Original Article

Application of Fuzzy Sets for Taking Optional Subjects in Any Graduation based on Marks Obtained in SSC and Intermediate those who are Having Uncertainty

SVS. Chaitanya Varma[1], B. Nageswara Rao[2]

[1]*Department of Computer Science Engineering, Lendi Institute of Engineering and Technology, Jonnada, Vizinagaram (Dist.) A.P, India.*
[2]*Department of Mathematics, Lendi Institute of Engineering and Technology, Jonnada, Vizinagaram (Dist.) A.P, India.*

Corresponding Author : chaitanyasirivuri@gmail.com

Abstract - This paper focuses on addressing the issue of uncertainty that students face when choosing optional subjects for their higher studies in various graduation programs such as B.A, BCom, B. Sc, and B.Tech. We propose the use of the Hamming distance and Euclidean distance between fuzzy sets to assist students in making informed decisions based on their marks obtained in SSC and intermediate exams. The Hamming distance is a measurement that quantifies the dissimilarity between two fuzzy sets. By utilizing this distance measure, we aim to evaluate the discrepancy between the subjects chosen by students and the subjects recommended based on their academic performance. This approach takes into account the uncertainty inherent in the decision-making process, enabling students to make choices that align better with their academic strengths and interests. By combining these distance measures, the paper aims to provide a comprehensive framework for students to evaluate their options for higher studies. The proposed approach considers both the uncertainty associated with subject selection and the academic performance of the students. Ultimately, this methodology assists students in making more informed decisions regarding their optional subjects, thereby facilitating a smoother transition to their chosen graduate program.

Keywords - Membership function, Fuzzy set, Fuzzy values.

AIMS Mathematics Subject Classification: 03G10, 06B23, 0C21

1. Introduction

Fuzzy set theory has gained significant attention in recent years due to its ability to handle uncertainty and imprecision in real-world problems. Fuzzy sets are a mathematical tool used to represent uncertain or vague information that is difficult to express using classical set theory. This research paper explores the application of fuzzy sets to analyze students' academic performance and interests. This research's primary objective is to analyze students' performance and interests by constructing fuzzy sets from their grade 10, grade 12, and undergraduate mark lists. We focused on three subjects, namely mathematics, science, and English for grades 10 and 12, and core subjects for undergraduate studies. The fuzzy sets were named S, I, and B for grade 10, grade 12, and undergraduate studies; respectively, we calculated the membership values of the three subjects for each student and constructed fuzzy sets. Then, we measured the distance between the fuzzy sets of each student using the Hamming and Euclidean distance metrics. The Hamming distance is a measure of the dissimilarity between two fuzzy sets, while the Euclidean distance measures the distance between two points in a multi-dimensional space.

Finally, we analyzed the results to determine the students' interests and performance in their majors. The research aims to provide valuable insights into the usefulness of fuzzy sets in analyzing academic performance and interests. The findings could help educators and academic advisors to better understand the student's interests and performance and guide them toward their desired academic path. We have converted the student's grades into membership degrees and constructed fuzzy sets from their academic grades of SSC (S), Intermediate (I), and B. Tech (B). From SSC and Intermediate, we have taken Mathematics (M), Science(S), and English(E) into consideration, and from B. Tech, we considered Python Programming (PP), Database Management Systems (DBMS), and Digital Logic Design (DLD). In this paper, we applied the Hamming distance, the Euclidean distance between the fuzzy sets for taking the optional subject of any graduation (B.A, B.Com, B.Sc., B. Tech) based on marks obtained in SSC and intermediate those are having uncertainty for choosing the option for higher studies. The marks obtained in the subjects of Mathematics, Science, and English of ten students in SSC are converted to membership values of fuzzy sets.

2. Preliminaries

We give here a review of some definitions and results which are in Gau.w.L and Buhere. D.j [3], Nageswararao. B, Ramakrishana. N and Eswarlal.T[6],[7],[8],[9].

Definition [14] 2.1: Let X be a non-empty set. A fuzzy set A drawn from X is defined as A={x, μ$_A$(x): x ∈ X }, where μ $_A$ (x): X→ [0, 1] is the membership function of the fuzzy set A

The membership degrees are calculated using membership functions with $\mu_A(X)$ where X is considered the subject.

