Team Member1:

Jagadish Tarla

Tarla.2@wright.edu

U0110479

Team Member2:

Dileep Kumar Reddy kollareddy

kollareddy.2@wright.edu

U01094135

Team Member3:

Lakshma Reddy Chada

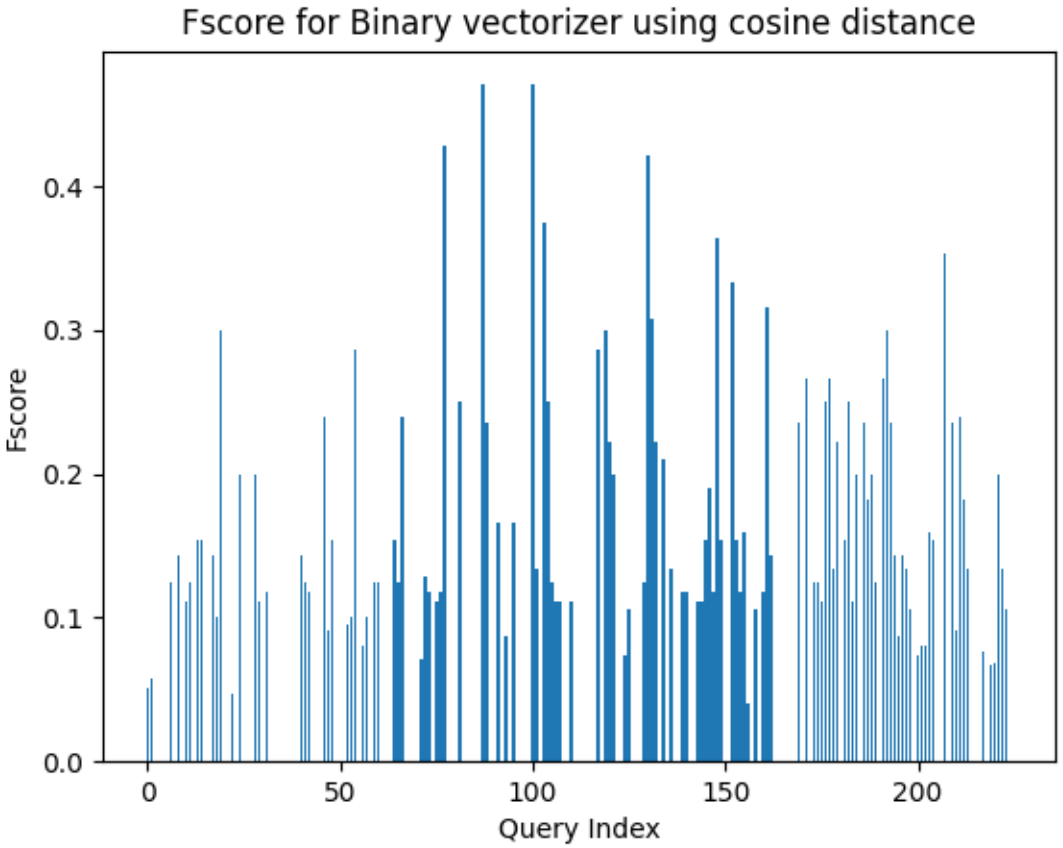
