

Informe – Criterios de Aceptación

Introducción

Durante el desarrollo del programa se implementó la funcionalidad para cambiar dinámicamente los pesos asignados a distintas columnas (**first_name**, **last_name** y **email**) y recalcular los resultados de similitud entre registros. Esto permite evaluar cómo influye cada campo en la detección de coincidencias y qué configuración es más adecuada según el caso de uso.

Se realizaron varias pruebas modificando los pesos y observando los resultados en diferentes combinaciones.

Gráficas y capturas

SIMBOLOGIA
First_name (FN)
Last_name (LN)
email (EM)

SIMBOLOGIA DE COLOR
Pasa todos los filtros
Pasa solo el filtro de 70%
No pasa ningún filtro

	Evelyn	Robert	Maryna	Brenda	Yvonne
FN=2, LN=2, EM=8	100	89.83	78.99	65.63	59.92
FN=1, LN=3, EM=6	100	89.7	79.61	66.41	60.41
FN=6, LN=1, EM=3	100	91.62	73.14	58.71	56.71
FN=9, LN=6, EM=1	100	87.81	68.01	48.27	40.77
FN=2, LN=4, EM=10	100	90.36	80.15	67.99	62.99
FN=1, LN=1, EM=9	100	93.96	92.58	92.78	92.58
FN=1, LN=1, EM=1	100	88.81	73.86	57.45	50.78