Definition [13] 2.2: The Hamming distance between the fuzzy sets can be defined by

$$d(A,B) = \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|, \ where \ x_i \in X$$

Definition [12] 2.3: The Euclidean distance between the fuzzy sets can be defined by

$$d(A,B) = \sum_{i=1}^{n} \sqrt{(\mu_A(x_i))^2 - (\mu_B(x_i))^2}, \ where \ x_i \in X$$

The marks obtained in the subjects of Mathematics, Science and English of ten students in SSC are converted to membership values of fuzzy sets.

3. Fuzzy values of the Marks obtained in SSC

The following table (3.1) shows the membership values of fuzzy sets of the ten students who obtained their SSC marks for the subjects of Mathematics, Science, and English.

Table 3.1

Name	Mathematics	Science	English
K. Purandhareswari	0.7	0.8	0.8
T. Lokesh	0.9	0.9	0.9
T. Neelima	0.9	0.9	0.9
K. Nitish Kumandan	0.9	0.8	0.9
T. Devanshu Kumar	0.9	0.8	0.9
K. Sai Vamsi	0.9	0.8	0.9
P. Sasank Shanmuk	0.6	0.7	0.8
D. Bhavya Sri	0.7	0.7	0.7
P.V.N.L. Srinija	0.9	0.9	0.9
P. Teja	0.7	0.6	0.8

3.2. Marks obtained in Intermediate

The following table (3.2) shows the membership values of fuzzy sets of the ten students who have obtained their IPE marks for the subjects of Mathematics, Physics and Chemistry, and English

Name	Mathematics	Science(P &C)	English
K. Purandhareswari	0.9	0.9	0.9
T. Lokesh	0.8	0.8	0.9
T. Neelima	0.9	0.9	0.9
K. Nitish Kumandan	0.5	0.5	0.3
T. Devanshu Kumar	0.9	0.8	0.8
K. Sai Vamsi	0.7	0.5	0.7
P. Sasank Shanmuk	0.5	0.4	0.8
D. Bhavya Sri	0.8	0.8	0.9
P.V.N.L. Srinija	0.7	0.9	0.9
P. Teja	0.9	0.9	0.9

3.3. Marks obtained in B. Tech
The following table (3.3) shows the membership values of fuzzy sets of the ten students who obtained their B-Tech marks for the subjects of Python, DBMS, and DLD.

Name	Python	DBMS	DLD
K. Purandhareswari	0.8	0.7	0.8
T. Lokesh	0.8	0.7	0.7
T. Neelima	0.9	0.8	0.9
K. Nitish Kumandan	0.3	0.3	0.3
T. Devanshu Kumar	0.8	0.8	0.8
K. Sai Vamsi	0.8	0.7	0.6
P. Sasank Shanmuk	0.6	0.7	0.8
D. Bhavya Sri	0.6	0.6	0.5
P.V.N.L. Srinija	0.8	0.8	0.8
P. Teja	0.8	0.7	0.9

3.4. Distances between marks of SSC (S) & Intermediate (I)
The following table (3.4) shows the distance between the membership values of fuzzy sets of ten students who obtained their SSC and IPE marks for the subjects of Mathematics, Physics and chemistry, and English.

1. Hamming Distance, Euclidean Distance of membership function of K. Purandhareswari:

$$S = (M, 0.7), (S, 0.8), (E, 0.8)$$

$$I = (M, 0.9), (S, 0.9), (E, 0.9)$$

$$B = (PP, 0.8), (DBMS, 0.7), (DLD, 0.8)$$

(a) Hamming Distance:

$$d(S, I) = |0.7 - 0.9| + |0.8 - 0.9| + |0.8 - 0.9| = 0.2 + 0.1 + 0.1 = 0.4$$

$$d(S, B) = |0.7 - 0.8| + |0.8 - 0.7| + |0.8 - 0.8| = 0.1 + 0.1 + 0 = 0.2$$

$$d(I, B) = |0.9 - 0.8| + |0.9 - 0.7| + |0.9 - 0.8| = 0.1 + 0.2 + 0.1 = 0.4$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.7 - 0.9)^2 + (0.8 - 0.9)^2 + (0.8 - 0.9)^2} = \sqrt{(0.2)^2 + (0.1)^2 + (0.1)^2} = 0.245$$

$$d(S, B) = \sqrt{(0.7 - 0.8)^2 + (0.8 - 0.7)^2 + (0.8 - 0.8)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0)^2} = 0.141$$

$$d(I, B) = \sqrt{(0.9 - 0.8)^2 + (0.9 - 0.7)^2 + (0.9 - 0.8)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0.1)^2} = 0.245$$

