

EXTROVERTIDO VS INTROVERTIDO



Amaia Duo Resano, Nicole Estefanía Sarmiento, Erika Lairi Vera y Andrea Zamboran

ÍNDICE

1

INTRODUCCIÓN

2

INserción de
DATOS

3

NORMALIZACIÓN
DE DATOS

4

COMPARACIÓN
DE LOS DATOS

5

K-1

INserción
NUEVO EJEMPLO

6

7

K-NN

8

WILSON

9

ESTUDIO
PRECISIÓN

INTRODUCCIÓN

Nuestro Dataset muestra el comportamiento de una persona (Extrovertido vs Introvertido) teniendo en cuenta su situación y la de su entorno.

El objetivo del proyecto es predecir correctamente la personalidad a partir de esas variables mediante K=1, K-NN y Wilson.

Elegimos este tema porque resulta actual, cercano y relevante para comprender mejor el comportamiento humano.

INSERCIÓN DE DATOS

Procesar cada línea del fichero mediante la librería fragmenta

Metemos los datos en un struct para poder usarlos fácilmente

Metemos los datos en una lista para poder acceder a los datos y los eliminamos si es necesario (útil para Wilson)



NORMALIZACIÓN DE DATOS

Después de introducir los datos al struct los normalizamos con la fórmula, exceptuando aquellos que no son números:

$$V = (v - \text{min_atributo}) / (\text{max_atributo} - \text{min_atributo})$$

Una vez metidos los datos al struct y normalizados los insertamos en una lista.

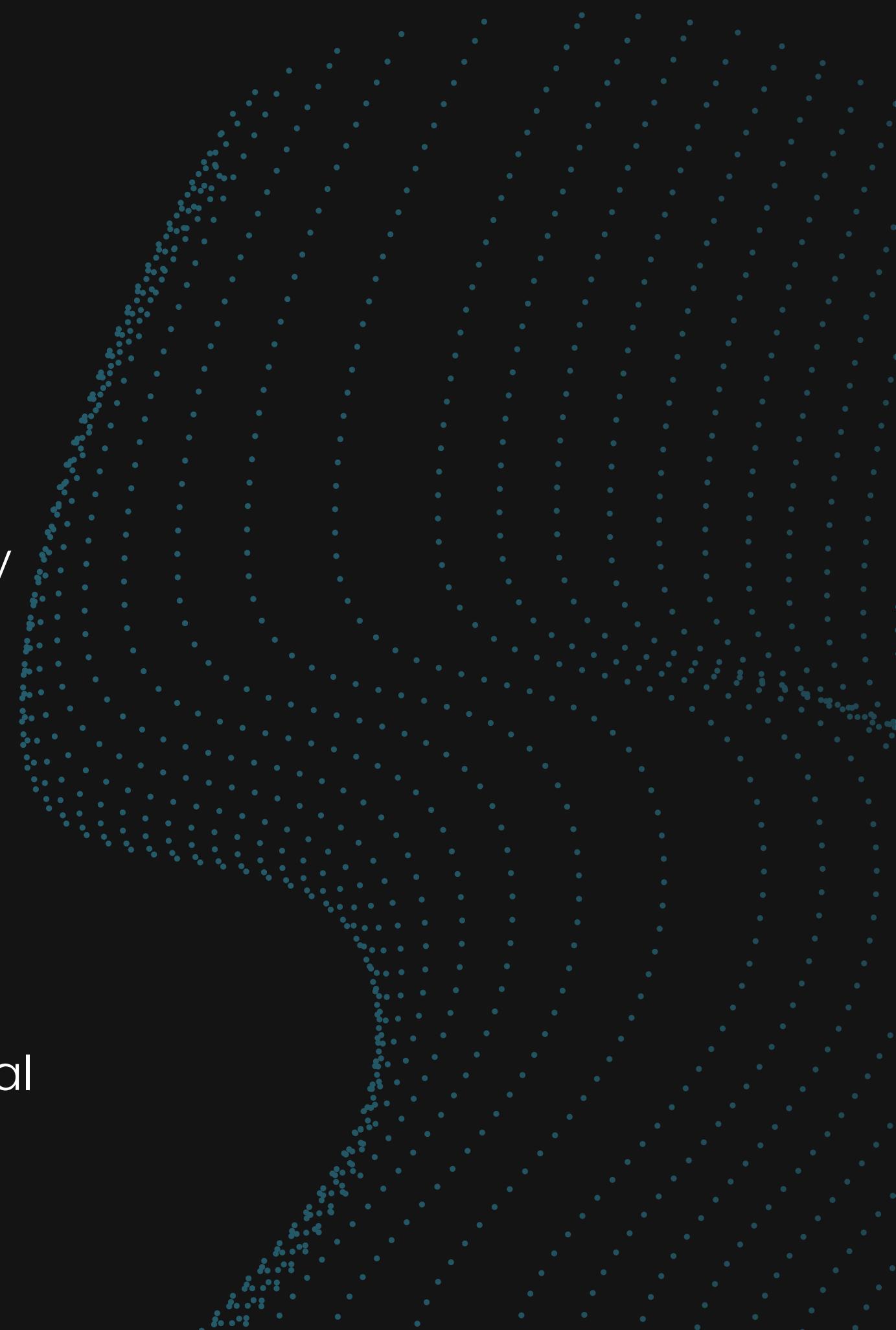
- Solo normalizamos los atributos numéricos
- Los categóricos (sí/no) se mantienen sin cambios.