Chada.12@wright.edu

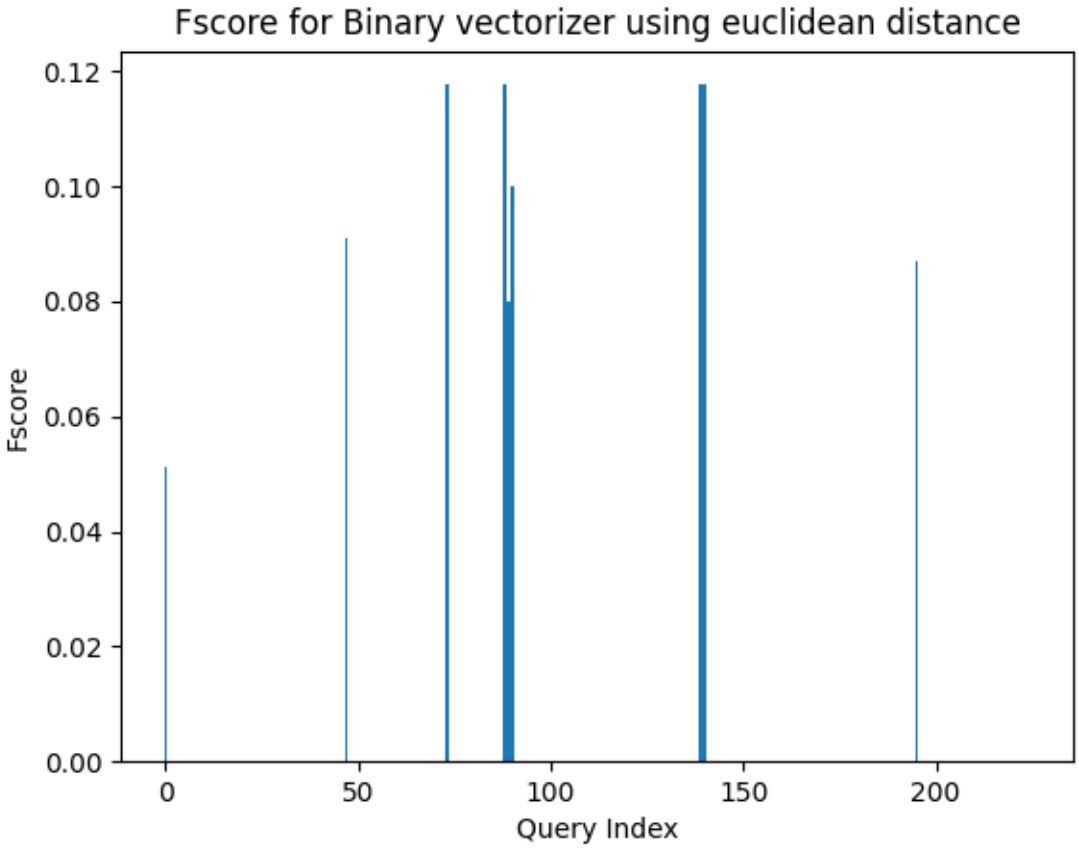
U01095958

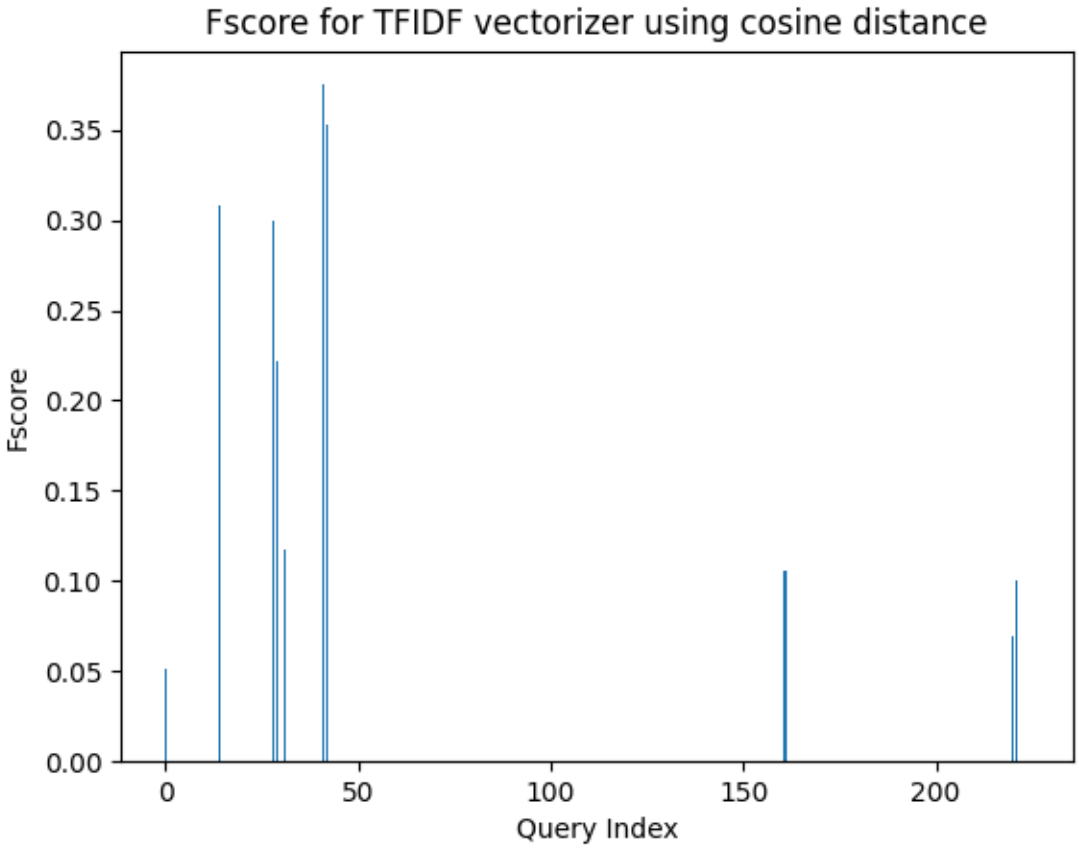
F-Score:

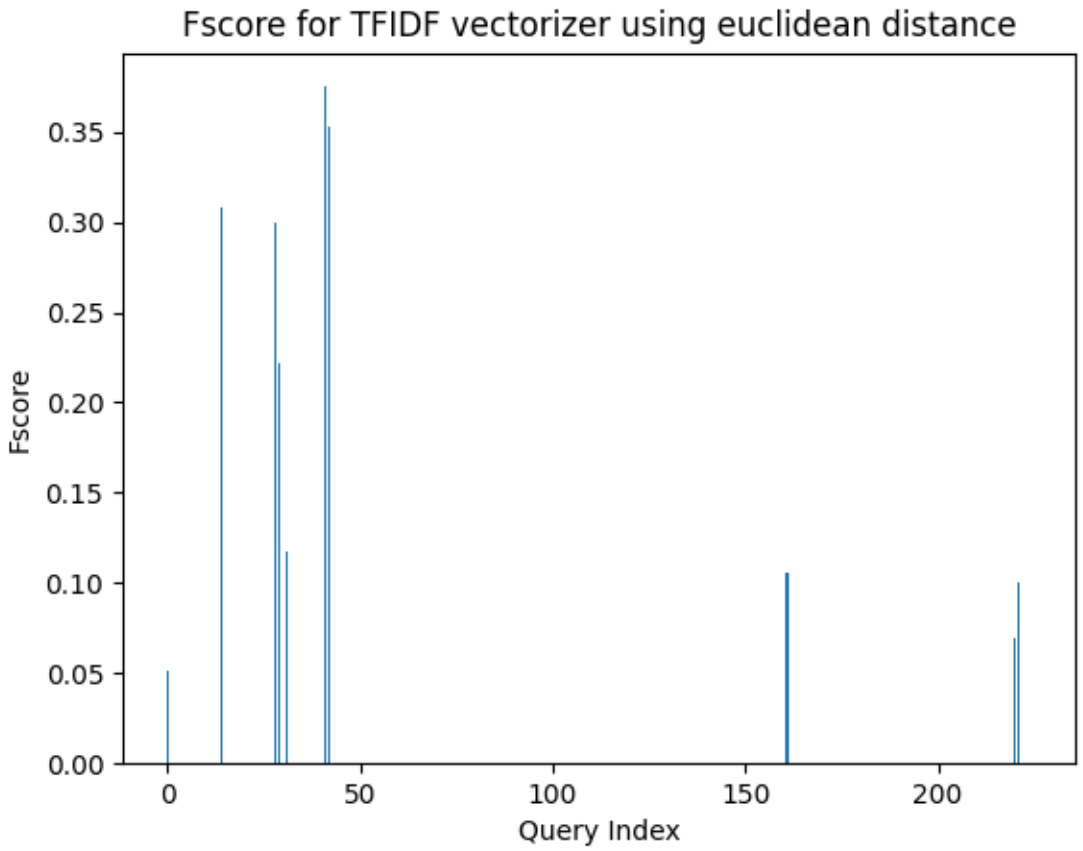
the harmonic mean of recall and precision