2. Hamming Distance, Euclidean Distance of membership function of T. Lokesh

$$I = (M, 0.8), (S, 0.8), (E, 0.9)$$

$$B = (PP, 0.8), (DBMS, 0.7), (DLD, 0.7)$$

(a) Hamming Distance:

$$d(S, I) = |0.9 - 0.8| + |0.9 - 0.8| + |0.9 - 0.9| = 0.1 + 0.1 + 0 = 0.2$$

$$d(S, B) = |0.9 - 0.8| + |0.9 - 0.7| + |0.9 - 0.7| = 0.1 + 0.2 + 0.2 = 0.5$$

$$d(I, B) = |0.8 - 0.8| + |0.8 - 0.7| + |0.9 - 0.7| = 0 + 0.1 + 0.2 = 0.3$$

(b) Euclidean Distance:

$$d(S,I) = \sqrt{(0.9-0.8)^2 + (0.9-0.8)^2 + (0.9-0.9)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0)^2} = 0.141$$

$$d(S,B) = \sqrt{(0.9-0.8)^2 + (0.9-0.7)^2 + (0.9-0.7)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0.2)^2} = 0.3$$

$$d(I,B) = \sqrt{(0.8-0.8)^2 + (0.8-0.7)^2 + (0.9-0.7)^2} = \sqrt{(0)^2 + (0.1)^2 + (0.2)^2} = 0.223$$

3. Hamming Distance, Euclidean Distance of membership function of T. Neelima

$$S = (M, 0.9), (S, 0.9), (E, 0.9)$$

$$I = (M, 0.9), (S, 0.9), (E, 0.9)$$

$$B = (PP, 0.9), (DBMS, 0.8), (DLD, 0.9)$$

(a) Hamming Distance:

$$d(S,I) = |0.9-0.9| + |0.9-0.9| + |0.9-0.9| = 0 + 0 + 0 = 0$$

$$d(S,B) = |0.9-0.9| + |0.9-0.8| + |0.9-0.9| = 0 + 0.1 + 0 = 0.1$$

$$d(I,B) = |0.9-0.9| + |0.9-0.8| + |0.9-0.9| = 0 + 0.1 + 0 = 0.1$$

(b) Euclidean Distance:

$$d(S,I) = \sqrt{(0.9-0.9)^2 + (0.9-0.9)^2 + (0.9-0.9)^2} = \sqrt{(0)^2 + (0)^2 + (0)^2} = 0$$

$$d(S,B) = \sqrt{(0.9-0.9)^2 + (0.9-0.8)^2 + (0.9-0.9)^2} = \sqrt{(0)^2 + (0.1)^2 + (0)^2} = 0.1$$

$$d(I,B) = \sqrt{(0.9-0.9)^2 + (0.9-0.8)^2 + (0.9-0.9)^2} = \sqrt{(0)^2 + (0.1)^2 + (0)^2} = 0.1$$

4. Hamming Distance, Euclidean Distance of membership function of K. Nitish

$$S = (M, 0.9), (S, 0.8), (E, 0.9)$$

$$I = (M, 0.5), (S, 0.5), (E, 0.3)$$

$$B = (PP, 0.3), (DBMS, 0.3), (DLD, 0.3)$$

(a) Hamming Distance:

$$d(S,I) = |0.9-0.5| + |0.8-0.5| + |0.9-0.3| = 0.4 + 0.3 + 0.6 = 1.3$$

$$d(S,B) = |0.9-0.3| + |0.8-0.3| + |0.9-0.3| = 0.6 + 0.5 + 0.6 = 1.7$$

$$d(I,B) = |0.5-0.3| + |0.5-0.3| + |0.3-0.3| = 0.2 + 0.2 + 0 = 0.4$$

(b) Euclidean Distance:

$$d(S,I) = \sqrt{(0.9-0.5)^2 + (0.8-0.5)^2 + (0.9-0.3)^2} = \sqrt{(0.4)^2 + (0.3)^2 + (0.6)^2} = 0.78$$

$$d(S,B) = \sqrt{(0.9-0.3)^2 + (0.8-0.3)^2 + (0.9-0.3)^2} = \sqrt{(0.6)^2 + (0.5)^2 + (0.6)^2} = 0.98$$

$$d(I,B) = \sqrt{(0.5-0.3)^2 + (0.5-0.3)^2 + (0.3-0.3)^2} = \sqrt{(0.2)^2 + (0.2)^2 + (0)^2} = 0.28$$

