Person-Target Profiling #1: Binary Summed Percent

A Cognadev Profiling Series publication

© 2022 Cognadev Ltd.
Paul Barrett, Cognadev Ltd.

1. Definition 5

1.1 Properties 7

2. Finessing a Target Profile 9

2.1 Weightings 10

2.2 Conditionals 11

3. Examples 13

3.1 Example_1: Small Candidate N 14

3.1.1 Unweighted matching 14

3.1.2 Weighted matching 17

3.1.3 Conditional matching 19

3.2 Example_2: Large Candidate N 21

3.2.1 Unweighted matching 21

3.2.2 Weighted matching 22

3.2.3 Conditional matching 23

4. Benefits and Drawbacks 25

5. My Conclusions 31

Index 33

 © 2022 Cognadev Ltd.

1. Definition

1. Definition

Summed Binary Percent Person-Target-Profile Matching is no more than summing the number of attributes in a person's vector of attributes which match *a class or ordered category* target profile consisting of multiple attributes, then dividing through by the total number of attributes in that target profile, and expressing the result as a percentage.

That is, a match is simply counted as 1 (a match) where a person's attribute category, or attribute score/rank falls within the bounds set in the target profile. Usually a match is no more a than 'present or absent' indicator, as in matching target soft-skills against those extracted using natural language processing from an individual's CV, or text-fragments of their social media activity.

For example, given we have 10 attributes assessed in total, with 5 selected as a Target Profile, and 3 which match target attributes in a Person's profile:

Attributes	Target Profile	Person Profile	Match
A1	[5 and above]	[5 and above]	☑
A2		[3]	
A3	[3 and above]	[2]	
A4		[5]	
A5	[Above Average]	[Above Average]	☑
A6		[4]	
A7	[Achiever]	[Challenger]	
A8	[Team Player]	[Team Player]	☑
A9		[4]	
A10		[Bold]	

The match % is calculated as:

$$N_matches := 3$$

$$N_attributes_in_target := 5$$

$$match := \left(\frac{N_matches}{N_attributes_in_target} \right) \cdot 100 = 60\%$$

© 2022 Cognadev Ltd.

1.1 Properties

Being a *unifinal* profile match method, it assumes there is only one optimal profile for a 'target'. i.e. there is only one profile which is deemed causal or associated with 'success' or which characterises the best employees in a job-role/group.

It is also most often a form of *non-compensatory* profiling; where each attribute's qualities or ranks are considered as independent of any other attribute i.e. the presence or absence (match or no match) of one attribute cannot 'compensate' for presence or absence of another. However, it is possible to adjust the procedure to build in some compensatory component, as explained in the section: 2. Finessing a Target Profile 10.

 © 2022 Cognadev Ltd.

2. Finessing a Target Profile

2. Finessing a Target Profile

Where no conditionals or weightings are applied to any attribute within a target profile, and without any diagnostics as to which targets were matched, the reporting of a simple % match is hardly useful in practice (except where its value is very high - although any such notional 'bound' value is conditional upon the number of targets in a target profile. For example, if there are only 4 attributes to be matched, a person can only attain match values of [0, 25, 50, 75, and 100]% . If 10 attributes, then [0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]%.

But it is possible to 'finesse' a simple binary sum - using attribute weightings and conditional 'rules' which allow for a more 'sensitive', more variability, and a more informative index.

This is also known as a *hybrid target profile*, where person-target matching is not just an arithmetic or simple 'present-absent' count process but also contains '*production-rules' (If-then statements*) and rules adding 'compensatory[11]' features which are designed to moderate the match %.

2.1 Weightings

This is where instead of a simple "present-absent" approach is undertaken, a weight is applied to the binary [0, 1] choice - which moderates the effect it has in any summed % match. So, if you think an attribute is more important to be matched than another, you can weight the 'occurrence' to make it so in any resulting match %.

However, things get really complicated very quickly because those weights affect the maximum-possible match values which can be attained, so weighted binary profiling needs some careful arithmetic adjustment-formulae to keep things within a 0-100% match range. Examples 3.1.2[17] and 3.2.2[22] demonstrate this kind of 'adjusted' binary sum profiling.

The calculation/formula for weighting is applied to case #1 in the small-sample dataset[14]:

© 2022 Cognadev Ltd.

The relevant data

- the target attributes [Attribute], the target profile [*Target*],
- a person's scores on those target attributes [*Person*],
- a binary indicator showing whether a person's scores match the profile target for each attribute [*Match*],
- the weight applied to each target match [*Weight*].

$$Attribute = \begin{bmatrix} \text{“Ambition”} \\ \text{“Tolerance”} \\ \text{“Curiosity”} \\ \text{“Sincerity”} \\ \text{“Rigidity”} \\ \text{“Extraversion”} \\ \text{“Openness”} \end{bmatrix} \quad Target = \begin{bmatrix} \text{“>= 4”} \\ \text{“> = 4”} \\ \text{“<= 5”} \\ \text{“between 3 \& 6”} \\ \text{“between 4 \& 6”} \\ \text{“<=4”} \\ \text{“>=4”} \end{bmatrix} \quad Person = \begin{bmatrix} 5 \\ 1 \\ 3 \\ 1 \\ 2 \\ 5 \\ 4 \end{bmatrix} \quad Match = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad Weight = \begin{bmatrix} 2 \\ 1.5 \\ 3 \\ 1 \\ 1 \\ 1.5 \\ 0.5 \end{bmatrix}$$

The match % must be between 0 to 100%, which requires taking into account the sum of the weights applied to the binary matches [0, 1]. In this example, the person only matches 3 of the profile targets.