Below are the F scores that are obtained


Fscore for Binary vectorizer using cosine distance
Fscore
0.0
0.1
0.2
0.3
0.4
0
50
100
150
200
Query Index


Fscore for Binary vectorizer using euclidean distance
Fscore
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
Query Index

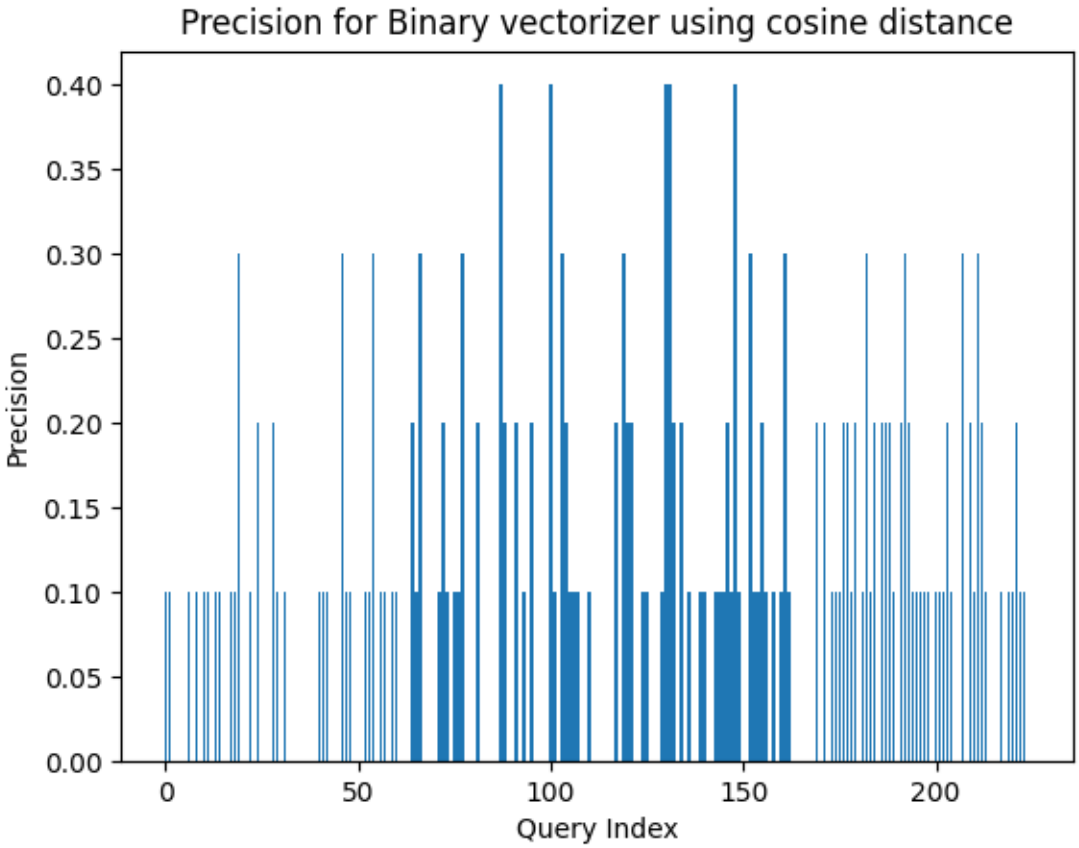

Fscore for TFIDF vectorizer using cosine distance
Fscore
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
50
100
150
200
Query Index

