

Raport z przeprowadzonych testów  
ze statystyką znalezionych błędów  
**Anty Plagiator 3000**

Mateusz Adamek, Arkadiusz Datka, Filip Gawęł, Szymon Greń  
29 stycznia 2020



**Politechnika  
Śląska**

# 1 Raport błędów

- 1 stycznia 2020 - W tablicy ze wzorami niektóre elementy są puste

**Wiadomość:** "Dostałem następujący błąd: System.IndexOutOfRangeException: Indeks wykraczał poza granice tablicy w klasie Algorytm.cs. Myślę, że jest to spowodowane tym, że niektóre elementy w tablicy ze wzorami są puste.

```
Dane wyjściowe
Pokaż dane wyjściowe z: Debugowanie
Wzory matematyczne:
0 e1: h=\frac{a\sqrt{3}}{2}
1 e1: \lim\left(a_n+b_n\right)=a+b
2 e1: F\left(s\right)=\left\{L\left(s\right)\right\}\left\{L\left(s\right)\right\}=\int_0^{\infty}e^{-st}f\left(t\right)dt
3 e1: x^{2+a}
4 e1: n_{1}=\{n_{2}-n_{3}\}
5 e1: x^2=4
6 e1: \lim\left(a_n-b_n\right)=a-b
-----BAZA PLIKOW:-----
SCIEZKA DO PLIKU:
C:\Users\fig\Documents\GitHub\Skradzione_wzory\Plagiator3000\TEXfiles\copy1\1.tex
WZORY Z PLIKU:
Wzory matematyczne:
0 e1: n_{k+1}
1 e1: a^{\frac{m}{n}}=\sqrt[n]{a^m}
2 e1: U(g_1-g_2)=k\cdot\sqrt{[u(g_1)]^2+[u(g_2)]^2}
3 e1: h=\frac{a\sqrt{3}}{2}
4 e1:
5 e1:
6 e1: F=q\left(E+v\right)
Wzor oryginalny: h=\frac{a\sqrt{3}}{2}
Wzor plagiatu: n_{k+1}
SAME OR NOT----- ???
6,08276253029822
Wzor oryginalny: h=\frac{a\sqrt{3}}{2}
Wzor plagiatu: a^{\frac{m}{n}}=\sqrt[n]{a^m}
SAME OR NOT----- ???
3,74165738677394
Wzor oryginalny: h=\frac{a\sqrt{3}}{2}
Wzor plagiatu: U(g_1-g_2)=k\cdot\sqrt{[u(g_1)]^2+[u(g_2)]^2}
SAME OR NOT----- ???
5,29150262212918
Wzor oryginalny: h=\frac{a\sqrt{3}}{2}
Wzor plagiatu: h=\frac{a\sqrt{3}}{2}
SAME OR NOT----- ???
0
Wzor oryginalny: h=\frac{a\sqrt{3}}{2}
Wzor plagiatu:
„Plagiator3000.exe” (CLR v4.0.30319: Plagiator3000.exe): załadowano „C:\Windows\Microsoft.Net\assembly\GAC_MSIL\mscorlib.resources
Zgłoszony wyjątek: „System.IndexOutOfRangeException” w Plagiator3000.exe
Wystąpił nieobsługiwany wyjątek typu „System.IndexOutOfRangeException” w Plagiator3000.exe
Indeks wykraczał poza granice tablicy.
```

Widok w konsoli

```
private static Dictionary<char, int> CountFrequentlyOfLetters(char[] letters)
{
    var freqLet = new Dictionary<char, int>();
    Array.Sort(letters);
    var actualChar = letters[0];
    var sum = 0;
    foreach (var c in letters)
    {
        if (c == actualChar)
        {
            sum++;
        }
        else
        {
            freqLet.Add(actualChar, sum);
            actualChar = c;
            sum = 1;
        }
    }
    freqLet.Add(actualChar, sum);
    return freqLet;
}
```

Nieobsługiwany wyjątek

System.IndexOutOfRangeException: „Indeks wykraczał poza granice tablicy.”