The weighted sum formula

$$Sum_weights := \sum Weight = 10.5 \qquad N_attributes := 7$$

$$Weighted_Sum := \left| \begin{array}{l} \text{for } i \in 1..N_attributes \\ \quad sum \leftarrow sum + Match_i \cdot Weight_i \\ sum \end{array} \right| = 5.5$$

$$Match_percent := \left(\frac{Weighted_Sum}{Sum_weights} \right) = 52.381\%$$

If the profile matches had not been weighted, the simple binary sum match would have been 3/7=**42.9%**

Ultimately, I wouldn't use this approach unless I was really convinced I needed to weight matches proportionately. Given this Binary sum profiling is the most basic approach to person-target profiling, I'd prefer to stay within the simple binary logic and use Conditionals[11]. That is, apply rules as to what's "must have" vs "nice to have".

2.2 Conditionals

These are rules supplied by a user and embedded in the matching algorithm in order to enable a match coefficient to reflect whether or not some attributes 'not-matched' are terminal to the entire process, or somehow reduce or attenuate a match coefficient to reflect 'nice to have' matches rather than 'must-haves'.

These rules can be qualitative or quantitative. That is, a simple "if-then" statement can terminate an entire matching process, if the attribute not-matched is specified as 'must-have'.

© 2022 Cognadev Ltd.

And, as in examples 3.1.3 19 and 3.2.3 23, the rules can include compensatory statements such as:

if designated "must have" attribute X is not matched, but "'compensatory" attribute Y is, then do not terminate the match algorithm.

The problem of course is that things can get very complicated and confusing very quickly!

And, as with any adjustments beyond merely summing binary [0, 1] values, the potential adverse consequences of the rules *(no candidates make it through the algorithm; match values are always pitiful etc.)* need to be carefully sense-checked.

So, as with Weightings 10, finessing any binary match like this must be done sparingly and with some care/due-diligence modeling wherever possible - otherwise just keep things simple!

© 2022 Cognadev Ltd.

3. Examples

3. Examples

Three examples have been created to show how this profiling method behaves with small and large candidate database sizes.

Two artificial random-data datasets are used in Examples 1 and 2, with integer scores on each attribute varying between 1 and 7:

- A 15 variable by 20 person dataset, with all 15 variables comprising data sampled from a normal distribution, but not actually distributed normally given the very small sample size (just 20 cases)
- A 15 variable by 10,000 person dataset, the first 10 variables with scores distributed normally, the final 5 distributed as a right-skewed beta distribution

The point of these datasets is to highlight specific features of the profile methodology.

3.1 Example_1: Small Candidate N

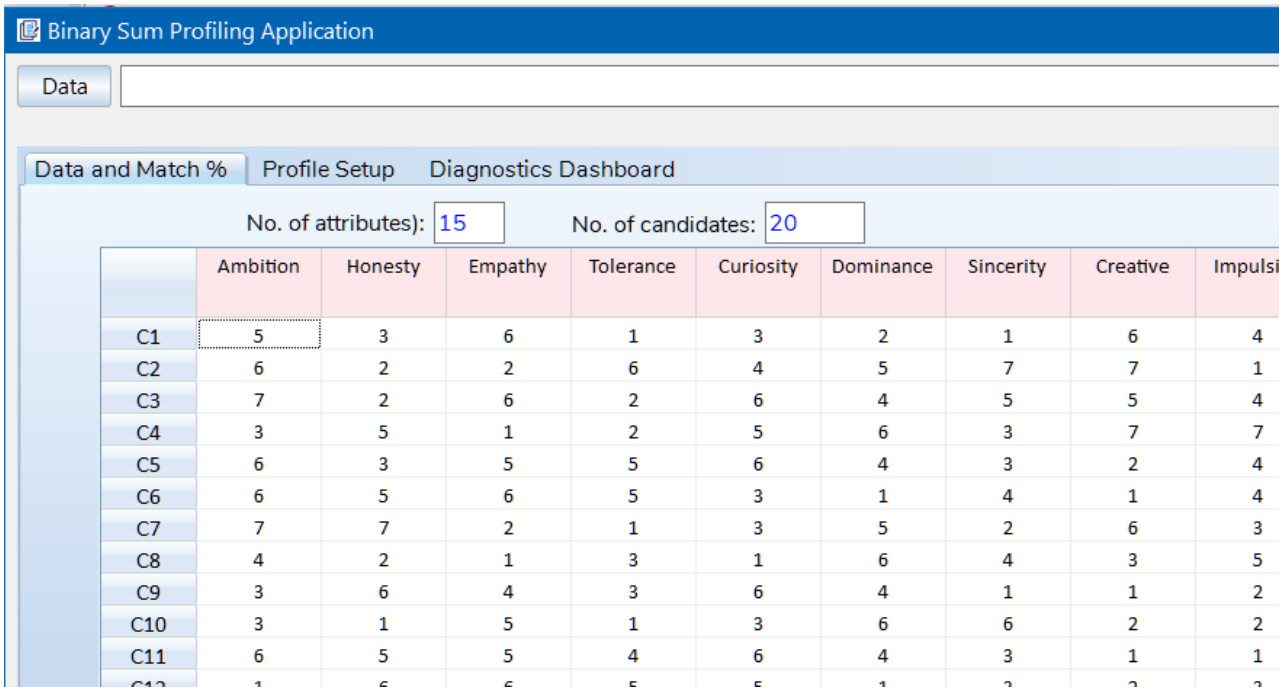
In this section, I use the 15 variable by 20 person dataset, with all 15 variables comprising data sampled from a normal distribution, values ranging between 1 and 7, but given the very small sample size (just 20 cases), 'normality' is just a token gesture!

3.1.1 Unweighted matching

This is the most straightforward use of binary sum profiling, where a "1" is recorded if a candidate score (or category) matches the target profile setting. A "0" is otherwise recorded.