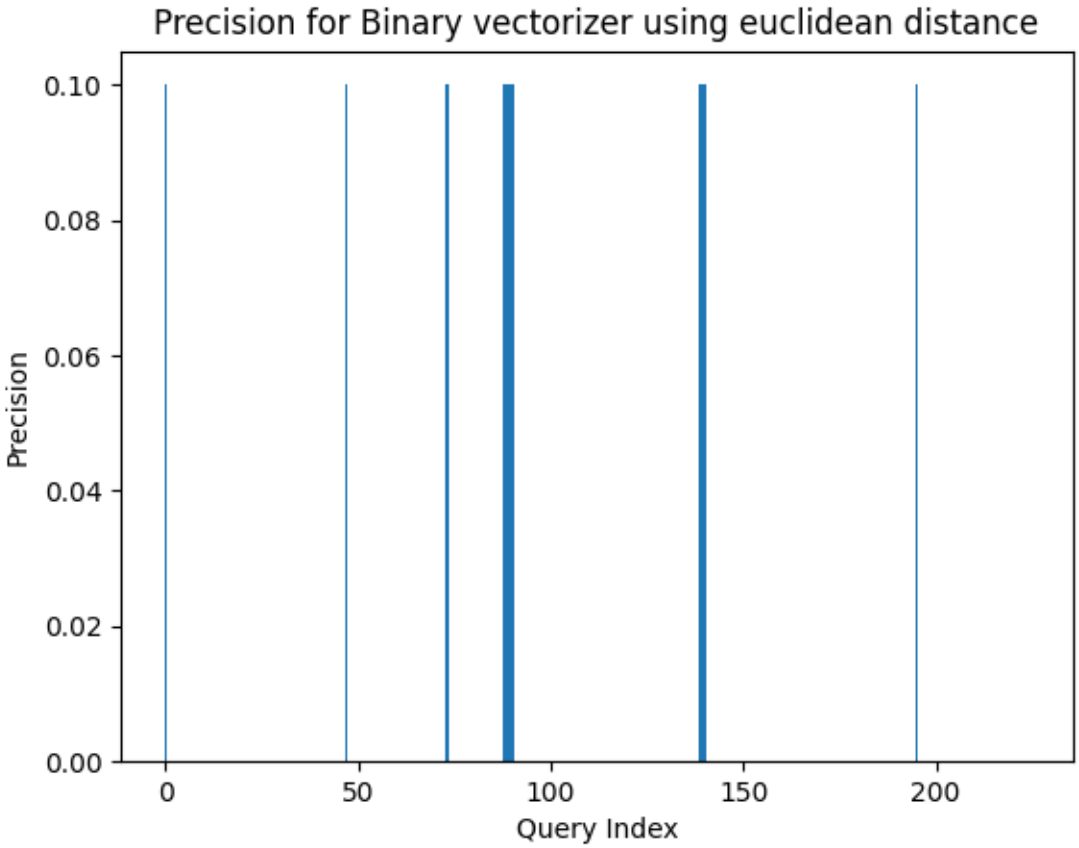


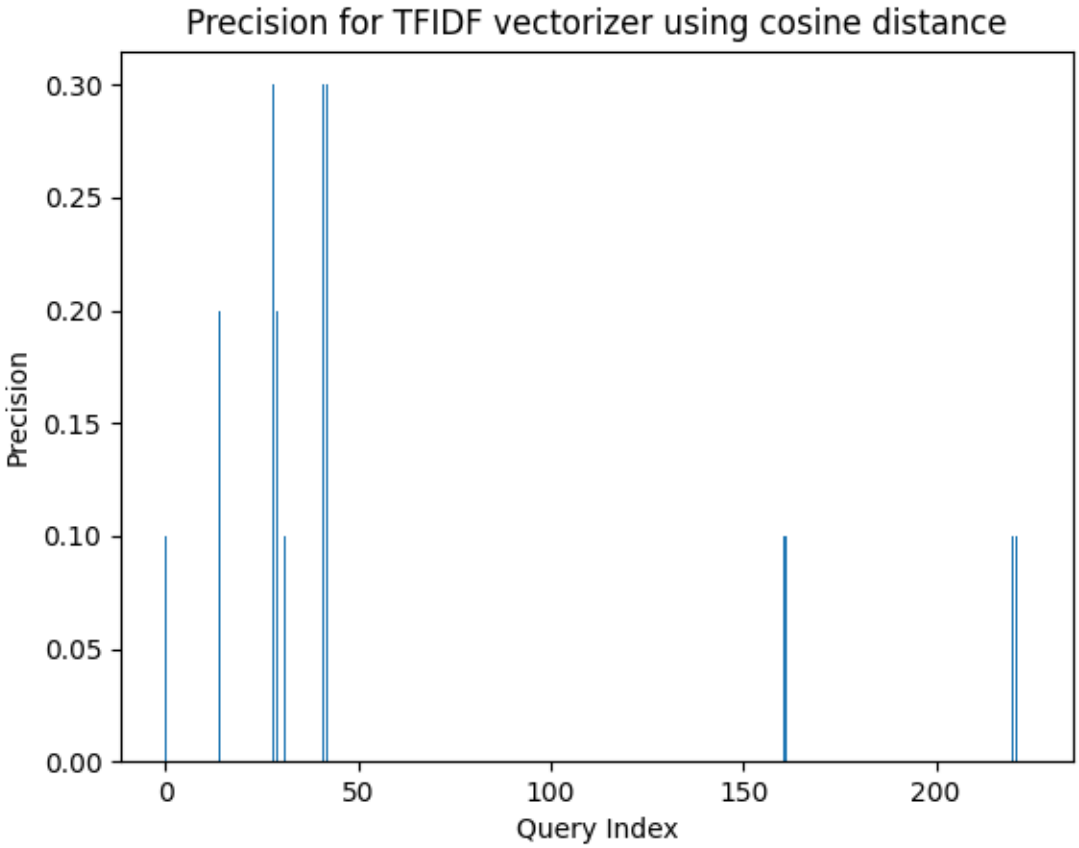
Precision:

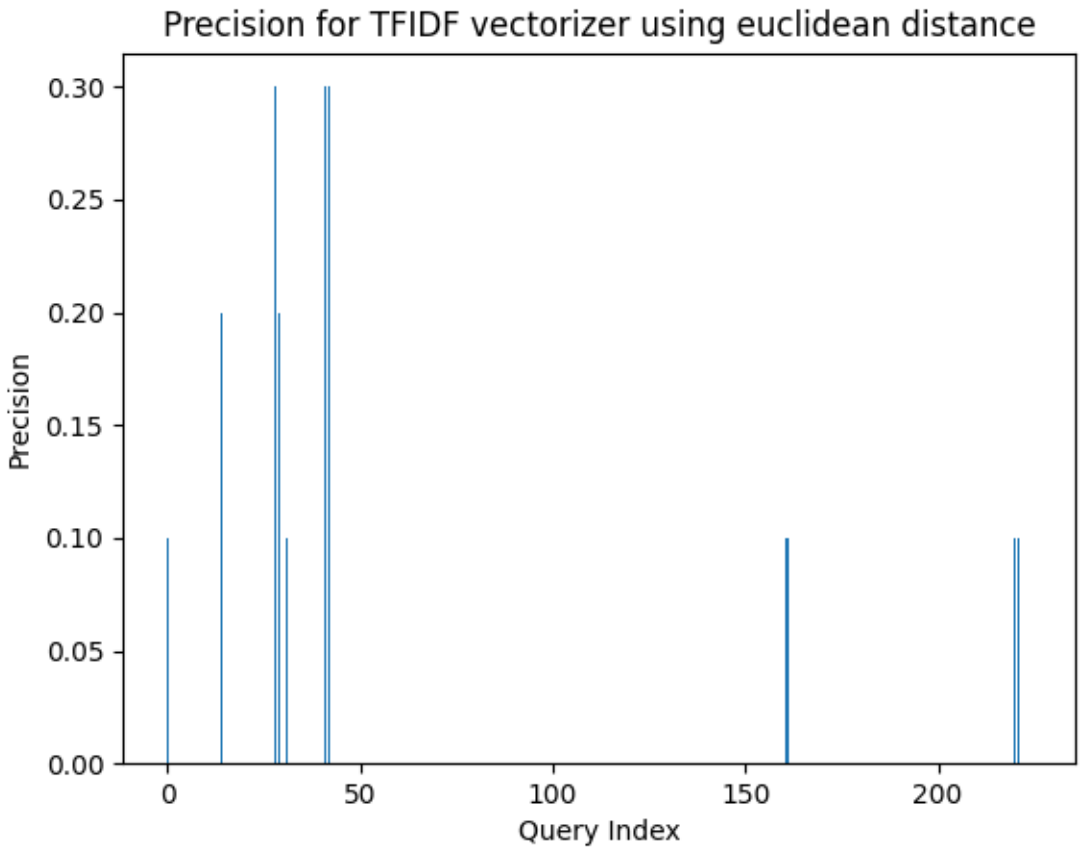
Precision is a metric used to assess the accuracy of positive predictions in classification issues. It is the proportion of actual positive observations to all positive predictions.

Below are the Precision Graphs


Precision for Binary vectorizer using cosine distance
Precision
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
50
100
150
200
Query Index


Precision for Binary vectorizer using euclidean distance
Precision
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
Query Index