Wyświetl szczegóły | Kopiuj szczegóły | Rozpocznij sesję rozszerzenia Live Share

Ustawienia wyjątków

Widok w Klasie Algorytm.cs

**Rozwiązanie:** Błędem okazało się złe wczytywanie wzorów. Na kilka możliwości wczytywać dwie działały nieprawidłowo. Możliwości które działały złe to wzory umieszczane między "(" a ")" oraz "[" i "]". Problemem okazało się złe zamykanie. To jest program widział kiedy otworzyć wczytywanie jednak po pierwszym znaku \ zamykał. Stąd, gdy w pliku latexa wzór zaczynał się od znaku "" program automatycznie zamykał wczytywanie wzoru przez co zostawał on pusty. Do naprawy wystarczyło dodać zmienną bool, która zmienia wartość na true, gdy po znaku "" następuje znak "]" lub "]". Tylko wtedy pozwala programowi opuścić wczytywanie wzorów.

- 28 stycznia 2020 – Wynik algorytmu cosine distance czasami wynosi NaN

**Wiadomość:** “Czasami, gdy użyję algorytmu cosine distance w raporcie latex oraz html dostaje wynik NaN.”

$n_1 = n_2 - n_3$	0,151973505030525
$n_1 = n_2 - n_3$	0,854135008502105
$x^2 = 4$	NaN
$x^2 = 4$	0,4

## Raport ogólny porównania plików

Plik bazowy: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\1.tex

---

Plik: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1\1.tex

Stopień podobieństwa:

NaN

---

Plik: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1\2.tex

Stopień podobieństwa:

0,377184476253723

---

Plik: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1\3.tex

Stopień podobieństwa:

0,346877724898933

---

Plik: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1\4.tex

Stopień podobieństwa:

NaN

---

Plik: C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1\5.tex

Stopień podobieństwa:

NaN

---

**Rozwiązanie:** „Błąd występował, gdy w algorytmie zachodziło dzielenie przez 0. Algorytm dzieli przez 0 tylko, gdy wzór w plagiacie nie posiada ani jednej litery, która występuje w oryginalnym wzorze. Dodanie warunku rozwiązało problem.”

## 2 Przeprowadzone testy

- Algorytmy:
  - Cosine distance

Po wprowadzeniu identycznych danych jak na stronie internetowej:  
<https://neo4j.com/docs/graph-algorithms/current/labs-algorithms/cosine/>  
otrzymaliśmy identyczny wynik.

Strona internetowa:

We can use it to compute the similarity of two hardcoded lists.

The following will return the cosine similarity of two lists of numbers:

```
RETURN algo.similarity.cosine([3,8,7,5,2,9], [10,8,6,6,4,5]) AS similarity
```

Table 9.97. Results

similarity
0.8638935626791596

These two lists of numbers have a Cosine similarity of 0.863. We can see how this result is derived by breaking down the formula:

$$\left( \text{similarity}(A,B) = \frac{3 \cdot 10 + 8 \cdot 8 + 7 \cdot 6 + 5 \cdot 6 + 2 \cdot 4 + 9 \cdot 5}{\sqrt{3^2 + 8^2 + 7^2 + 5^2 + 2^2 + 9^2} \times \sqrt{10^2 + 8^2 + 6^2 + 6^2 + 4^2 + 5^2}} = \frac{219}{15.2315 \times 16.6433} = 0.8639 \right)$$

Nasz program:

### Raport szczegółowy porównania plików

Plik bazowy:

C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\Testy\1.tex

Plik: C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\Testy\copy\1.tex

Wzór	Jest podobny do wzoru oryginalnego	Procent podobieństwa
aaaaaaaaabbbbbbbccccccddddddeeeffffff	aaabbbbbbbccccccddddeeffffffff	86,389356267916

Wynik testu pozytywny

## ■ Euclidean distance

Po wprowadzeniu identycznych danych uzyskaliśmy prawidłowy wynik.