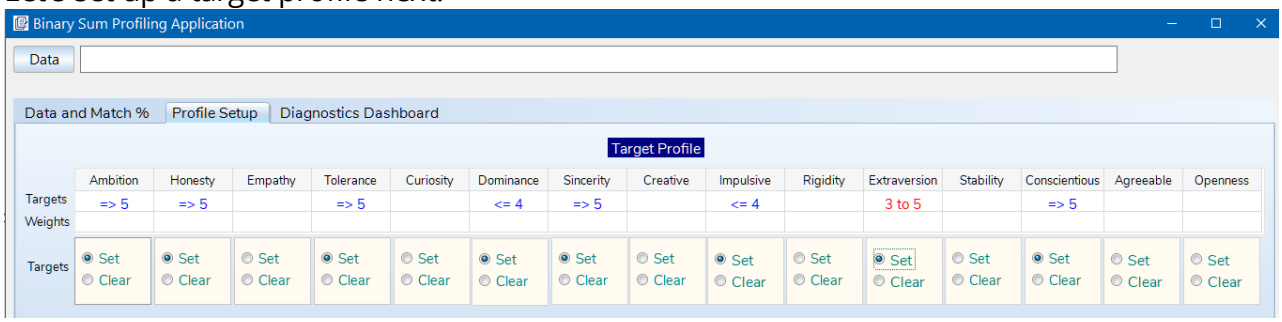
The data look like:

© 2022 Cognadev Ltd.



	Ambition	Honesty	Empathy	Tolerance	Curiosity	Dominance	Sincerity	Creative	Impulsi
C1	5	3	6	1	3	2	1	6	4
C2	6	2	2	6	4	5	7	7	1
C3	7	2	6	2	6	4	5	5	4
C4	3	5	1	2	5	6	3	7	7
C5	6	3	5	5	6	4	3	2	4
C6	6	5	6	5	3	1	4	1	4
C7	7	7	2	1	3	5	2	6	3
C8	4	2	1	3	1	6	4	3	5
C9	3	6	4	3	6	4	1	1	2
C10	3	1	5	1	3	6	6	2	2
C11	6	5	5	4	6	4	3	1	1

Let's set up a target profile next:



	Ambition	Honesty	Empathy	Tolerance	Curiosity	Dominance	Sincerity	Creative	Impulsive	Rigidity	Extraversion	Stability	Conscientious	Agreeable	Openness
Targets	=> 5	=> 5		=> 5		<= 4	=> 5		<= 4		3 to 5		=> 5		
Weights															
Targets	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear

We have 8 attributes in our selection profile e.g. if a person's Ambition score is equal to or greater than 5, then that counts as a match.

Matching our 20 candidates to this profile reveals:

© 2022 Cognadev Ltd.

	Creative	Impulsive	Rigidity	Extraversion	Stability	Conscientious	Agreeable	Openness	Unweighted Match %
C1	6	4	2	5	4	7	7	4	62.5
C2	7	1	1	3	5	3	4	2	62.5
C3	5	4	3	3	7	1	1	3	62.5
C4	7	7	1	4	6	2	4	6	25.0
C5	2	4	7	1	7	1	2	3	50.0
C6	1	4	3	2	2	7	7	3	75.0
C7	6	3	5	1	4	4	6	2	37.5
C8	3	5	5	2	6	7	7	1	12.5
C9	1	2	5	2	5	7	7	1	50.0
C10	2	2	7	5	7	4	4	7	37.5
C11	1	1	2	2	7	3	7	1	50.0
C12	2	3	4	2	7	7	4	4	62.5
C13	3	4	4	7	5	3	7	2	50.0
C14	3	1	2	7	4	6	4	3	50.0
C15	2	6	3	2	4	5	5	3	37.5
C16	3	5	4	6	6	7	1	7	50.0
C17	6	1	7	5	7	5	3	1	62.5
C18	5	2	2	7	3	6	6	6	50.0
C19	2	3	6	7	6	4	7	4	50.0
C20	6	6	5	7	3	1	1	3	0.0

with the histogram of matches:

Binary Sum profiling: Unweighted match frequencies				
Match %	Count	% Count	Cumulative Count	Cumulative Count %
75	1	5.0	1	5.0
62.5	5	25.0	6	30.0
50	8	40.0	14	70.0
37.5	3	15.0	17	85.0
25	1	5.0	18	90.0
12.5	1	5.0	19	95.0
0	1	5.0	20	100.0

So, just one person has a match index of 75% (case C6). If we were using a 70% minimum target match to indicate 'persons of interest', then only Case #6 would be identified as 'meeting/exceeding' the target profile.

The calculation for this individual is:

© 2022 Cognadev Ltd.

The relevant data

- the target attributes [Attribute], the target profile [*Target*],
- a person's scores on those target attributes [*Person*],
- a binary indicator showing whether a person's scores match the profile target for each attribute [*Match*],
- the weight applied to each target match [*Weight*].

$$Attribute = \begin{bmatrix} \text{“Ambition”} \\ \text{“Honesty”} \\ \text{“Tolerance”} \\ \text{“Dominance”} \\ \text{“Sincerity”} \\ \text{“Impulsive”} \\ \text{“Extraversion”} \\ \text{“Conscientious”} \end{bmatrix} \quad Target = \begin{bmatrix} \text{“>= 5”} \\ \text{“>= 5”} \\ \text{“>= 5”} \\ \text{“<= 4 ”} \\ \text{“>= 5”} \\ \text{“<= 4”} \\ \text{“between 3 \& 5”} \\ \text{“>= 5”} \end{bmatrix} \quad Person = \begin{bmatrix} 6 \\ 5 \\ 5 \\ 1 \\ 4 \\ 4 \\ 2 \\ 7 \end{bmatrix} \quad Match = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad Weight = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

The match % must be between 0 to 100%, which requires taking into account the sum of the weights applied to the binary matches [0, 1]. In this example, the person matches 6 of the 8 profile targets.