5. Hamming Distance, Euclidean Distance of membership function of T. Devanshu Kumar

$$S = (M, 0.8), (S, 0.6), (E, 0.9)$$

$$I = (M, 0.9), (S, 0.8), (E, 0.8)$$

$$B = (PP, 0.8), (DBMS, 0.8), (DLD, 0.8)$$

(a) Hamming Distance:

$$d(S, I) = |0.8 - 0.9| + |0.6 - 0.8| + |0.9 - 0.8| = 0.1 + 0.2 + 0.1 = 0.4$$

$$d(S, B) = |0.8 - 0.8| + |0.6 - 0.8| + |0.9 - 0.8| = 0 + 0.2 + 0.1 = 0.3$$

$$d(I, B) = |0.9 - 0.8| + |0.8 - 0.8| + |0.8 - 0.8| = 0.1 + 0 + 0 = 0.1$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.8 - 0.9)^2 + (0.6 - 0.8)^2 + (0.9 - 0.8)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0.1)^2} = 0.245$$

$$d(S, B) = \sqrt{(0.8 - 0.8)^2 + (0.6 - 0.8)^2 + (0.9 - 0.8)^2} = \sqrt{(0)^2 + (0.2)^2 + (0.1)^2} = 0.223$$

$$d(I, B) = \sqrt{(0.9 - 0.8)^2 + (0.8 - 0.8)^2 + (0.8 - 0.8)^2} = \sqrt{(0.1)^2 + (0)^2 + (0)^2} = 0.1$$

6. Hamming Distance, Euclidean Distance of membership function of K. Sai Vamsi

$$S = (M, 0.9), (S, 0.8), (E, 0.9)$$

$$I = (M, 0.7), (S, 0.5), (E, 0.7)$$

$$B = (PP, 0.8), (DBMS, 0.7), (DLD, 0.6)$$

(a) Hamming Distance:

$$d(S, I) = |0.9 - 0.7| + |0.8 - 0.5| + |0.9 - 0.7| = 0.2 + 0.3 + 0.2 = 0.7$$

$$d(S, B) = |0.9 - 0.8| + |0.8 - 0.7| + |0.9 - 0.6| = 0.1 + 0.1 + 0.3 = 0.5$$

$$d(I, B) = |0.7 - 0.8| + |0.5 - 0.7| + |0.7 - 0.6| = 0.1 + 0.2 + 0.1 = 0.4$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.9 - 0.7)^2 + (0.8 - 0.5)^2 + (0.9 - 0.7)^2} = \sqrt{(0.2)^2 + (0.3)^2 + (0.2)^2} = 0.412$$

$$d(S, B) = \sqrt{(0.9 - 0.8)^2 + (0.8 - 0.7)^2 + (0.9 - 0.6)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.3)^2} = 0.332$$

$$d(I, B) = \sqrt{(0.7 - 0.8)^2 + (0.5 - 0.7)^2 + (0.7 - 0.6)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0.1)^2} = 0.244$$

7. Hamming Distance, Euclidean Distance of membership function of P.S. Sasank

$$S = (M, 0.6), (S, 0.7), (E, 0.8)$$

$$I = (M, 0.5), (S, 0.4), (E, 0.5)$$

$$B = (PP, 0.5), (DBMS, 0.5), (DLD, 0.6)$$

(a) Hamming Distance:

$$d(S, I) = |0.6 - 0.5| + |0.7 - 0.4| + |0.8 - 0.5| = 0.1 + 0.3 + 0.3 = 0.7$$

$$d(S, B) = |0.6 - 0.5| + |0.7 - 0.5| + |0.8 - 0.6| = 0.1 + 0.2 + 0.2 = 0.5$$

$$d(I, B) = |0.5 - 0.5| + |0.4 - 0.5| + |0.5 - 0.6| = 0 + 0.1 + 0.1 = 0.2$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.6 - 0.5)^2 + (0.7 - 0.4)^2 + (0.8 - 0.5)^2} = \sqrt{(0.1)^2 + (0.3)^2 + (0.3)^2} = 0.436$$

$$d(S, B) = \sqrt{(0.6 - 0.5)^2 + (0.7 - 0.5)^2 + (0.8 - 0.6)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0.2)^2} = 0.3$$

$$d(I, B) = \sqrt{(0.5 - 0.5)^2 + (0.4 - 0.5)^2 + (0.5 - 0.6)^2} = \sqrt{(0)^2 + (0.1)^2 + (0.1)^2} = 0.141$$