Strona internetowa

Features	Coord1	Coord2	Coord3	Coord4	Coord5	Coord6
Object A	1	3	4	5	2	4
Object B	7	6	3	1	3	1

Get Euclidean Distance

Object A coordinate = {1,3,4,5,2,4}

Object B coordinate = {7,6,3,1,3,1}

You inputted 6 dimensions

The Euclidean Distance is 8.485This program is presented to you by [@ Kardi Teknomo](#)

Nasz program:

## Raport szczegółowy porównania plików

Plik bazowy:

C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\Testy\1.tex

Plik: C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\Testy\copy\1.tex

Wzór	Jest podobny do wzoru oryginalnego	Procent podobieństwa
<i>aaaaaaabbbbccccdeef</i>	<i>abbbccccdddeeffff</i>	80,9978148228733

Nasz wynik wynosi: 80,9978148228733, jest to procent podobieństwa. Sam algorytm euclidean distance zwraca dystans z przedziału od 0 do nieskończoności. Wyzaczyliśmy maksymalną granicę poprzez "przepuszczenie" algorytmu przez całą bazę danych. MaxOfEuclidean wynosi dla naszej bazy 44.65423. Procent podobieństwa wyznaczony jest wg następującego kodu:

lb - wartość zwrócona przez algorytm

```
double pom = (double) (maxOfEuclidean - lb);
result = Scale(pom, 0.0, maxOfEuclidean, 0.0, 1.0);

private static double Scale(double value, double min, double max, double toMin,
double toMax)
{ return (value - min) / (max - min) * (toMax - toMin) + toMin; }
```