The weighted sum formula

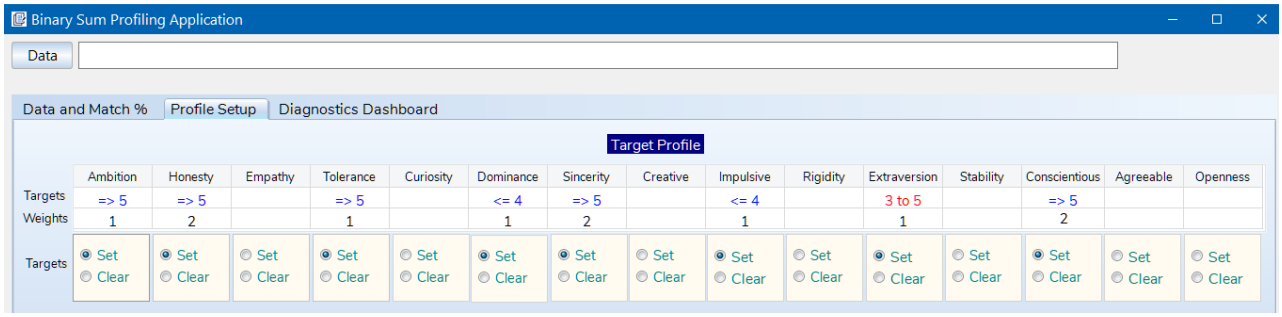
$$Sum_weights := \sum Weight = 8 \qquad N_attributes := 8$$

$$Weighted_Sum := \left\| \begin{array}{l} \text{for } i \in 1..N_attributes \\ \quad sum \leftarrow sum + Match_i \cdot Weight_i \\ sum \end{array} \right| = 6$$

$$Match_percent := \left(\frac{Weighted_Sum}{Sum_weights} \right) = 75\%$$

3.1.2 Weighted matching

Using the same small-sample dataset as in Example 3.1.1 14, and the same target profile - but now with weights defined for each target attribute:



the results are:

© 2022 Cognadev Ltd.

	Creative	Impulsive	Rigidity	Extraversion	Stability	Conscientious	Agreeable	Openness	Unweighted Match %	Weighted Match %
C1	6	4	2	5	4	7	7	4	62.5	54.5
C2	7	1	1	3	5	3	4	2	62.5	54.5
C3	5	4	3	3	7	1	1	3	62.5	54.5
C4	7	7	1	4	6	2	4	6	25.0	27.3
C5	2	4	7	1	7	1	2	3	50.0	36.4
C6	1	4	3	2	2	7	7	3	75.0	72.7
C7	6	3	5	1	4	4	6	2	37.5	36.4
C8	3	5	5	2	6	7	7	1	12.5	18.2
C9	1	2	5	2	5	7	7	1	50.0	54.5
C10	2	2	7	5	7	4	4	7	37.5	36.4
C11	1	1	2	2	7	3	7	1	50.0	45.5
C12	2	3	4	2	7	7	4	4	62.5	63.6
C13	3	4	4	7	5	3	7	2	50.0	45.5
C14	3	1	2	7	4	6	4	3	50.0	45.5
C15	2	6	3	2	4	5	5	3	37.5	45.5
C16	3	5	4	6	6	7	1	7	50.0	54.5
C17	6	1	7	5	7	5	3	1	62.5	63.6
C18	5	2	2	7	3	6	6	6	50.0	54.5
C19	2	3	6	7	6	4	7	4	50.0	45.5
C20	6	6	5	7	3	1	1	3	0.0	0.0

with the histogram of matches:

Binary Sum profiling: Weighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
72.7	1	5.00	1	5.00
63.6	2	10.00	3	15.00
54.5	6	30.00	9	45.00
45.5	5	25.00	14	70.00
36.4	3	15.00	17	85.00
27.3	1	5.00	18	90.00
18.2	1	5.00	19	95.00
0	1	5.00	20	100.00

The calculation for case #6 (72.7% profile match) is:

© 2022 Cognadev Ltd.

The relevant data

- the target attributes [Attribute], the target profile [*Target*],
- a person's scores on those target attributes [*Person*],
- a binary indicator showing whether a person's scores match the profile target for each attribute [*Match*],
- the weight applied to each target match [*Weight*].

$$Attribute = \begin{bmatrix} \text{“Ambition”} \\ \text{“Honesty”} \\ \text{“Tolerance”} \\ \text{“Dominance”} \\ \text{“Sincerity”} \\ \text{“Impulsive”} \\ \text{“Extraversion”} \\ \text{“Conscientious”} \end{bmatrix} \quad Target = \begin{bmatrix} \text{“>= 5”} \\ \text{“> = 5”} \\ \text{“> = 5”} \\ \text{“< =4 ”} \\ \text{“> = 5”} \\ \text{“< = 4”} \\ \text{“between 3 \& 5”} \\ \text{“> = 5”} \end{bmatrix} \quad Person = \begin{bmatrix} 6 \\ 5 \\ 5 \\ 1 \\ 4 \\ 4 \\ 2 \\ 7 \end{bmatrix} \quad Match = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad Weight = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

The match % must be between 0 to 100%, which requires taking into account the sum of the weights applied to the binary matches [0, 1]. In this example, the person matches 6 of the 8 profile targets.

