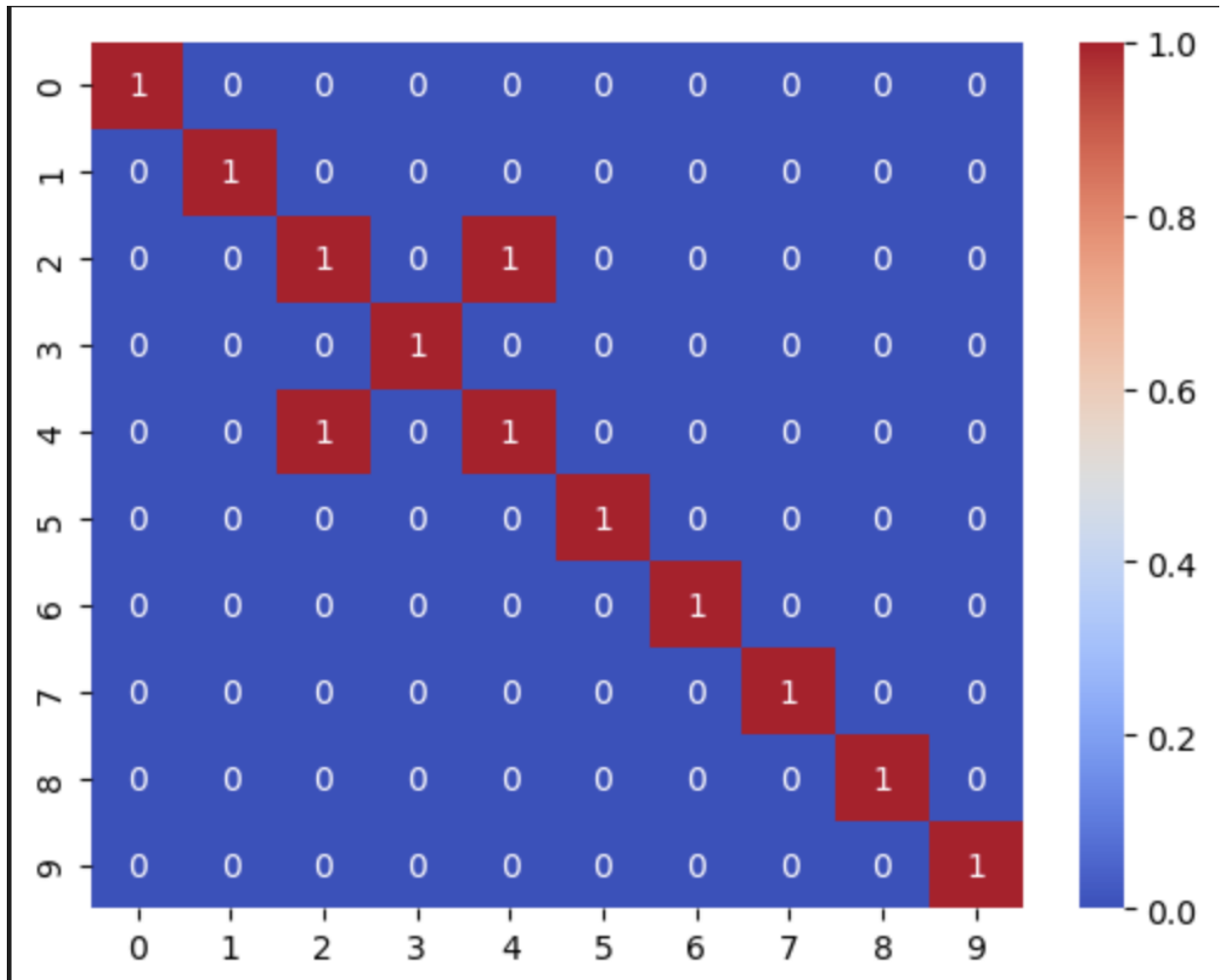
(m=10)



(m=100)

