**Plagiarism checking from text files**

● To be done in team of 2

● Write an essay of 10 line on the same topic as text fie (using file functions only))

● Read both text files

● Find frequency of common words between them (ignore case sensitivity)

● % of plagiarism= total frequency of common words/ max(len(file)

**Mother1.txt**

my mother

my mother is everything for me.

My mother is an ordinary woman she is my superhero.

In every step of my, she supported and encouraged me.

Whether day or night she was always there for me no matter what the condition is.

Furthermore, her every work, persistence, devotion, dedication, conduct is an inspiration for me.

In this essay on my mother, I am going to talk about my mother and why she is so special to me.

She is an essence of truthfulness, love, and sincerity. Another reason is that she showers her family with her blessing and live.

Furthermore, she gives us everything but never demand anything in return.

The way she cares for everyone in the family inspires me to the same in my future.

Also, her love is not just for the family she treats every stranger and animals the same way she did to me.

Due to, this she is very kind and sensible towards the environment and animals.

**Mother2.txt**

MY MOTHER

My mother is my god.

She takes care of me a lot.

She cooks very good food.

She always buy me new things whenever i want.

She always comes with me for shopping.

She is very sweet,kind and caring.

My mom is a housewife.

Also she is very fast at her household work.

She is my everything.

I love my mom a lot.

#creating file1 with its content

file1=open("mother1.txt","w")

line=["my mother\n","my mother is everything for me.\n","My mother is an ordinary woman she is my superhero.\n","In every step of my, she supported and encouraged me.\n",

"Whether day or night she was always there for me no matter what the condition is.\n","Furthermore, her every work, persistence, devotion, dedication, conduct is an inspiration for me. \n",

"In this essay on my mother, I am going to talk about my mother and why she is so special to me.\n","She is an essence of truthfulness, love, and sincerity. Another reason is that she showers her family with her blessing and live.\n",

"Furthermore, she gives us everything but never demand anything in return.\n","The way she cares for everyone in the family inspires me to the same in my future.\n",

"Also, her love is not just for the family she treats every stranger and animals the same way she did to me.\n","Due to, this she is very kind and sensible towards the environment and animals."]

file1.writelines(line)

file1.close()

#creating file2 with its content

file2=open("mother2.txt","w")

line=["MY MOTHER\n",

"My mother is my god.\n",

"She takes care of me a lot.\n",

"She cooks very good food.\n",

"She always buy me new things whenever i want.\n",

"She always comes with me for shopping.\n",

"She is very sweet,kind and caring.\n",

"My mom is a housewife.\n",

"Also she is very fast at her household work.\n",

"She is my everything.\n",

"I love my mom a lot."]

file2.writelines(line)

file2.close()

#creating dictionary m1 for file1

file1=open("mother1.txt","r")

m1={}

con=0

for i in file1.read().lower().split():

  con=0

  file2=open("mother1.txt","r")

  for jc in file2.read().lower().split():

    if(i==jc):

      con=con+1

    m1[i]=con

    file2.close()

#creating dictionary m2 for file2

file1=open("mother2.txt","r")

m2={}

con=0

for i in file1.read().lower().split():

  con=0

  file2=open("mother2.txt","r")

  for jc in file2.read().lower().split():

    if(i==jc):

      con=con+1

    m2[i]=con

    file2.close()

#find total frequency of common words

sum=0

for i,j in m1.items():

  for p,k in m2.items():

    if(i==p):

      if(j>k):

        sum=sum+k

#find  maximum length of the file

lenm1=len(m1)

lenm2=len(m2)

pmax=0

if(lenm1>lenm2):

  pmax=lenm1

else:

  pmax=lenm2

  #find plagiarism of both the file.

plg=(sum/pmax)\*100

print("your document have ","{:.2f}".format(plg),"% plagiarism")

**output:-**

your document have 25.26 % plagiarism