University of Science and Technology of Southern Philippines

Alubijid | Cagayan de Oro | Claveria | Jasaan | Oroquieta | Panaon

COLLEGE OF
INFORMATION
TECHNOLOGY
AND COMPUTING

Name: Iverson M. Noja   Section: _____DS3A_____
Course code: **DS312**   Professor: **Jocelyn B. Barbosa, PhD**
Course Description: **Data Mining and Applications**

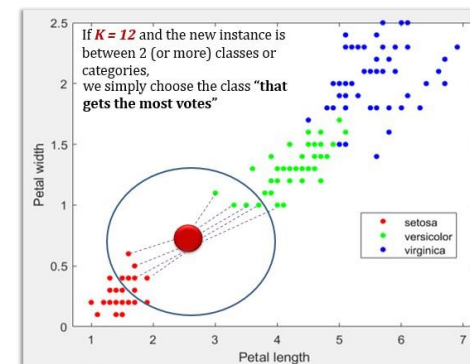## k-Nearest Neighbors (kNN) Algorithm

Objectives:

- To apply kNN algorithm to a fisheriris dataset.
- To identify similarity distance of the two (2) data points in a given dataset by applying distance function.
- To perform data analysis on fisheriris dataset.

Directions:

1. Open fisheriris dataset (see attached **fisheriris.xlsx**) or you may download the csv it here:
   https://gist.githubusercontent.com/curran/a08a1080b88344b0c8a7/raw/0e7a9b0a5d22642a06d3d5b9bcbad9890c8ee534/iris.csv

2. Determine the k-Nearest Neighbors by the finding the similarity distance (using Euclidian distance) of the two (2) data points (i.e. from the given new instance to each of the elements/data points in the whole dataset).
   Note: *Check attached file for the value of* **k** *and* **new instance** *assigned to you.*

   Sample final output:

Deadline: <Check your USTeP account>.
Filename: <surname_firstname_kNN>.pdf
        Example: **barbosa_jocelyn_kNN.pdf.**

University of Science and Technology of Southern Philippines

Alubijid | Cagayan de Oro | Claveria | Jasaan | Oroquieta | Panaon

COLLEGE OF
INFORMATION
TECHNOLOGY
AND COMPUTING

Solutions:

	Name	K	petal_length	petal_width	$(a_1 - b_1)^2$	$(a_2 - b_2)^2$	$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$
1	AGBALOG, PRINCESS HONEY G	8	5.5	1.7	2.25	0.04	1.5133
2	BAGUIO, CHRISTIAN JAY	9	5	2	1	0.01	1.005
3	BALISTOY, RAYMOND B	10	4	1.7	0	0.04	0.2
4	BANDIVAS, ROXANNE MAE D	11	2.5	1.3	2.25	0.36	1.6155
5	BARRIO, REYJANE MAE G.	12	4.5	1.6	0.25	0.09	0.5831
6	BASA, Ma. Sophia C	13	4.5	2.3	0.25	0.16	0.6403
7	BAYBAY, Alliana Eli U	14	4.5	2	0.25	0.01	0.5099
8	CAGAS, Nephea R	15	4.5	1.8	0.25	0.01	0.5099
9	CAÑETE , Vince Ryan Malagusan	16	5	1.4	1	0.25	1.118
10	CESPEDES, Karsci Vincze A	17	2.6	1.2	1.96	0.49	1.5652
11	DAUG, Nicole Mikka MACAPAYAG	18	5.2	1.8	1.44	0.01	1.2042
12	GALINATO, Jan Carlo H	8	3	1.8	1	0.01	1.005
13	GONZALES , Kyle Ray Cabanlas	9	2.6	1.2	1.96	0.49	1.5652
14	LAURON, Aubrey Faye Dadole	10	5	1.4	1	0.25	1.118
15	LLENARES, John Paolo M	11	4.5	1.8	0.25	0.01	0.5099
16	LOPEZ, Eurice Jean C	12	5.5	1.6	2.25	0.09	1.5297
17	MEGALBIO, DJ Alen P	13	3	1.5	1	0.16	1.077
18	NAVARRO, Charles Vince Cespon	14	2.5	1.3	2.25	0.36	1.6155
19	NOJA , Iverson Maga	15	4	1.9	0	0	0
20	ODCHIGUE, Lein Loisse Marundan	16	5	1	1	0.81	1.3454
21	OPALLA, ERICA B	17	5.5	1.7	2.25	0.04	1.5133
22	PADILLO, Rodmar Kent G	18	4	1.4	0	0.25	0.5
23	PALMES, Jhibsam B	8	5.5	2.1	2.25	0.04	1.5133
24	POCOL, Czarr Vic Israel P	9	2.5	1.1	2.25	0.64	1.7
25	RALUTO , KierDaniel U	10	3.6	2	0.16	0.01	0.4123
26	REYES, Lester Jane	11	4	1.5	0	0.16	0.4
27	ROA, arousfe franco d	12	5.5	1	2.25	0.81	1.7493
28	SABACAJAN, Cle Michael T	13	5.2	2	1.44	0.01	1.2042
29	SAQUILAYAN, Erykha Mae G	16	5	1	1	0.81	1.3454
30	TAMPARONG, Fatima O	8	3	1.8	1	0.01	1.005
31	TAPAO, Mark Vincent S	9	2.6	1.2	1.96	0.49	1.5652
32	TORRALBA, Colin Elijah J	10	5	1.4	1	0.25	1.118
33	VELONO, Hannah T	11	4.5	1.8	0.25	0.01	0.5099
34	VIUDOR, Jan Rey B	12	5.5	1.6	2.25	0.09	1.5297

Final output:

Nearest Neighbors (K=15)	
Distance	Owner
0.2	BALISTOY, RAYMOND B
0.4	REYES, Lester Jane
0.4123	RALUTO , KierDaniel U
0.5	PADILLO, Rodmar Kent G
0.5099	BAYBAY, Alliana Eli U
0.5099	CAGAS, Nephea R
0.5099	LLENARES, John Paolo M
0.5099	VELONO, Hannah T
0.5831	BARRIO, REYJANE MAE G.
0.6403	BASA, Ma. Sophia C
1.005	GALINATO, Jan Carlo H
1.005	BAGUIO, CHRISTIAN JAY
1.005	TAMPARONG, Fatima O
1.077	MEGALBIO, DJ Alen P
1.118	LAURON, Aubrey Faye Dadole