Tabla 1. Resultados de ponderación con diferentes pesos

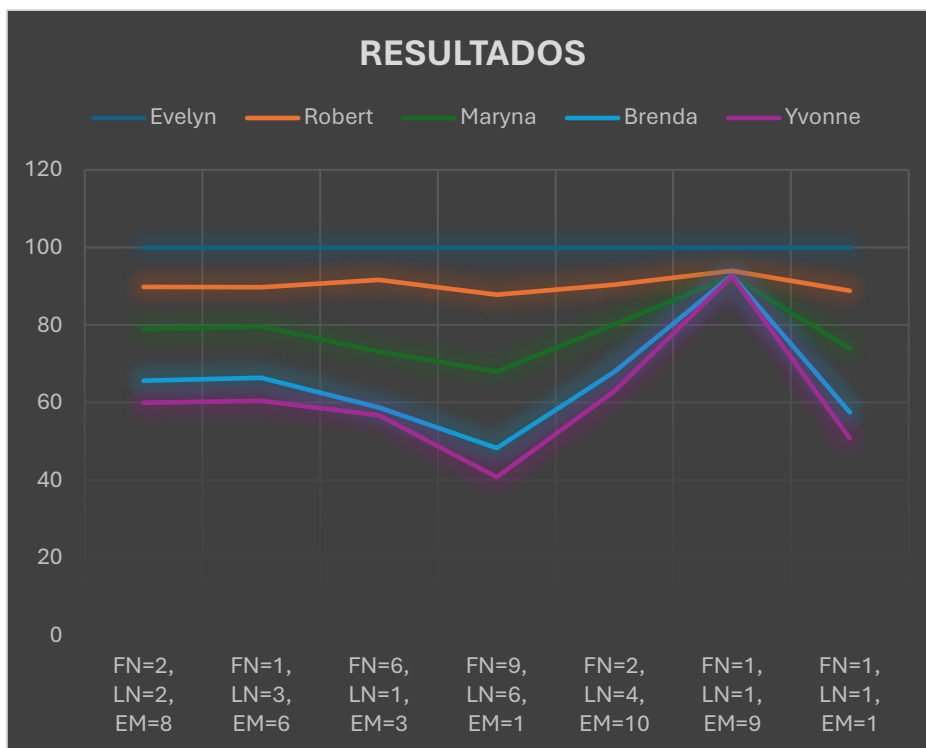


Figura 1. Gráfica de líneas de distribución de scores ponderados por usuario

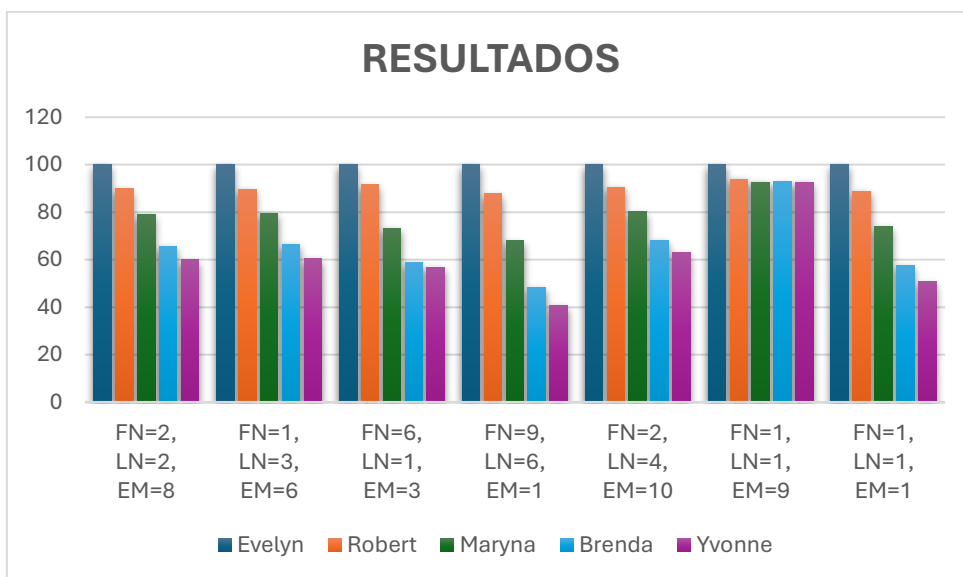


Figura 2. Gráfica de barras de distribución de scores ponderados por usuario

```

Se usan los pesos de la base de datos (más recientes).
gg {'first_name': 4, 'last_name': 4, 'email': 2, 'fecha_modificacion': datetime.datetime(2025, 9, 23, 9, 59, 49)}
Nuevo peso para first_name (1-99): 1
Nuevo peso para last_name (1-99): 1
Nuevo peso para email (1-99): 1

Simulación de impacto con nuevos pesos:

```

	query	match	weighted_score
0	{'first_name': 'Evelyn', 'last_name': 'fryda',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	100.00
1	{'first_name': 'Robert', 'last_name': 'Davis',...	{'first_name': 'Roberto', 'last_name': 'Daviz'...	88.81
2	{'first_name': 'Maryna', 'last_name': 'Wells',...	{'first_name': 'Matina', 'last_name': 'Wils', ...	73.86
3	{'first_name': 'Brenda', 'last_name': 'Meyer',...	{'first_name': 'Brinte', 'last_name': 'naier',...	57.45
4	{'first_name': 'Yvonne', 'last_name': 'Jesen',...	{'first_name': 'Yviclo', 'last_name': 'yazin',...	50.78

```

¿Desea cambiar los pesos? (y/n): █

```

Figura 3. Captura 1 Ejemplo de recalcu de pesos en las columnas

```

Nuevo peso para first_name (1-99): 1
Nuevo peso para last_name (1-99): 3
Nuevo peso para email (1-99): 6

Simulación de impacto con nuevos pesos:

```

	query	match	weighted_score
0	{'first_name': 'Evelyn', 'last_name': 'fryda',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	100.00
1	{'first_name': 'Robert', 'last_name': 'Davis',...	{'first_name': 'Roberto', 'last_name': 'Daviz'...	89.70
2	{'first_name': 'Maryna', 'last_name': 'Wells',...	{'first_name': 'Matina', 'last_name': 'Wils', ...	79.61
3	{'first_name': 'Brenda', 'last_name': 'Meyer',...	{'first_name': 'Brinte', 'last_name': 'naier',...	66.41
4	{'first_name': 'Yvonne', 'last_name': 'Jesen',...	{'first_name': 'Yviclo', 'last_name': 'yazin',...	60.41

```

¿Desea cambiar los pesos? (y/n): y
Nuevo peso para first_name (1-99): 6
Nuevo peso para last_name (1-99): 1
Nuevo peso para email (1-99): 3

Simulación de impacto con nuevos pesos:

```

	query	match	weighted_score
0	{'first_name': 'Evelyn', 'last_name': 'fryda',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	100.00
1	{'first_name': 'Robert', 'last_name': 'Davis',...	{'first_name': 'Roberto', 'last_name': 'Daviz'...	91.62
2	{'first_name': 'Maryna', 'last_name': 'Wells',...	{'first_name': 'Matina', 'last_name': 'Wils', ...	73.14
3	{'first_name': 'Brenda', 'last_name': 'Meyer',...	{'first_name': 'Brinte', 'last_name': 'naier',...	58.71
4	{'first_name': 'Yvonne', 'last_name': 'Jesen',...	{'first_name': 'Yviclo', 'last_name': 'yazin',...	56.71

Figura 4. Captura 2 Ejemplo de recalcu de pesos en las columnas

```

¿Desea cambiar los pesos? (y/n): y
Nuevo peso para first_name (1-99): 2
Nuevo peso para last_name (1-99): 4
Nuevo peso para email (1-99): 10

Simulación de impacto con nuevos pesos:

```

	query	match	weighted_score
0	{'first_name': 'Evelyn', 'last_name': 'fryda',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	100.00
1	{'first_name': 'Robert', 'last_name': 'Davis',...	{'first_name': 'Roberto', 'last_name': 'Daviz'...	90.36
2	{'first_name': 'Maryna', 'last_name': 'Wells',...	{'first_name': 'Matina', 'last_name': 'Wils', ...	80.15
3	{'first_name': 'Brenda', 'last_name': 'Meyer',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	67.99
4	{'first_name': 'Yvonne', 'last_name': 'Jesen',...	{'first_name': 'Evelyn', 'last_name': 'fryda',...	62.99

Figura 5. Captura 3 Ejemplo de recalcu de pesos más justo en las columnas

CONCLUSIÓN

Configuración más justa

Tras las pruebas realizadas con diferentes configuraciones de pesos, la conclusión es que la configuración más justa es la siguiente:

- **first_name:** peso bajo (1 o 2).
- **last_name:** un poco mayor que first_name (entre 2 y 5).
- **email:** debe ser el campo con mayor peso. Como mínimo debe equivaler a la suma de first_name + last_name, y para ser más justo aún, debe ser esa suma más la mitad entera más cercana de la suma.

Esta configuración refleja que el correo electrónico es el campo más representativo y menos repetitivo, mientras que nombre y apellido son campos más genéricos y repetibles.

Posibles Problemas Detectados

- Aunque existen validaciones para el filtrado, todavía es posible ingresar filtros incorrectos, lo que puede ocasionar resultados vacíos o inconsistentes.
- La ponderación de pesos nunca será 100% exacta: en algunos escenarios puede marcar como coincidencia más probable a un registro equivocado, sobre todo si el usuario anteriormente cambió datos clave como correo o apellido solo en uno de los 2 registros.
- Requiere cierto nivel de capacidad de interpretación y pensamiento crítico por parte del administrador, para encontrar una ponderación justa y equitativa.

Recomendaciones de Mejora

- **Agregar más columnas únicas:** incluir datos como teléfono, CURP, RFC o identificadores más difíciles de repetir.
- **Usar solo los nombres y apellidos como complementos** y no como guía principal, para reducir falsos positivos.
- **Dar flexibilidad al administrador** para elegir qué columnas incluir o excluir en el cálculo.
- **Mejorar las validaciones en filtros y entradas de peso**, para evitar errores de captura.