8. Hamming Distance, Euclidean Distance of membership function of D. Bhavya Sri

$$S = (M, 0.7), (S, 0.7), (E, 0.7)$$

$$I = (M, 0.8), (S, 0.8), (E, 0.9)$$

$$B = (PP, 0.6), (DBMS, 0.6), (DLD, 0.5)$$

(a) Hamming Distance:

$$d(S, I) = |0.7 - 0.8| + |0.7 - 0.8| + |0.7 - 0.9| = 0.1 + 0.1 + 0.2 = 0.4$$

$$d(S, B) = |0.7 - 0.6| + |0.7 - 0.6| + |0.7 - 0.5| = 0.1 + 0.1 + 0.2 = 0.4$$

$$d(I, B) = |0.8 - 0.6| + |0.8 - 0.6| + |0.9 - 0.5| = 0.2 + 0.2 + 0.4 = 0.8$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.7 - 0.8)^2 + (0.7 - 0.8)^2 + (0.7 - 0.9)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.2)^2} = 0.245$$

$$d(S, B) = \sqrt{(0.7 - 0.6)^2 + (0.7 - 0.6)^2 + (0.7 - 0.5)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.2)^2} = 0.245$$

$$d(I, B) = \sqrt{(0.8 - 0.6)^2 + (0.8 - 0.6)^2 + (0.9 - 0.5)^2} = \sqrt{(0.2)^2 + (0.2)^2 + (0.4)^2} = 0.49$$

9. Hamming Distance, Euclidean Distance of membership function of P.V.N.L. Srinija

$$S = (M, 0.9), (S, 0.9), (E, 0.9)$$

$$I = (M, 0.7), (S, 0.9), (E, 0.9)$$

$$B = (PP, 0.8), (DBMS, 0.8), (DLD, 0.8)$$

(a) Hamming Distance:

$$d(S, I) = |0.9 - 0.7| + |0.9 - 0.9| + |0.9 - 0.9| = 0.2 + 0 + 0 = 0.2$$

$$d(S, B) = |0.9 - 0.8| + |0.9 - 0.8| + |0.9 - 0.8| = 0.1 + 0.1 + 0.1 = 0.3$$

$$d(I, B) = |0.7 - 0.8| + |0.9 - 0.8| + |0.9 - 0.8| = 0.1 + 0.1 + 0.1 = 0.3$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.9 - 0.7)^2 + (0.9 - 0.9)^2 + (0.9 - 0.9)^2} = \sqrt{(0.2)^2 + (0)^2 + (0)^2} = 0.2$$

$$d(S, B) = \sqrt{(0.9 - 0.8)^2 + (0.9 - 0.8)^2 + (0.9 - 0.8)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.1)^2} = 0.173$$

$$d(I, B) = \sqrt{(0.7 - 0.8)^2 + (0.9 - 0.8)^2 + (0.9 - 0.8)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.1)^2} = 0.173$$

10. Hamming Distance, Euclidean Distance of membership function of P. Teja

$$S = (M, 0.7), (S, 0.6), (E, 0.8)$$

$$I = (M, 0.9), (S, 0.9), (E, 0.9)$$

$$B = (PP, 0.8), (DBMS, 0.7), (DLD, 0.9)$$

(a) Hamming Distance:

$$d(S, I) = |0.7 - 0.9| + |0.6 - 0.9| + |0.8 - 0.9| = 0.2 + 0.3 + 0.1 = 0.6$$

$$d(S, B) = |0.7 - 0.8| + |0.6 - 0.7| + |0.8 - 0.9| = 0.1 + 0.1 + 0.1 = 0.3$$

$$d(I, B) = |0.9 - 0.8| + |0.9 - 0.7| + |0.9 - 0.9| = 0.1 + 0.2 + 0 = 0.3$$

(b) Euclidean Distance:

$$d(S, I) = \sqrt{(0.7 - 0.9)^2 + (0.6 - 0.9)^2 + (0.8 - 0.9)^2} = \sqrt{(0.2)^2 + (0.3)^2 + (0.1)^2} = 0.374$$

$$d(S, B) = \sqrt{(0.7 - 0.8)^2 + (0.6 - 0.7)^2 + (0.8 - 0.9)^2} = \sqrt{(0.1)^2 + (0.1)^2 + (0.1)^2} = 0.173$$

$$d(I, B) = \sqrt{(0.9 - 0.8)^2 + (0.9 - 0.7)^2 + (0.9 - 0.9)^2} = \sqrt{(0.1)^2 + (0.2)^2 + (0)^2} = 0.223$$