COMPARACIÓN DE LOS DATOS

Tras tener los datos normalizados y insertados en el array podemos compararlos de 2 en 2.

- Si son números -> sumamos el cuadrado de la resta al total
- Si son caracteres -> sumamos 0 en caso de que sean iguales y 1 si son distintos al total

Terminamos haciendo la raíz cuadrada sobre la suma total



K-1

Coge un elemento de la lista y lo compara con el resto de la lista eligiendo así el que menor distancia tenga con el elemento elegido.

K-1

2559 ACIERTOS
341 FALLOS

PORCENTAJE ACIERTOS
88,24 %

¿Quéquieres hacer?

- 0 - Salir
- 1 - K = 1 con todos
- 2 - KNN con K elegido
- 3 - Wilson con la mejor K
- 4 - Wilson con K cualquiera
- 5 - Estudio precisión vs K y nº de ejemplos

Opción: 1

Clasificación de K=1 con todos los datos

Total datos: 2900, Aciertos: 2559, Fallos: 341
88.24% de acierto

INSERCIÓN NUEVO EJEMPO

```
--- Introducción de un ejemplo nuevo ---
Tiempo a solas (0-11): 9.0
Miedo escénico (Yes/No): Yes
Asistencia a eventos (0-10): 1.0
Salir fuera (0-7): 1.0
Agotado tras socializar (Yes/No): Yes
Número de amigos cercanos (0-15): 2.0
Frecuencia publicación redes (0-10): 1.0
Elige K para clasificar el ejemplo nuevo (entero >=1): 5

--- Clasificación de ejemplo nuevo (K=5) ---
Clase predicha: introvert
-----
```

K-NN

Hace lo mismo que K-N solo que se pide al usuario que inserte un número, siendo este el número de distancias mínimas que se almacenarán.

Creamos un struct que almacena la distancia y el valor de cada elemento para después poder insertarlo al maxMonticulo.

En caso de que haya empate , cogemos el que tenga la menor distancia

WILSON

- Limpiar y depurar el Conjunto de Entrenamiento
- Bien clasificado → Se mantiene (nueva Lista)
 - Mal clasificado → Se descarta
- Resultado: Dataset mas pequeño y más coherente

Wilson con la Mejor K

¿Quéquieres hacer?

- 0 - Salir
- 1 - K = 1 con todos
- 2 - KNN con K elegido
- 3 - Wilson con la mejor K
- 4 - Wilson con K cualquiera
- 5 - Estudio precisión vs K y nº de ejemplos