The weighted sum formula

$$Sum_weights := \sum Weight = 11 \qquad N_attributes := 8$$

$$Weighted_Sum := \left\| \begin{array}{l} \text{for } i \in 1..N_attributes \\ \quad sum \leftarrow sum + Match_i \cdot Weight_i \\ sum \end{array} \right| = 8$$

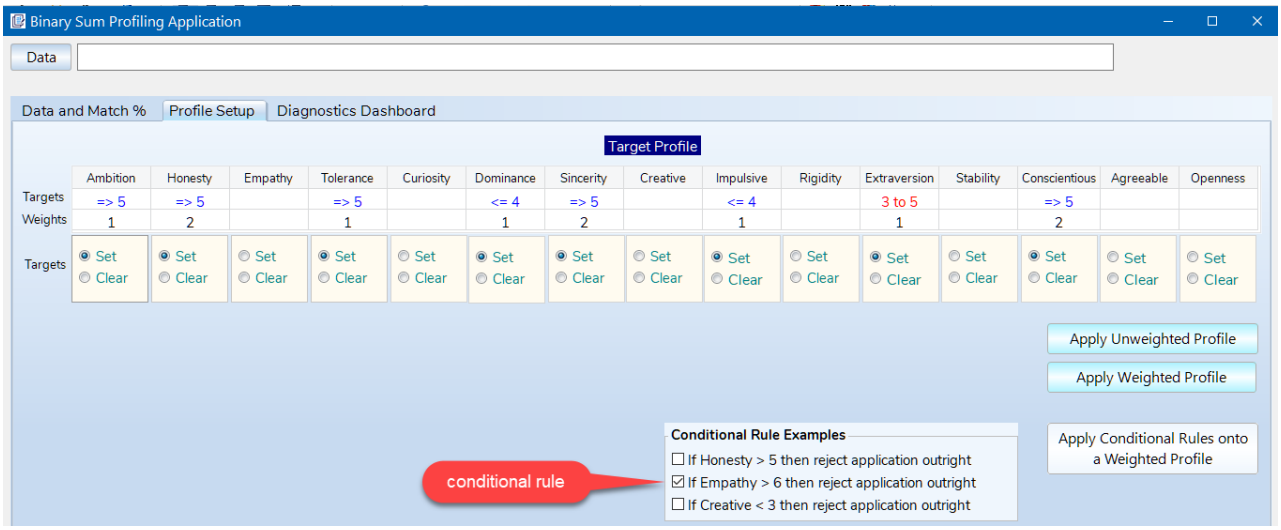
$$Match_percent := \left(\frac{Weighted_Sum}{Sum_weights} \right) = 72.727\%$$

If the profile matches had not been weighted, binary sum match for this case would have been 6/8=**75**%

3.1.3 Conditional matching

Again, using the same small-sample dataset as in Example 3.1.1 14, and the same target profile with weights defined for each target attribute as in Section 3.1.2 17, and now with a conditional rule applied:

© 2022 Cognadev Ltd.



The results are:

	Creative	Impulsive	Rigidity	Extraversion	Stability	Conscientious	Agreeable	Openness	Unweighted Match %	Weighted Match %	Conditional Match %
C1	6	4	2	5	4	7	7	4	62.5	54.5	54.5
C2	7	1	1	3	5	3	4	2	62.5	54.5	54.5
C3	5	4	3	3	7	1	1	3	62.5	54.5	54.5
C4	7	7	1	4	6	2	4	6	25.0	27.3	27.3
C5	2	4	7	1	7	1	2	3	50.0	36.4	36.4
C6	1	4	3	2	2	7	7	3	75.0	72.7	72.7
C7	6	3	5	1	4	4	6	2	37.5	36.4	36.4
C8	3	5	5	2	6	7	7	1	12.5	18.2	18.2
C9	1	2	5	2	5	7	7	1	50.0	54.5	54.5
C10	2	2	7	5	7	4	4	7	37.5	36.4	36.4
C11	1	1	2	2	7	3	7	1	50.0	45.5	45.5
C12	2	3	4	2	7	7	4	4	62.5	63.6	63.6
C13	3	4	4	7	5	3	7	2	50.0	45.5	45.5
C14	3	1	2	7	4	6	4	3	50.0	45.5	45.5
C15	2	6	3	2	4	5	5	3	37.5	45.5	0.0
C16	3	5	4	6	6	7	1	7	50.0	54.5	54.5
C17	6	1	7	5	7	5	3	1	62.5	63.6	63.6
C18	5	2	2	7	3	6	6	6	50.0	54.5	54.5
C19	2	3	6	7	6	4	7	4	50.0	45.5	45.5
C20	6	6	5	7	3	1	1	3	0.0	0.0	0.0

with match index histogram:

Binary Sum profiling: Conditionals match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
72.7	1	5.00	1	5.00
63.6	2	10.00	3	15.00
54.5	6	30.00	9	45.00
45.5	4	20.00	13	65.00
36.4	3	15.00	16	80.00
27.3	1	5.00	17	85.00
18.2	1	5.00	18	90.00
0	2	10.00	20	100.00

As can be seen compared to a weighted profile matching (no conditional rules - section 3.1.2 17), the only difference is that two cases' match indexes were set to 0, as their scores for Empathy exceeded the rule value of 6.

© 2022 Cognadev Ltd.

Any variety of rule/s can be set up as filters which can be applied either before or after the target profiling calculations.

3.2 Example_2: Large Candidate N

In this section, I use the 15 variable by 10,000 person dataset, the first 10 variables with scores distributed normally, the final 5 distributed as a right-skewed beta distribution.