Table 3.5. Average of hamming and euclidean distances from all combinations

Name	Hamming Distance	Euclidean Distance
K. Purandhareswari	0.333	0.21
T. Lokesh	0.333	0.222
T. Neelima	0.067	0.067
K. Nitish Kumandan	1.133	0.683
T. Devanshu Kumar	0.133	0.114
K. Sai Vamsi	0.533	0.33
P. Sasank Shanmuk	0.267	0.211
D. Bhavya Sri	0.533	0.327
P.V.N.L. Srinija	0.267	0.182
P. Teja	0.4	0.257

4. Algorithm

Input: Student marks in terms of Fuzzy sets A and B

Output: Student performance according to their scores

1. Define the function euclidean_distance(a, b) that takes two lists (a and b) as input and calculates the Euclidean distance between them.
2. Define the function hamming_distance(a, b) that takes two lists (a and b) as input and calculates the Hamming distance between them.
3. Define the function difference(a, b) that takes two numbers (a and b) as input and calculates their difference.
4. Read the CSV files using pd.read_csv() and assign them to the respective variables: mark_ssc, mark_inter, mark_btech, dist_si, dist_sb, and dist_ib.
5. Iterate over the rows of mark_ssc using a for a loop.
   5.1. Calculate the Hamming distance between the marks in mark_ssc and mark_inter for the current row using the hamming_distance() function. Store the result in the corresponding row and column of dist_si.
   5.2. Calculate the Euclidean distance between the marks in mark_ssc and mark_inter for the current row using the euclidean_distance() function. Store the result in the corresponding row and column of dist_si.
   5.3. Calculate the Hamming distance between the marks in mark_ssc and mark_btech for the current row using the hamming_distance() function. Store the result in the corresponding row and column of dist_sb.
   5.4. Calculate the Euclidean distance between the marks in mark_ssc and mark_btech for the current row using the euclidean_distance() function. Store the result in the corresponding row and column of dist_sb.
   5.5. Calculate the Hamming distance between the marks in mark_inter and mark_btech for the current row using the hamming_distance() function. Store the result in the corresponding row and column of dist_ib.
   5.6. Calculate the Euclidean distance between the marks in mark_inter and mark_btech for the current row using the euclidean_distance() function. Store the result in the corresponding row and column of dist_ib.
6. Save the modified dist_si, dist_sb, and dist_ib DataFrames to CSV files.
7. Read the CSV file avg_distances.csv and assign it to the variable avg_distances.
8. Iterate over the rows of avg_distances using a for loop.
   8.1. Calculate the average Hamming distance for the current row by adding the corresponding values from dist_si, dist_sb, and dist_ib and dividing by 3. Store the result in the corresponding row and column of avg_distances.

8.2. Calculate the average Euclidean distance for the current row by adding the corresponding values from dist_si, dist_sb, and dist_ib and dividing by 3. Store the result in the corresponding row and column of avg_distances.
9. Iterate over the rows of avg_distances using a for loop.
   9.1. Check if the marks in mark_ssc, mark_inter, and mark_btech for the current row meet the condition of being greater than or equal to 0.7.
   9.2. Check if the difference between the average Hamming distance and the average Euclidean distance for the current row is less than or equal to 0.