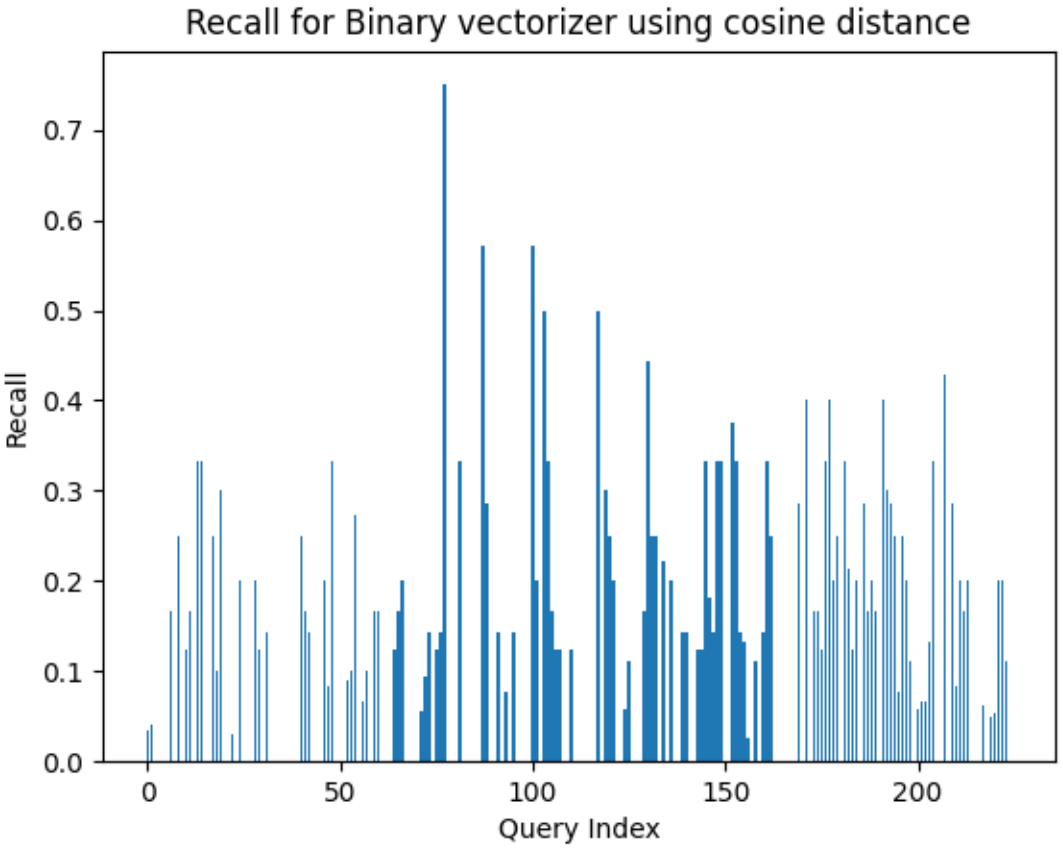

Precision for TFIDF vectorizer using cosine distance
Precision
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
50
100
150
200
Query Index