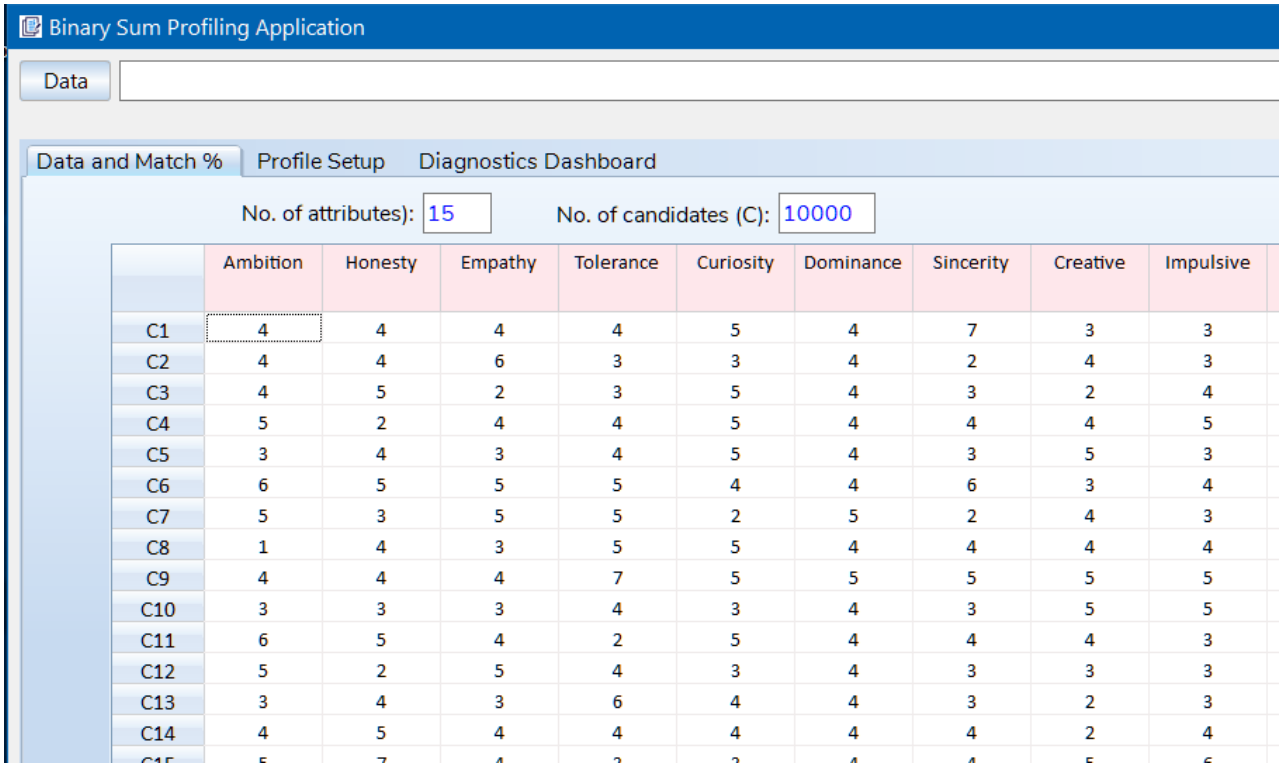
The examples in section 3.1 are deceptive in that with such few candidate numbers, it's straightforward to select the top few cases who best meet a target profile. As we see in this section, with 10,000 candidates being matched, the results are no longer of any practical utility if only a few attributes are chose for the target profile, as far too many candidates can possess very high index match coefficients.

Let's assume we are looking for the top 100 candidates who will be selected for assessment centre performance evaluation and interview.

3.2.1 Unweighted matching

Again, this the most straightforward use of binary sum profiling, where a "1" is recorded if a candidate score (or category) matches the target profile setting. A "0" is otherwise recorded.

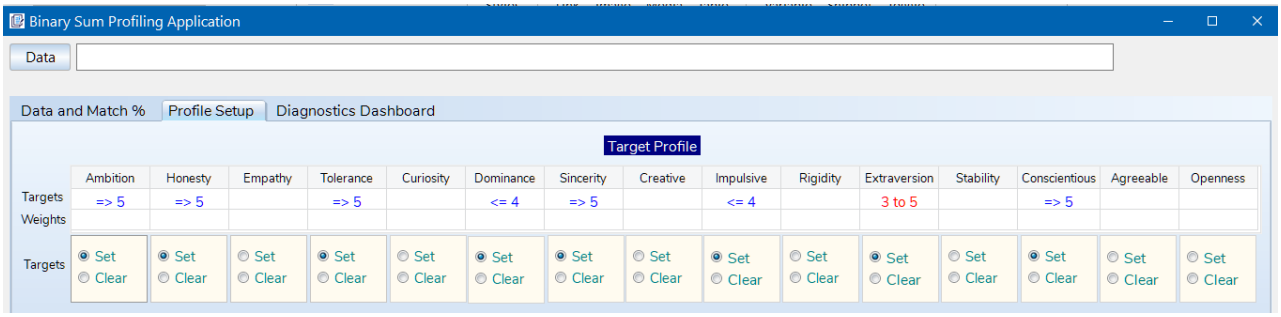
The data look like:



	Ambition	Honesty	Empathy	Tolerance	Curiosity	Dominance	Sincerity	Creative	Impulsive
C1	4	4	4	4	5	4	7	3	3
C2	4	4	6	3	3	4	2	4	3
C3	4	5	2	3	5	4	3	2	4
C4	5	2	4	4	5	4	4	4	5
C5	3	4	3	4	5	4	3	5	3
C6	6	5	5	5	4	4	6	3	4
C7	5	3	5	5	2	5	2	4	3
C8	1	4	3	5	5	4	4	4	4
C9	4	4	4	7	5	5	5	5	5
C10	3	3	3	4	3	4	3	5	5
C11	6	5	4	2	5	4	4	4	3
C12	5	2	5	4	3	4	3	3	3
C13	3	4	3	6	4	4	3	2	3
C14	4	5	4	4	4	4	4	2	4

For comparability purposes, I'm using the same profile setup as in section 3.1.1 14:

© 2022 Cognadev Ltd.



