Mathematics 3
Distance Measures for Machine Learning

Name
Muhammad Baihaqi Aulia Asy'ari

NIM
2241720145

Class
2I

Department
Information Technology

Study Program
D4 Informatics Engineering

Task 1

```python
def hammingDistance(n1,n2) :

  x = n1 ^ n2
  setBits = 0

  while (x > 0) :
    setBits += x & 1
    x >>= 1

  return setBits

if __name__ == '__main__' :
  n1 = 9
  n2 = 14
  print(hammingDistance(9, 14))
```

```
3
```

$$
\begin{aligned}
9 &= 1001 \\
14 &= 1110 \\
&= 0111 \\
&= 3
\end{aligned}
$$

Task 2

```python
def hammingDistance(x, y) :
  ans = 0
  m = max(x, y)
  while (m):
    c1 = x & 1
    c2 = y & 1
    if (c1 != c2):
      ans += 1
    m = m >> 1
    x = x >> 1
    y = y >> 1
  return ans

n1 = 4
n2 = 8
hdist = hammingDistance(n1, n2)
print(hdist)
```

```
2
```

Additional Task

1. count the hamming distance of these

   BEEN
BEAN
0010

   CEREAL
SERIAL
100100

   hamming distance of BEEN and BEAN is 1
   hamming distance of CEREAL and SERIAL is 2

2. hamming distance from 10 and 15 (binary)

$$\begin{aligned} 10 &= 1010 \\ 15 &= 1111 \\ &= 0101 \\ &= 2 \end{aligned}$$

3. hamming distance from 6 and 11 (binary)

$$\begin{aligned} 6 &= 0110 \\ 11 &= 1011 \\ &= 1101 \\ &= 3 \end{aligned}$$

Task 3

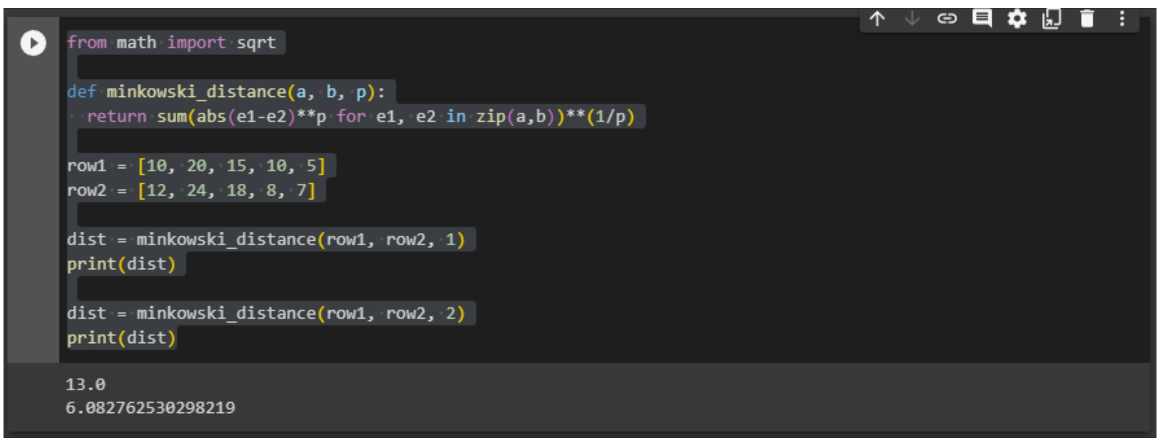
```python
from math import sqrt

def minkowski_distance(a, b, p):
return sum(abs(e1-e2)**p for e1, e2 in zip(a,b))**(1/p)

row1 = [10, 20, 15, 10, 5]
row2 = [12, 24, 18, 8, 7]

dist = minkowski_distance(row1, row2, 1)
print(dist)

dist = minkowski_distance(row1, row2, 2)
print(dist)
```



p = 1

$$
\begin{aligned}
d(x,y) &= (|10-12|^1 + |20-24|^1 + |15-18|^1 + |10-8|^1 + |5-7|^1)^{(1/1)} \\
&= (|2|^1 + |4|^1 + |3|^1 + |2|^1 + |2|^1)^1 \\
&= (2+4+3+2+2) \\
&= 13
\end{aligned}
$$

p = 2

$$
\begin{aligned}
d(x,y) &= (|10-12|^2 + |20-24|^2 + |15-18|^2 + |10-8|^2 + |5-7|^2)^{(1/2)} \\
&= \sqrt{(|2|^2 + |4|^2 + |3|^2 + |2|^2 + |2|^2)} \\
&= \sqrt{(4+16+9+4+4)} \\
&= \sqrt{37} = 6.082
\end{aligned}
$$

Task 4

A Resume on Minkowski and Chebyshev Distance

The Minkowski distance is a mathematical tool used to measure how far apart two sets of numbers are from each other. It's like a flexible ruler that can be adjusted based on a special number called "order" or p. This allows us to calculate distances in different ways depending on the situation. So, Minkowski distance is a helpful tool in math and data analysis that lets us customize how we measure distances to fit our needs.

Chebyshev's distance is a metric used to quantify the dissimilarity between two points in a grid or vector space. It measures the maximum absolute difference between corresponding coordinates of these points it provides a straightforward way to gauge the "worst-case scenario" for movement between two points within a multi-dimensional space.

Minkowski's mathematical equation goes as follows

$$\sqrt[r]{\sum_{k=1}^{n} |x_k - y_k|^r}$$

Chebyshev's mathematical equation goes as follows

$$\lim_{q \to \infty} \sqrt[r]{\sum_{i=1}^{n} |x_i - y_i|^q}$$

Minkowski distance plays a vital role in portfolio diversification. Minkowski distance helps with this by measuring their similarity. A larger distance indicates that these assets have distinct risk and return characteristics, making them suitable for diversification.

In the world of financial security, Chebyshev distance is a useful tool for making fraud detection better. This metric allows us to pinpoint potential irregularities by measuring the most significant deviation between transaction attributes. When integrated into machine learning models and real-time monitoring systems, it aids in swiftly flagging transactions that fall beyond predefined thresholds.