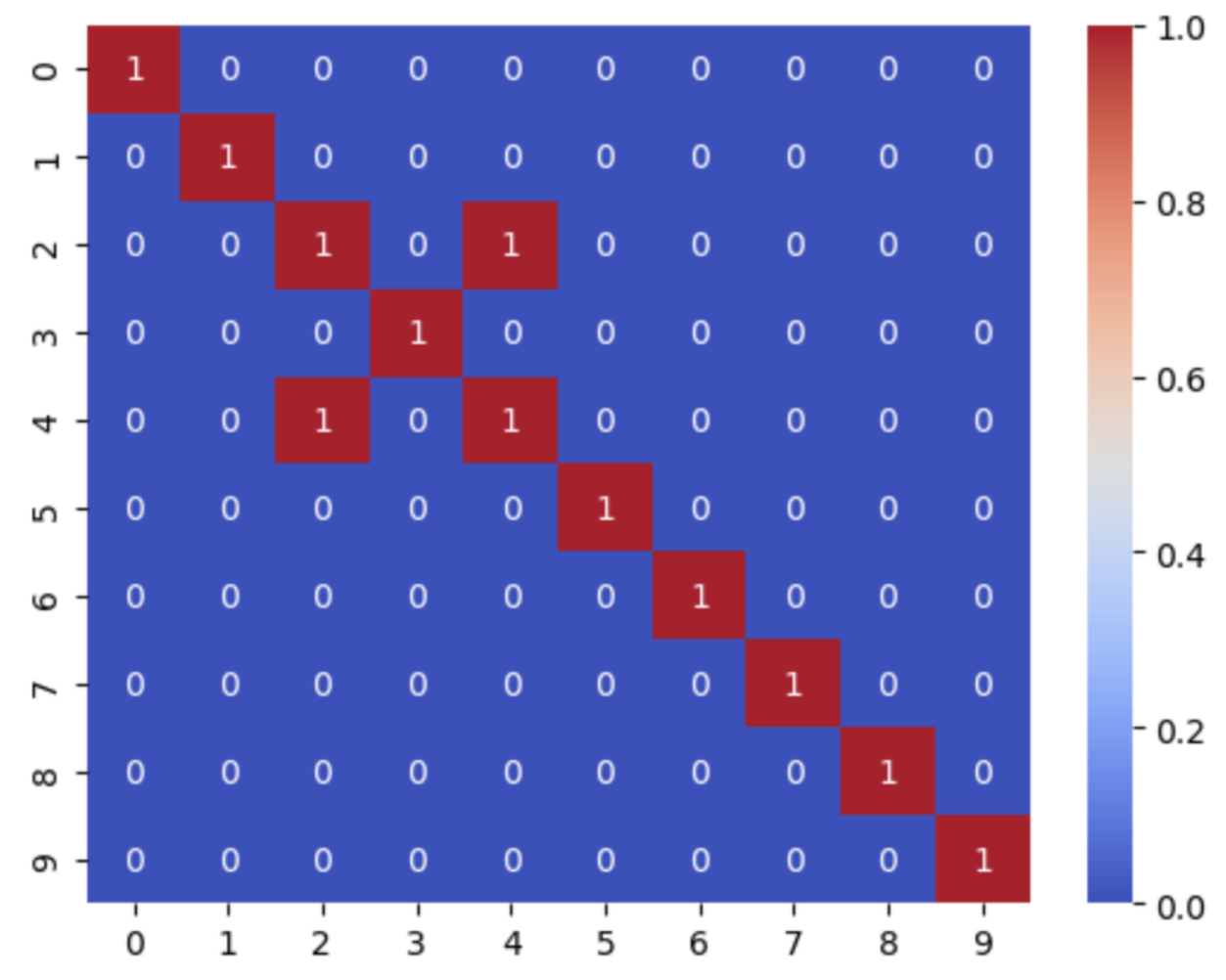


(m=1000)



(m=10000)

3. Minhash Jaccard



Extra.

```
# Checking BloomFilter
```

```
def bloomFilter(inSet, m, hs1=0, hs2=10):
```

```
    bvvector = bv.BitVector(size = m)
```

```
    for i in inSet:
```

```
        b=[]
```

```
        h1=(seededHash(i,hs1))%m
```

```
        h2=(seededHash(i,hs2))%m
```

```
        bvvector[h1]=1
```

```
        bvvector[h2]=1
```

```
    return bvvector
```

```
def InBloomFilter(actor, drama, m, hs1=0, hs2=10):
```

```
    inSet = set(list(df["Main Role"])[list(df["Title"]).index(drama)].split(",
```

```
    bvvector = bloomFilter(inSet, m, hs1, hs2)
```

```
    h1 = (seededHash(actor, hs1)) % m
```

```
    h2 = (seededHash(actor, hs2)) % m
```

```
    if bvvector[h1] == 1 and bvvector[h2] == 1:
```

```
        return True
```

```
    else:
```

```
        return False
```

```
InBloomFilter("Park Ji Hoon", "Weak Hero Class 1 " , 1000)
```

✓ 0.0s

True