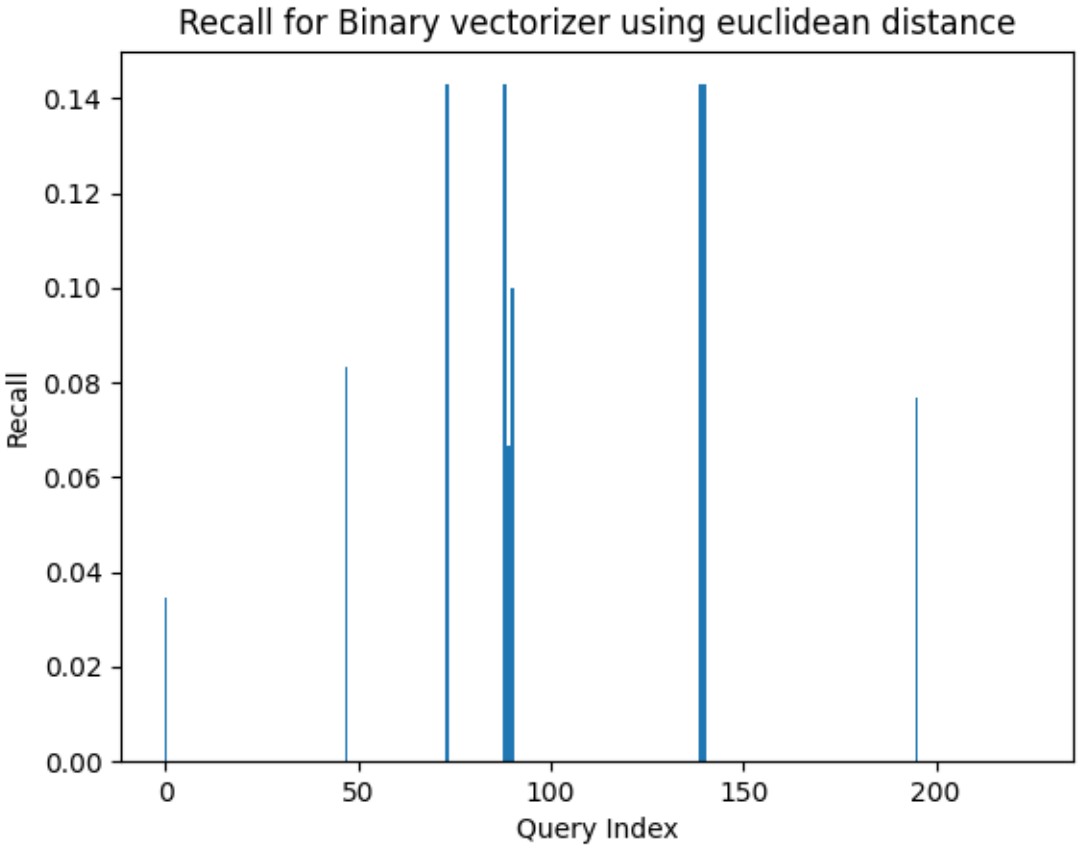


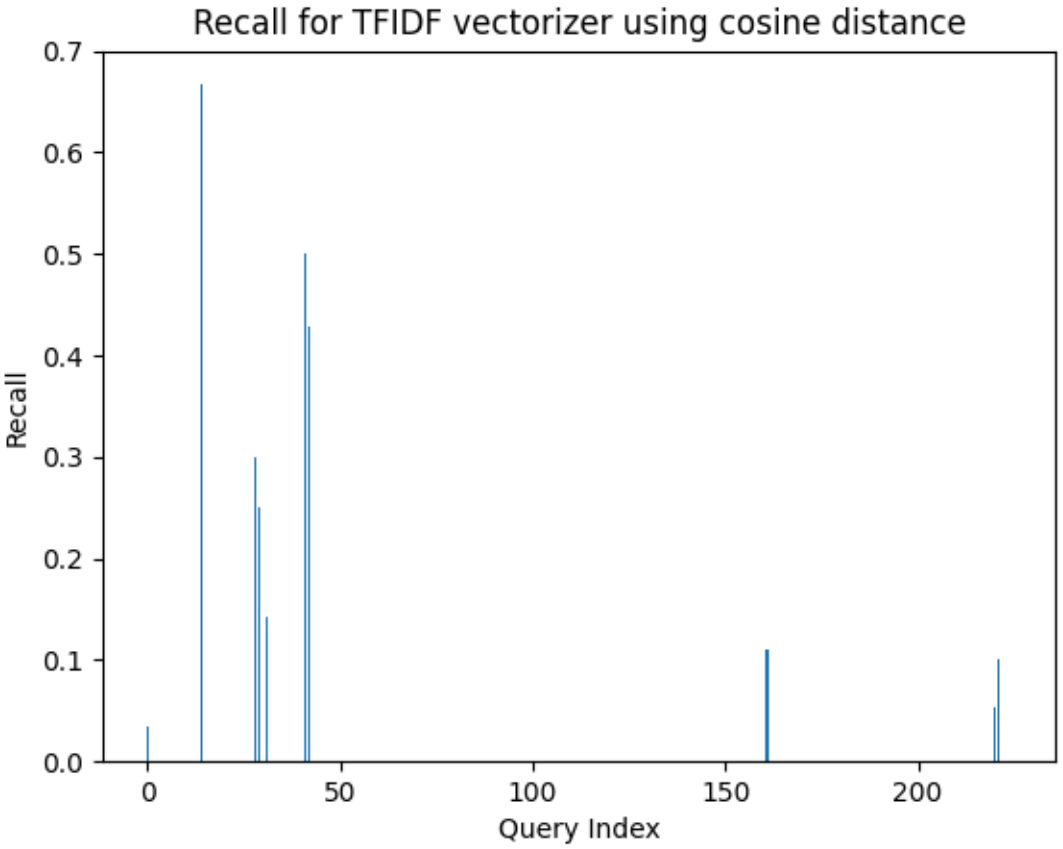
Recall:

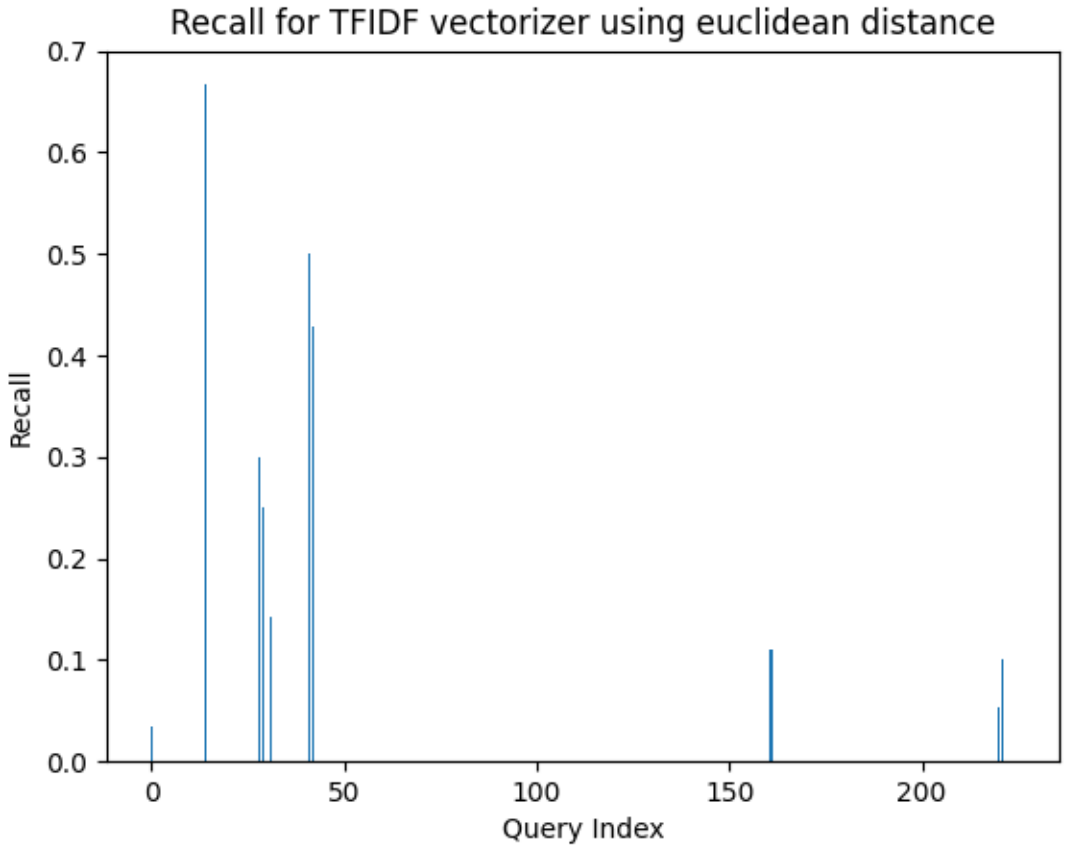
It is the proportion of real positives to all true positive observations. Stated differently, recall indicates the proportion of real positive events that the model properly detected.