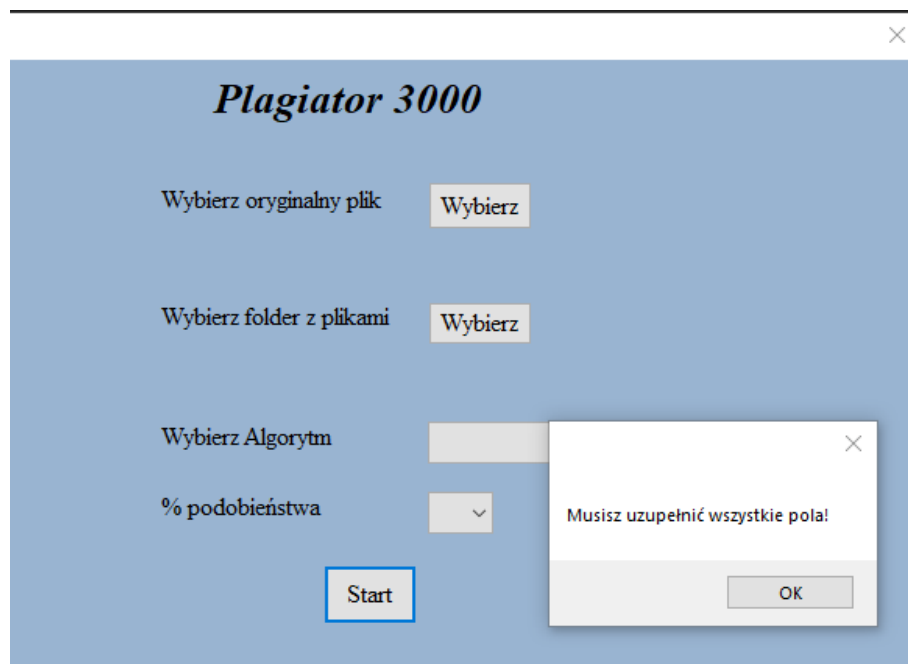
$pom = 44.65423 - 8.485 = 36.16923$

$result = (36.16923 - 0) / (44.65423 - 0) * (1 - 0) + 0$

$result = 0.80998$

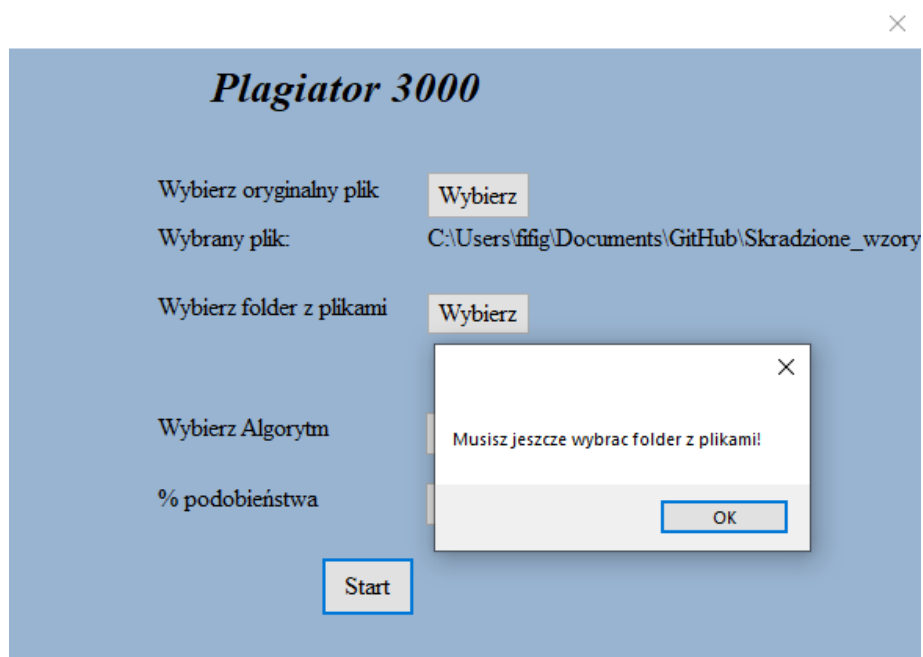
Wynik testu pozytywny

- Oprogramowanie:
  - Brak jakiegokolwiek wyboru



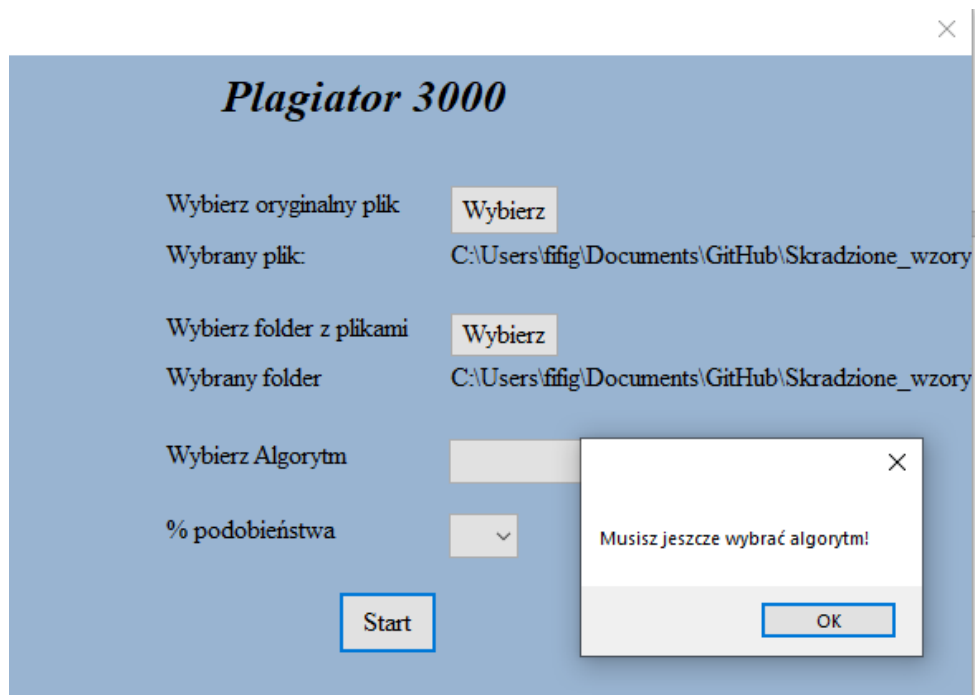
Wynik testu pozytywny

- Wybór tylko oryginalnego pliku



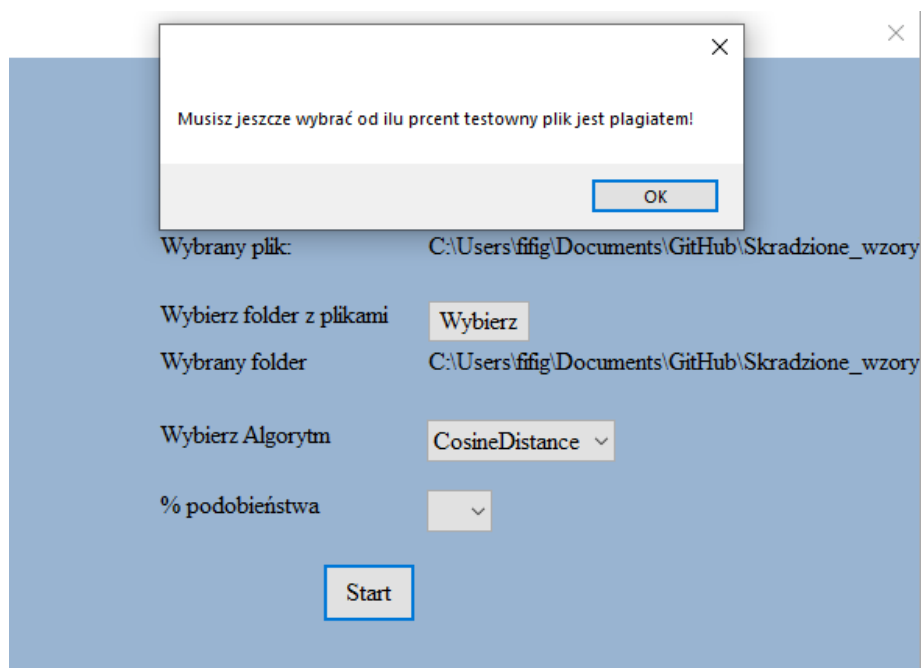
Wynik testu pozytywny

- Wybór oryginalnego pliku oraz folderu z plikami podejrzanymi



Wynik testu pozytywny

- Wybór oryginalnego pliku, folderu z plikami podejrzanymi oraz algorytmu



Wynik testu pozytywny



- Wybór niezgodnego pliku oryginalnego

×

## Plagiator 3000

Wybierz oryginalny plik

Wybrany plik: C:\Users\ffig\Desktop\glow.JPG

Wybierz folder z plikami

Wybrany folder C:\Users\ffig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\copy1