Opción: 3

Clasificación Wilson con la mejor K

Porcentaje: 93.41% Mejor K vecinos: 16

Wilson Con K Cualquiera

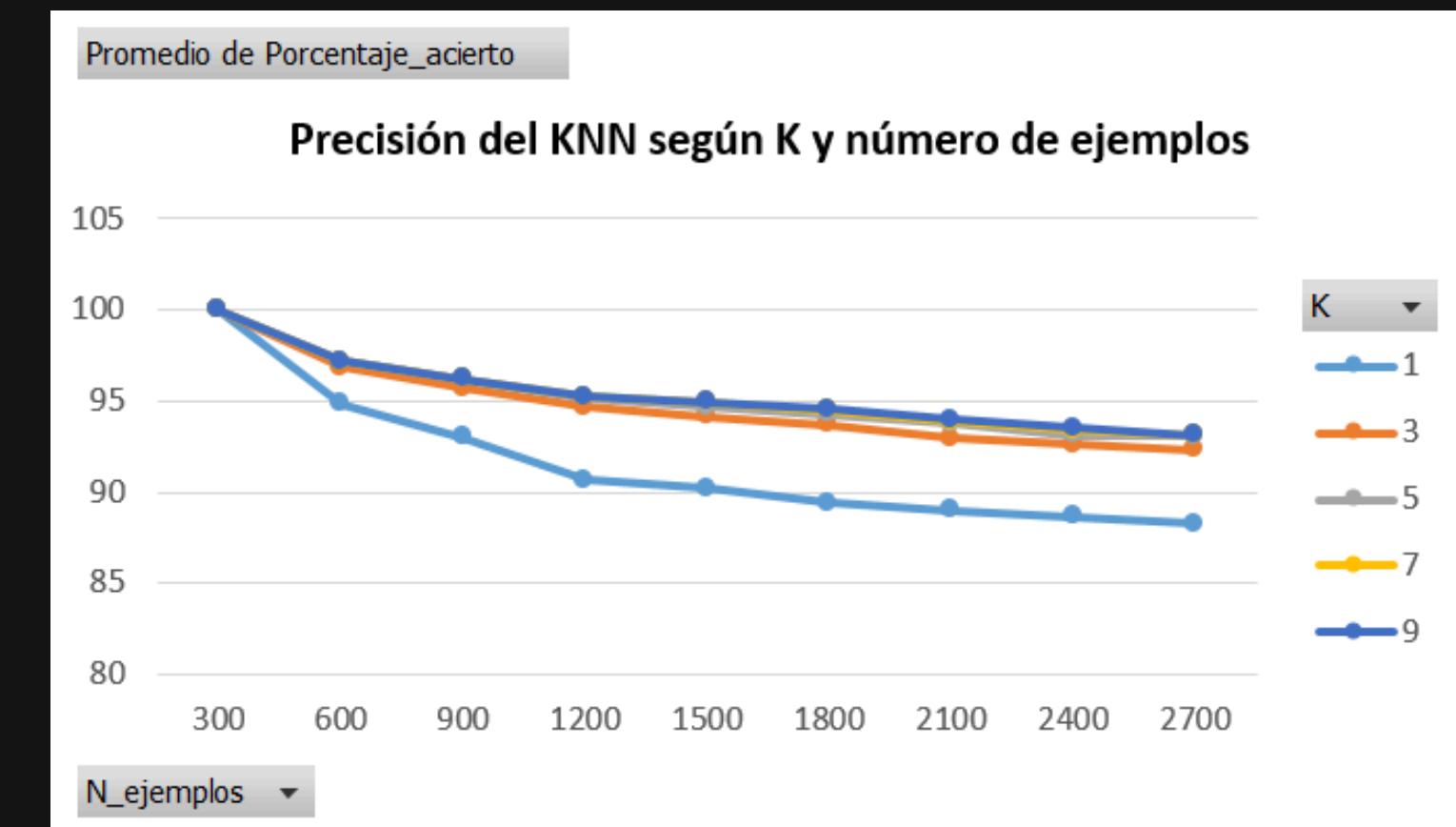
```
¿Quéquieres hacer?  
0 - Salir  
1 - K = 1 con todos  
2 - KNN con K elegido  
3 - Wilson con la mejor K  
4 - Wilson con K cualquiera  
5 - Estudio precisión vs K y nº de ejemplos  
Opción: 4  
¿Con cuántos vecinosquieres comparar (K)?: 3
```

```
0.82 yes 0.20 0.14 yes 0.33 0.00 introvert  
0.00 no 0.60 0.43 no 0.47 0.50 extrovert  
0.36 yes 0.10 0.29 yes 0.87 0.20 introvert  
0.09 no 0.90 0.43 no 0.47 0.60 extrovert  
0.64 yes 0.20 0.29 yes 0.87 0.10 introvert  
1.00 yes 0.00 0.29 yes 0.20 0.10 introvert  
0.91 yes 0.00 0.00 no 0.13 0.00 introvert  
0.64 yes 0.00 0.00 yes 0.33 0.20 introvert  
0.00 no 0.90 0.86 no 0.40 0.30 extrovert  
0.45 yes 0.00 0.14 yes 0.87 0.00 introvert  
0.55 yes 0.00 0.00 yes 0.20 0.00 introvert  
0.73 yes 0.20 0.00 yes 0.87 0.20 introvert  
0.55 yes 0.30 0.14 yes 0.33 0.10 introvert  
0.82 yes 0.20 0.00 yes 0.87 0.20 introvert  
0.82 no 0.20 0.00 yes 0.27 0.20 introvert  
0.00 no 0.90 0.43 no 0.80 0.36 extrovert  
0.27 no 0.70 0.86 no 0.40 0.60 extrovert  
0.27 no 0.80 0.43 no 0.93 0.90 extrovert  
0.36 yes 0.10 0.14 yes 0.27 0.00 introvert  
1.00 yes 0.10 0.43 yes 0.13 0.00 introvert  
0.27 no 0.60 0.86 no 0.40 0.90 extrovert  
Número de filas: 2669
```

ESTUDIO DE PRECISIÓN

```
5 - Estudio precisión vs K y nº de ejemplos
Opción: 5
La lista tiene 2669 ejemplos.
¿Cada cuántos ejemplos quieres muestrear? (p.ej. 100): 300

N_ejemplos, K, Porcentaje_acierto
300, 1, 100.00
300, 3, 100.00
300, 5, 100.00
300, 7, 100.00
300, 9, 100.00
600, 1, 99.83
600, 3, 99.83
600, 5, 99.83
600, 7, 99.83
600, 9, 99.83
900, 1, 99.78
900, 3, 99.78
900, 5, 99.78
900, 7, 99.78
900, 9, 99.78
1200, 1, 99.75
1200, 3, 99.75
1200, 5, 99.75
1200, 7, 99.75
1200, 9, 99.75
1500, 1, 99.53
1500, 3, 99.73
1500, 5, 99.73
1500, 7, 99.73
1500, 9, 99.73
1800, 1, 99.56
1800, 3, 99.72
1800, 5, 99.72
1800, 7, 99.72
1800, 9, 99.72
2100, 1, 99.67
2100, 3, 99.76
```



FIN

GRACIAS POR SU ATENCIÓN