5. Source Code

The provided code implements this research analysis in Python using various distance metrics to assess the academic performance of students. This code consists of Data Preparation, Distance Calculation and Storage, Average Distance Calculation, Success Determination, Results and Conclusion Overall, this implementation utilizes distance metrics to evaluate academic performance and determines success based on defined criteria. It provides insights into students' performance and their alignment with recommended subjects.

```python
import pandas as pd

from decimal import Decimal

def euclidean_distance(a, b):

    sum = 0

    for i in range(len(a)):

        sum += (Decimal(a[i]) - Decimal(b[i]))**2

    return round((sum.sqrt()), 3)

def hamming_distance(a, b):

    sum = 0

    for i in range(len(a)):

        sum += abs(Decimal(a[i]) - Decimal(b[i]))

    return round((sum), 3)

def difference(a, b):

    return round((Decimal(a) - Decimal(b)), 3)

mark_ssc = pd.read_csv('mark_ssc.csv')

mark_inter = pd.read_csv('mark_inter.csv')

mark_btech = pd.read_csv('mark_btech.csv')

dist_si = pd.read_csv('dist_si.csv')

dist_sb = pd.read_csv('dist_sb.csv')

dist_ib = pd.read_csv('dist_ib.csv')

for i in range(len(mark_ssc)):

    dist_si.loc[i, "Hamming Distance"] = hamming_distance(
```

```python
        mark_ssc.iloc[i, 1:].tolist(), mark_inter.iloc[i, 1:].tolist())
    dist_si.loc[i, "Euclidean Distance"] = euclidean_distance(
        mark_ssc.iloc[i, 1:].tolist(), mark_inter.iloc[i, 1:].tolist())
    dist_sb.loc[i, "Hamming Distance"] = hamming_distance(
        mark_ssc.iloc[i, 1:].tolist(), mark_btech.iloc[i, 1:].tolist())
    dist_sb.loc[i, "Euclidean Distance"] = euclidean_distance(
        mark_ssc.iloc[i, 1:].tolist(), mark_btech.iloc[i, 1:].tolist())
    dist_ib.loc[i, "Hamming Distance"] = hamming_distance(
        mark_inter.iloc[i, 1:].tolist(), mark_btech.iloc[i, 1:].tolist())
    dist_ib.loc[i, "Euclidean Distance"] = euclidean_distance(
        mark_inter.iloc[i, 1:].tolist(), mark_btech.iloc[i, 1:].tolist())
dist_ib.to_csv('dist_ib.csv', index=False)
dist_sb.to_csv('dist_sb.csv', index=False)
dist_si.to_csv('dist_si.csv', index=False)
avg_distnaces = pd.read_csv('avg_distances.csv')
for i in range(len(avg_distnaces)):
    avg_distnaces.loc[i, "Hamming Distance"] = round(
        (dist_si.iloc[i, 1] + dist_sb.iloc[i, 1] + dist_ib.iloc[i, 1])/3, 3)
    avg_distnaces.loc[i, "Euclidean Distance"] = round(
        (dist_si.iloc[i, 2] + dist_sb.iloc[i, 2] + dist_ib.iloc[i, 2])/3, 3)
for i in range(len(avg_distnaces)):
    if mark_ssc.iloc[i, 1] >= 0.7 and mark_inter.iloc[i, 1] >= 0.7 and mark_btech.iloc[i, 1] >= 0.7:
        if difference(avg_distnaces.iloc[i, 1], avg_distnaces.iloc[i, 2]) <= 0.1:
            avg_distnaces.loc[i, "Success"] = True
        else:
            avg_distnaces.loc[i, "Success"] = False
    else:
        avg_distnaces.loc[i, "Success"] = False
```

```
avg_distnaces.to_csv('avg_distances.csv', index=False)

print(avg_distnaces)

for i in range(len(avg_distnaces)):

    if avg_distnaces.iloc[i, 3] == True:

        print(

            f"{mark_ssc.iloc[i, 0]} has performed consistently well in all their
academics.")
```

6. Conclusion

Based on our observations and calculations, we have found that the least absolute difference between the Hammonton distance and the Euclidean distance serves as an indicator of a student's success. In other words, when the difference between these distances is smaller, it suggests that the student is more likely to excel in their chosen course of study.

A smaller difference between the Hammonton and Euclidean distances signifies a closer alignment between the subjects chosen by the student and the subjects recommended based on their academic performance. This alignment indicates that the student's selected subjects are well-matched with their strengths and interests, enhancing their chances of success.

On the other hand, when the difference between these distances is larger, it implies a greater complexity for the student to catch up with the pace of the course. A larger difference indicates a significant discrepancy between the subjects chosen by the student and the recommended subjects based on their academic performance. This misalignment suggests that the student may face challenges in keeping up with the coursework due to a lack of foundational knowledge or aptitude in the chosen subjects.

Therefore, our conclusion is that a lower absolute difference between the Hammonton and Euclidean distances is indicative of a more successful outcome for the student. It suggests a better fit between the student's subject choices and their academic performance, while a higher difference signifies a greater complexity and potential difficulties for the student in their chosen course of study

1. The student T. Neelima has obtained a path difference of 0, indicating that there is no significant difference between the Hammonton distance and the Euclidean distance for her chosen subjects. This suggests a high level of compatibility between the subjects she has selected and her academic performance. As a result, it is predicted that T. Neelima will have a successful future in both academics and research or a technical career. The close alignment between her subject choices and academic performance indicates that she is likely to excel and thrive in her chosen field.