Wybierz Algorytm

% podobieństwa

```

Wzor oryginalny: y=???N?!!?????k?+?L+I=
U?u?z?V?e?5JgR^?g????U?z?K?E?.?W????1?Y?C?=?[?X?z?~%?v????\#????i????<O?A????????XJ????K?!!?+????? ???Q?????'x????? ?&?????
?x??W??????z?MA'??A)??9?????uZ??Z)Vz=#tj?h?0??z?S?????S?D?+?c????W?U~?~?Dr|K^L<{?e?????Q????w?,o!!???O?G????K R-??z?Oo?
Wzor plagiatu: f\left(a\right)=\frac{1}{2}\pi\oint\frac{f\left(z\right)}{z-a}dz
SAME OR NOT----- ???
2,95865168966233

Wzor oryginalny: y=???N?!!?????k?+?L+I=
U?u?z?V?e?5JgR^?g????U?z?K?E?.?W????1?Y?C?=?[?X?z?~%?v????\#????i????<O?A????????XJ????K?!!?+????? ???Q?????'x????? ?&?????
?x??W??????z?MA'??A)??9?????uZ??Z)Vz=#tj?h?0??z?S?????S?D?+?c????W?U~?~?Dr|K^L<{?e?????Q????w?,o!!???O?G????K R-??z?Oo?
Wzor plagiatu: n_{k+1}
SAME OR NOT----- ???
9,30975008652817

Wzor oryginalny: y=???N?!!?????k?+?L+I=
U?u?z?V?e?5JgR^?g????U?z?K?E?.?W????1?Y?C?=?[?X?z?~%?v????\#????i????<O?A????????XJ????K?!!?+????? ???Q?????'x????? ?&?????
?x??W??????z?MA'??A)??9?????uZ??Z)Vz=#tj?h?0??z?S?????S?D?+?c????W?U~?~?Dr|K^L<{?e?????Q????w?,o!!???O?G????K R-??z?Oo?
Wzor plagiatu: \log_{a}(x\cdot y)=\log_{a}x+\log_{a}y
SAME OR NOT----- ???
3,43091238257779

Wzor oryginalny: y=???N?!!?????k?+?L+I=
U?u?z?V?e?5JgR^?g????U?z?K?E?.?W????1?Y?C?=?[?X?z?~%?v????\#????i????<O?A????????XJ????K?!!?+????? ???Q?????'x????? ?&?????
?x??W??????z?MA'??A)??9?????uZ??Z)Vz=#tj?h?0??z?S?????S?D?+?c????W?U~?~?Dr|K^L<{?e?????Q????w?,o!!???O?G????K R-??z?Oo?
Wzor plagiatu: \int_0^{2\pi}dx
SAME OR NOT----- ???
6,25274733648442

Wzor oryginalny: y=???N?!!?????k?+?L+I=
U?u?z?V?e?5JgR^?g????U?z?K?E?.?W????1?Y?C?=?[?X?z?~%?v????\#????i????<O?A????????XJ????K?!!?+????? ???Q?????'x????? ?&?????
?x??W??????z?MA'??A)??9?????uZ??Z)Vz=#tj?h?0??z?S?????S?D?+?c????W?U~?~?Dr|K^L<{?e?????Q????w?,o!!???O?G????K R-??z?Oo?
Wzor plagiatu: h=\frac{a\sqrt{3}}{2}
SAME OR NOT----- ???
5,57843144725572
-----KONIEC PLIKU-----
Program „[12460] Plagiator3000.exe” zakończył działanie z kodem -1 (0xffffffff).

```

Program nie kończy pracę. Nie tworzy raportu.

Wynik testu negatywny

- Wybór niezgodnego folderu z plikami podejrzаными

***Plagiator 3000***

Wybierz oryginalny plik	<input type="button" value="Wybierz"/>
Wybrany plik:	C:\Users\fifig\Documents\GitHub\Skradzione_wzory\Plagiator3000\TEXfiles\1.tex
Wybierz folder z plikami	<input type="button" value="Wybierz"/>
Wybrany folder	C:\Users\fifig\Desktop\graphs
Wybierz Algorytm	EuclideanDistan <input type="button" value="v"/>
% podobieństwa	93 <input type="button" value="v"/>
<input type="button" value="Start"/>	

## Raport szczegółowy porównania plików

**Plik bazowy:**

**C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\1.tex**

## Raport ogólny porównania plików

**Plik bazowy:**

**C:\Users\fifig\Documents\GitHub\Skradzione\_wzory\Plagiator3000\TEXfiles\1.tex**

---

Program kończy pracę. Tworzy raport bez jakichkolwiek danych.

Wynik testu pozytywny