Matching our 10,000 candidates to this profile produces the histogram of matches:

Binary Sum profiling: Large sample unweighted match frequencies

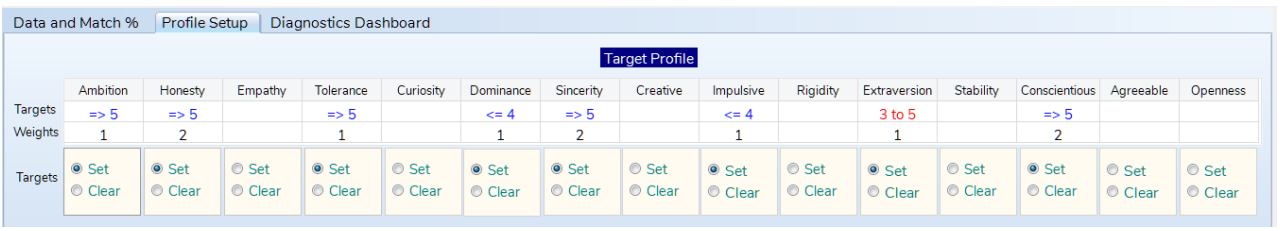
Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	11	0.0	11	0.0
87.5	123	1.0	134	1.0
75	646	7.0	780	8.0
62.5	1,911	19.0	2,691	27.0
50	3,060	31.0	5,751	58.0
37.5	2,746	28.0	8,497	85.0
25	1,263	13.0	9,760	98.0
12.5	229	2.0	9,989	100.0
0	11	0.0	10,000	100.0

So, we have 11 candidates with 100% profile match who are a 'shoe-in' for further assessment, with the next-best match of 87.5% possessed by 123 others. Given we only want to select 100 candidates for further assessment, we would have to choose the remaining 89 candidates from the 123 - using information other than that provided by our target profile.

Things might become clearer if we used weighted profile matching.

3.2.2 Weighted matching

So, here we use the same weights as applied in section 3.1.2 17:



which results in the following match statistics histogram:

© 2022 Cognadev Ltd.

Binary Sum profiling: Large sample weighted match frequencies				
Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	11	0	11	0
90.9	80	1	91	1
81.8	258	3	349	4
72.7	656	7	1,005	10
63.6	1,356	14	2,361	24
54.5	1,971	20	4,332	43
45.5	2,170	22	6,502	65
36.4	1,957	20	8,459	85
27.3	1,088	11	9,547	96
18.2	338	3	9,885	99
9.1	104	1	9,989	100
0	11	0	10,000	100

So, we have the same 11 candidates with 100% profile match who are a 'shoe-in' for further assessment, with the next-best match of 90.9% possessed by 80 others. Given we want to select 100 candidates for further assessment, we would have to choose the remaining 9 candidates from the 258 with the next best match (81.8%) - using information other than that provided by our target profile. But, having 91 "potentials" with ~91% match and above, things are looking a lot better in terms of candidate quality (given the design and weighting applied in that target profile).

3.2.3 Conditional matching

Let's assume we want to also screen-out those with a Creativity score <= 3, and a too high Empathy score (6 and above). So we apply two conditional rules to the weighted profiling. The results histogram is:

Binary Sum profiling: Large sample conditionals match frequencies				
Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	10	0	10	0
90.9	72	1	82	1
81.8	238	2	320	3
72.7	614	6	934	9
63.6	1,262	13	2,196	22
54.5	1,823	18	4,019	40
45.5	1,997	20	6,016	60
36.4	1,837	18	7,853	79
27.3	997	10	8,850	89
18.2	309	3	9,159	92
9.1	99	1	9,258	93
0	742	7	10,000	100

© 2022 Cognadev Ltd.

We now have 82 candidates with profile match indexes of over 90%, with 238 with 81.8% match from which to select the remaining 18 cases (our 100 to go forward to the assessment centre/interview stage).

But note that our conditional screening rules have removed 742 candidates from profile matching, compared to just 11 using unweighted or weighted profiling (some of whom now rejected would have been selected using just the weighted profiling solution). So, conditional screening is powerful but you have to be certain that the attributes you enter into such 'screen-out' rules really are what you want.

Of course, if you are not using dynamic autonomous person-target profiling (*candidates are matched against the profile as they apply*) but instead profiling a static database of applicants *(all candidates complete the initial screening assessment)* where on a closing application date profiling is conducted on the entire applicant pool data, then it is possible to tweak a target profile and 'try-out' several options according the characteristics/numbers of candidates being selected by each target profile. Likewise, this explorative approach can be used if you have a database of candidates or even employees which might be profiled against a range of target profiles.

© 2022 Cognadev Ltd.

4. Benefits and Drawbacks

4. Benefits and Drawbacks

The Benefits of Binary Sum Profiling are:

1. Simple to compute/apply and understand.
2. Weighting and applying conditional rules is an efficient way of finessing a simple profile.
3. It can be used with categorical as well as quantitative data, as the decision is always binary (the target is either met or not met).
4. Works well on small numbers of candidates where only a single or few high-match candidates are being selected.

But, it has one near-terminal drawback.

As the number of candidates *increase* and the number of attributes in a target *decrease*, the profile match statistics render any selection solution as untenable.

Example 1: the Large Sample dataset (n=10,000 cases) with a target profile using only the Big Five attributes:

Extraversion	Stability	Conscientious	Agreeable	Openness
3 to 5	=> 5	=> 5	=> 3	=> 3
◉ Set ○ Clear	◉ Set ○ Clear	◉ Set ○ Clear	◉ Set ○ Clear	◉ Set ○ Clear

which results in the following profile-match histogram:

Binary Sum profiling: Unweighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	2,051	21.0	2,051	21.0
80	6,185	62.0	8,236	82.0
60	1,634	16.0	9,870	99.0
40	125	1.0	9,995	100.0
20	5	0.0	10,000	100.0

And there's the problem .. **2,051** candidates who match the profile 100% .. with no means of choosing between them except using other information, which is entirely impractical with these numbers.

