

Universidad de Guadalajara
Centro Universitario de Ciencias Exactas e Ingenierías

Ingeniería en Computación
Seminario de Solución de Problemas de Inteligencia Artificial II
Profesor: Campos Peña Diego

Lepiz Hernández Manuel Alejandro
Código: 214797998



Proyecto Final

Introducción

Se realizará en Python la implementación de los Algoritmos Regresion Logistica, K-Vecinos, SVM, Naive-Bayes, y una Red Neuronal usando un archivo .csv para la entrada de datos.

Se hará uso de un data set sobre un zoológico.

Y se evaluarán los resultados usando las métricas accuracy, precision, sensitivity, specificity y f1 score.

Desarrollo

Tabla de variables

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
animal_name	ID	Categorical				no
hair	Feature	Binary				no
feathers	Feature	Binary				no
eggs	Feature	Binary				no
milk	Feature	Binary				no
airborne	Feature	Binary				no
aquatic	Feature	Binary				no
predator	Feature	Binary				no
toothed	Feature	Binary				no
backbone	Feature	Binary				no

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
breathes	Feature	Binary				no
venomous	Feature	Binary				no
fins	Feature	Binary				no
legs	Feature	Categorical				no
tail	Feature	Binary				no
domestic	Feature	Binary				no
catsize	Feature	Binary				no
type	Target	Categorical				no

Tipos de datos

Additional Variable Information

1. animal name: Unique for each instance
2. hair: Boolean
3. feathers: Boolean
4. eggs: Boolean
5. milk: Boolean
6. airborne: Boolean
7. aquatic: Boolean
8. predator: Boolean
9. toothed: Boolean
10. backbone: Boolean
11. breathes: Boolean
12. venomous: Boolean
13. fins: Boolean
14. legs: Numeric (set of values: {0,2,4,5,6,8})
15. tail: Boolean
16. domestic: Boolean
17. catsize: Boolean
18. type: Numeric (integer values in range [1,7])

Nombre de los animales

Class Labels

Class# -- Set of animals:

=====

1 -- (41) aardvark, antelope, bear, boar, buffalo, calf, cavy, cheetah, deer, dolphin, elephant, fruitbat, giraffe, girl, goat, gorilla, hamster, hare, leopard, lion, lynx, mink, mole, mongoose, opossum, oryx, platypus, polecat, pony, porpoise, puma, pussycat, raccoon, reindeer, seal, sealion, squirrel, vampire, vole, wallaby, wolf

2 -- (20) chicken, crow, dove, duck, flamingo, gull, hawk, kiwi, lark, ostrich, parakeet, penguin, pheasant, rhea, skimmer, skua, sparrow, swan, vulture, wren

3 -- (5) pitviper, seasnake, slowworm, tortoise, tuatara

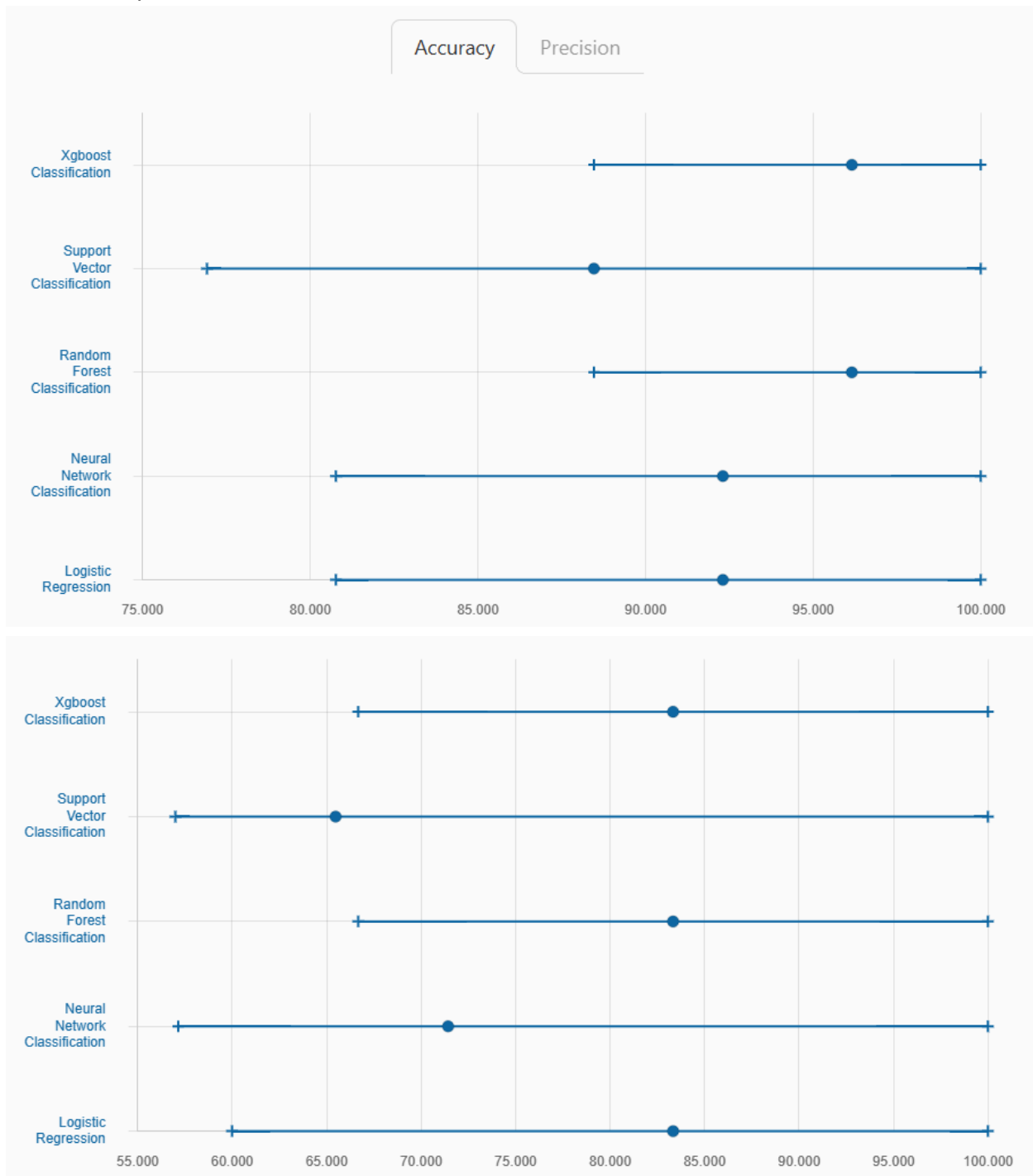
4 -- (13) bass, carp, catfish, chub, dogfish, haddock, herring, pike, piranha, seahorse, sole, stingray, tuna