2. T. Devanshu Kumar, with a path difference of 0.02 in both the Hammonton and Euclidean distance methods, exhibits a close alignment between the subjects chosen and academic performance. This indicates a strong compatibility and suggests a higher likelihood of success in academics, research, and a technical career The small path difference signifies that the subjects chosen by T. Devanshu Kumar are well-matched with their strengths and interests, providing a solid foundation for future success. It suggests that he has made appropriate subject choices based on their academic performance, positioning them favourably to excel in their chosen field.

3. The student P.V.N.L. Srinija has obtained a path difference of 0.085 in both the Hammonton and Euclidean distance methods. Although there is a slightly higher difference between the distances compared to the previous example, it is still relatively small. This indicates a reasonably good match between the subjects chosen by P.V.N.L. Srinija and her academic performance. As a result, she is also expected to have a successful future in academics, research, and a technical career. While there may be some minor complexity or challenges to overcome, the overall alignment suggests that she has made suitable subject choices based on her strengths and interests.

In all cases, the path values close to zero indicate strong compatibility between the student's subject choices and their academic performance. This alignment is a positive indicator of their future success in academics, research, and technical careers. However, it is important to note that while these path values suggest a higher likelihood of success, individual factors and circumstances can still influence their future outcomes.

References

[1] Ali H. Ganie, "A Picture Fuzzy Distance Measure and its Applications to Pattern Recognition Problem," *Irani Journal of Fuzzy Systems*, vol. 20, no. 1, pp. 71-85, 2023. [Google Scholar] [Publisher Link]

[2] Azriel Rosenfeld, "Fuzzy Groups," *Journal of Mathematical Analysis and Applications*, vol. 35, pp. 512-517, 1971. [Google Scholar] [Publisher Link]

[3] B. Nageswara Rao, N. Ramakrishana, and T. Eswarlal, "L-Vague Semi Rings of L-Semi Ring," *International Journal of Mathematics Trends and Technology*, vol. 57, no. 5, pp. 345-351, 2018. [CrossRef] [Publisher Link]

[4] B. Nageswara Rao, N. Ramakrishana, and T. Eswarlal, "Application of Vague Sets in the Carrier of the Students for their Opportunities," *A Journal of Composition Theory*, vol. 15, no. 12, pp. 181-187, 2022. [Publisher Link]

[5] B.Nageswara Rao, N. Ramakrishana, and T. Eswarlal, "Application of Translates of Elect Vague Sets on Elctoral Democracy Model," *International Journal of Chemical Sciences (Sadguru Publications)*, ISSN 0972-768X, pp. 762-768, (Impact Factor- 0.7, scoupus Indexed).

[6] B. Nageswara Rao, N. Ramakrishna, and T. Eswarlal, "Application of Vague sets in Carrier Decision Making," *International Journal of Chemical Sciences (Sadguru publications)*, ISSN:0972-768X, pp. 372-380, 2016.

[7] N. Ramakrishna, "Vague Groups and Vague Weights," *International Journal of Computational Cognition*, vol. 66 no.1, pp. 913-916, 2009.

[8] Ranjit Biswas, "Vague Groups," *International Journal of Computational Cognition*, vol. 4, no. 2, pp. 20-23, 2006.

[9] R.R. Zhao et al., "A Parametric Similarity Measure between Fuzzy Sets and Its Applications in Multi-attribute Decision Making," *Iranian Journal of Fuzzy Systems*, vol. 20, no. 1, pp. 87-102, 2023. [Google Scholar] [Publisher Link]

[10] S. Saghi, and A. Nazemi, "Simplex Algorithm for Hesitant Fuzzy Linear Problem with Hesitant Cost Coefficient," *Irani Journal of Fuzzy Systems*, vol. 20, no. 2, pp. 137-152, 2023.

[11] S.R. Islam, and M. Pal, "Second Zagreb Index Fuzzy Sets and Its Application Mathematical Chemistry," *Iranian Journal of Fuzzy Systems*, vol. 20, no. 1, pp. 119-136, 2023. [Google Scholar] [Publisher Link]

[12] W.L. Gahu, and D.J. Buehrer, "Vague Sets," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 2, pp. 610-614, 1993. [CrossRef] [Publisher Link]

[13] E. Smidt, and J. Kacprzzylc, "Distance between Fuzzy Sets and Systems," vol. 114, no. 3, pp. 505-518, 2000.

[14] L.A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965. [CrossRef] [Google Scholar]