Below are the Recall Graphs:

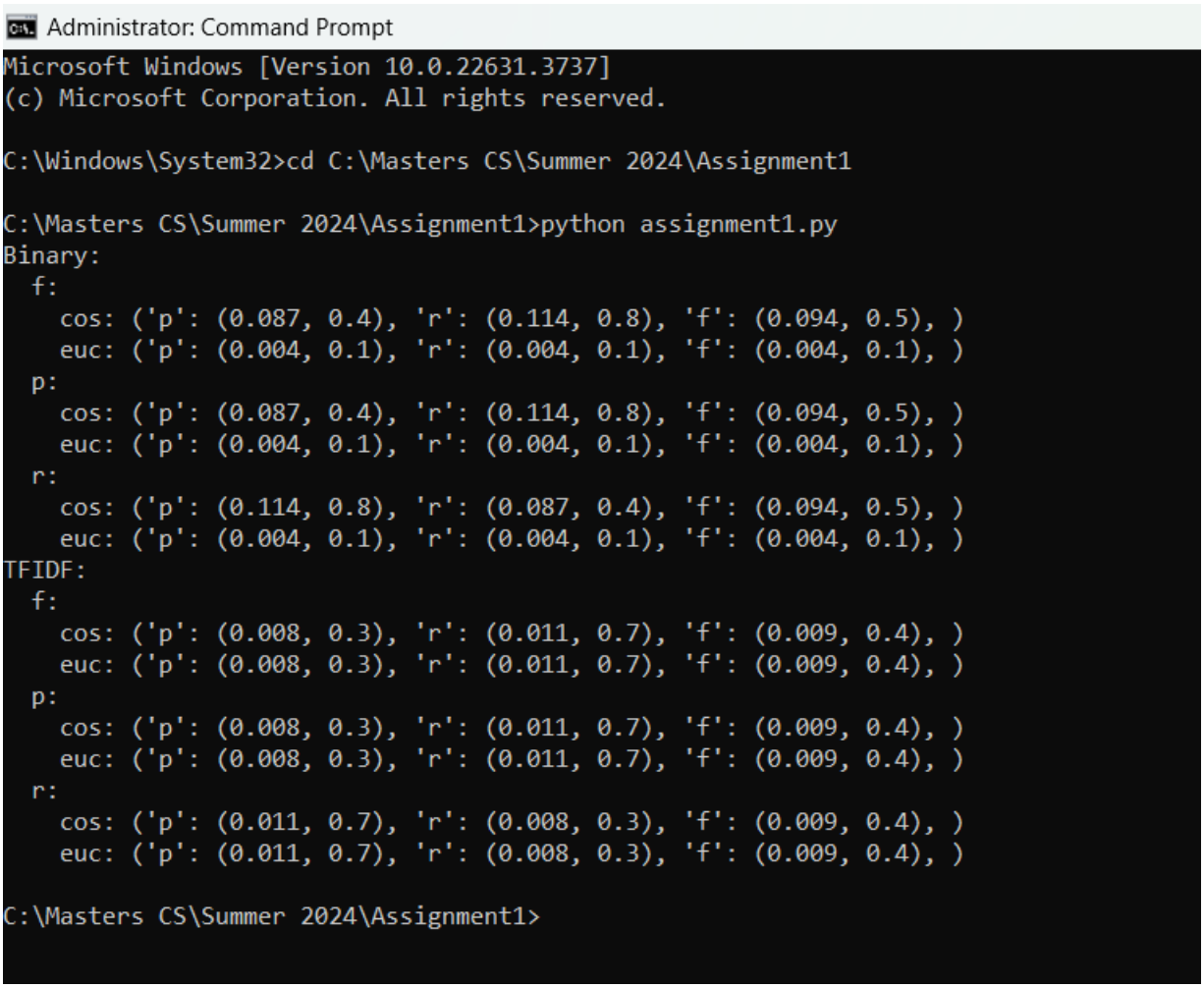

Recall for Binary vectorizer using cosine distance
Recall
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
50
100
150
200
Query Index


Recall for Binary vectorizer using euclidean distance
Recall
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
Query Index


Recall for TFIDF vectorizer using cosine distance
Recall
Query Index
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
50
100
150
200



Below is the Terminal Output that displays all the scores


Administrator: Command Prompt
Microsoft Windows [Version 10.0.22631.3737]
(c) Microsoft Corporation. All rights reserved.
C:\Windows\System32>cd C:\Masters CS\Summer 2024\Assignment1
C:\Masters CS\Summer 2024\Assignment1>python assignment1.py
Binary:
f:
cos: ('p': (0.087, 0.4), 'r': (0.114, 0.8), 'f': (0.094, 0.5),)
euc: ('p': (0.004, 0.1), 'r': (0.004, 0.1), 'f': (0.004, 0.1),)
p:
cos: ('p': (0.087, 0.4), 'r': (0.114, 0.8), 'f': (0.094, 0.5),)
euc: ('p': (0.004, 0.1), 'r': (0.004, 0.1), 'f': (0.004, 0.1),)
r:
cos: ('p': (0.114, 0.8), 'r': (0.087, 0.4), 'f': (0.094, 0.5),)
euc: ('p': (0.004, 0.1), 'r': (0.004, 0.1), 'f': (0.004, 0.1),)
TFIDF:
f:
cos: ('p': (0.008, 0.3), 'r': (0.011, 0.7), 'f': (0.009, 0.4),)
euc: ('p': (0.008, 0.3), 'r': (0.011, 0.7), 'f': (0.009, 0.4),)
p:
cos: ('p': (0.008, 0.3), 'r': (0.011, 0.7), 'f': (0.009, 0.4),)
euc: ('p': (0.008, 0.3), 'r': (0.011, 0.7), 'f': (0.009, 0.4),)
r:
cos: ('p': (0.011, 0.7), 'r': (0.008, 0.3), 'f': (0.009, 0.4),)
euc: ('p': (0.011, 0.7), 'r': (0.008, 0.3), 'f': (0.009, 0.4),)
C:\Masters CS\Summer 2024\Assignment1>