5 -- (4) frog, frog, newt, toad

6 -- (8) flea, gnat, honeybee, housefly, ladybird, moth, termite, wasp

7 -- (10) clam, crab, crayfish, lobster, octopus, scorpion, seawasp, slug, starfish, worm

Resultados esperados



Se usó la implementación de un Clasificador, que fue usado en las prácticas anteriores.

Y se añadió el método de Red Neuronal.

15

Regresion Logistica

Precision: 0.95

	precision	recall	f1-score	support
1	1.00	1.00	1.00	12
2	1.00	1.00	1.00	2
3	0.00	0.00	0.00	1
4	0.67	1.00	0.80	2
6	1.00	1.00	1.00	3
7	1.00	1.00	1.00	1
accuracy			0.95	21
macro avg	0.78	0.83	0.80	21
weighted avg	0.92	0.95	0.93	21

K Vecinos

Precision: 0.95

	precision	recall	f1-score	support
1	1.00	1.00	1.00	12
2	1.00	1.00	1.00	2
3	0.00	0.00	0.00	1
4	0.67	1.00	0.80	2
6	1.00	1.00	1.00	3
7	1.00	1.00	1.00	1
accuracy			0.95	21
macro avg	0.78	0.83	0.80	21
weighted avg	0.92	0.95	0.93	21

s



SVM

Precision: 0.71



	precision	recall	f1-score	support
1	1.00	0.92	0.96	12
2	0.33	1.00	0.50	2
3	0.00	0.00	0.00	1
4	0.50	1.00	0.67	2
6	0.00	0.00	0.00	3
7	0.00	0.00	0.00	1
accuracy			0.71	21
macro avg	0.31	0.49	0.35	21
weighted avg	0.65	0.71	0.66	21

Naive Bayes					
Precision: 0.86					
	precision	recall	f1-score	support	
1	1.00	0.92	0.96	12	
2	1.00	1.00	1.00	2	
3	0.00	0.00	0.00	1	
4	0.40	1.00	0.57	2	
6	1.00	1.00	1.00	3	
7	0.00	0.00	0.00	1	
accuracy			0.86	21	
macro avg	0.57	0.65	0.59	21	
weighted avg	0.85	0.86	0.84	21	

Red Neuronal					
Precision: 0.95					
	precision	recall	f1-score	support	
1	1.00	1.00	1.00	12	
2	1.00	1.00	1.00	2	
3	0.00	0.00	0.00	1	
4	0.67	1.00	0.80	2	
6	1.00	1.00	1.00	3	
7	1.00	1.00	1.00	1	
accuracy			0.95	21	
macro avg	0.78	0.83	0.80	21	
weighted avg	0.92	0.95	0.93	21	

Conclusión

Los resultados fueron acercados a los esperados, de manera general.

Siendo que en comparativa son bastante acertados en cuanto a performance.

Los Algoritmos realizaron bien las clasificación, aunque en Regresión Logística, K Vecinos, Red Neuronal se puede notar que son los mejores para realizar la clasificación.

El peor por mucha diferencia fue SVM.

No se pudo crear una gráfica para la visualización de los datos ya que eran muchas etiquetas las existentes.