Maybe if we weighted the attributes?

© 2022 Cognadev Ltd.

Extraversion	Stability	Conscientious	Agreeable	Openness
3 to 5	=> 5	=> 5	=> 3	=> 3
1	2	3	1	1.5
Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear

which results in the following profile-match histogram:

Binary Sum profiling: Weighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	2,051	21	2,051	21
88.2	5,621	56	7,672	77
82.4	20	0	7,692	77
76.5	320	3	8,012	80
70.6	69	1	8,081	81
64.7	1,047	11	9,128	91
58.8	2	0	9,130	91
52.9	711	7	9,841	98
47.1	7	0	9,848	99
41.2	37	0	9,885	99
35.3	7	0	9,892	99
29.4	103	1	9,995	100
17.6	1	0	9,996	100
11.8	4	0	10,000	100

which doesn't help at all as we still have 2,051 cases at 100% match accuracy. All we can do is to make the profile more 'discriminating' by adjusting the target values:

Extraversion	Stability	Conscientious	Agreeable	Openness
3 to 5	=> 6	4 to 6	4 to 6	=> 5
Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear

which results in the following profile-match histogram:

© 2022 Cognadev Ltd.

Binary Sum profiling: Unweighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	622	6.0	622	6.0
80	2,700	27.0	3,322	33.0
60	3,795	38.0	7,117	71.0
40	2,282	23.0	9,399	94.0
20	556	6.0	9,955	100.0
0	45	1.0	10,000	100.0

which reduces the 100% match candidates to 622 - which might be more useful if these are also considered for a 2nd-stage pre-screen before final selection, for say interviews or a more financially expensive assessment. But, by making the profile more 'discriminating', you get the feeling we are only doing this in order to cut back on 100% numbers rather than for more substantive psychological reasons.

Example 2: What if we use some real data (from a work-preferences + personality database), where each attribute is measured using a 1-100 range using a Visual Analogue Scale. Here we have 994 cases of data - from which we wished to choose our top 50 candidates. The target profile is:

Target Profile

	Clarity	Challenge	Affiliation	Autonomy	Conceptual	Intuitive	Evaluative	Curiosity	Humility	Accountable	Extraversion	Stability	Conscientious	Agreeable	Openness
Targets	=> 70				=> 70	=> 70		=> 70		50 to 90			40 to 80		=> 70
Weights															
Targets	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Clear / Set	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear	Set / Clear

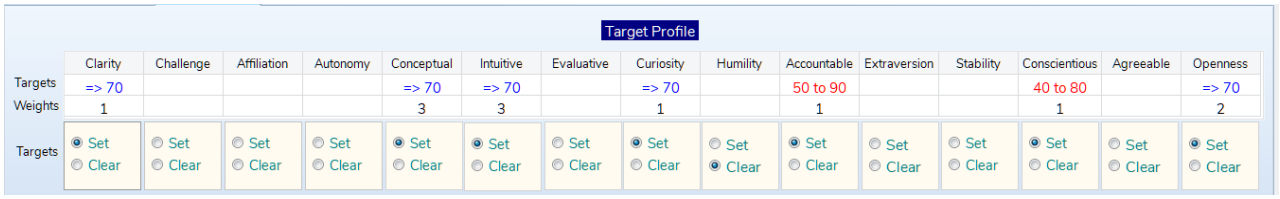
which results in the following profile-match histogram:

Binary Sum profiling: Real Data, unweighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	4	0.0	4	0.0
85.7	54	5.0	58	6.0
71.4	193	19.0	251	25.0
57.1	318	32.0	569	57.0
42.9	272	27.0	841	85.0
28.6	127	13.0	968	97.0
14.3	25	3.0	993	100.0
0	1	0.0	994	100.0

which makes our selection task both manageable and straightforward as we wanting to choose our top 50 candidates. We might even use a weighted profile to further finesse our selection:

© 2022 Cognadev Ltd.

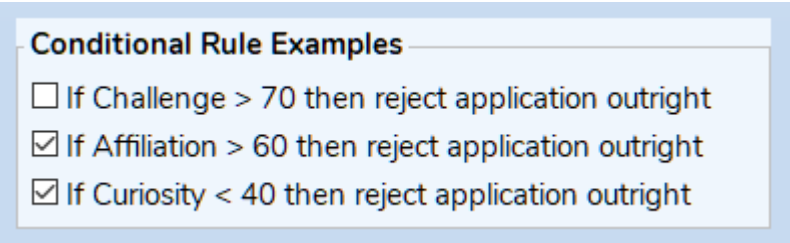


which results in the following profile-match histogram:

Binary Sum profiling: Real Data, weighted match frequencies

Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	4	0.0	4	0.0
91.7	12	1.0	16	2.0
83.3	48	5.0	64	6.0
75	93	9.0	157	16.0
66.7	70	7.0	227	23.0
58.3	89	9.0	316	32.0
50	177	18.0	493	50.0
41.7	107	11.0	600	60.0
33.3	99	10.0	699	70.0
25	165	17.0	864	87.0
16.7	105	11.0	969	98.0
8.3	24	2.0	993	100.0
0	1	0.0	994	100.0

Which again makes selection of the top 50 candidates very straightforward. Or be even more discriminating by screening out cases according to two rules, while retaining the weighting:



which results in the following profile-match histogram:

© 2022 Cognadev Ltd.

Binary Sum profiling: Real Data, conditionals match frequencies				
Match %	Count	% Count	Cumulative Count	Cumulative Count %
100	3	0.0	3	0.0
91.7	9	1.0	12	1.0
83.3	9	1.0	21	2.0
75	25	3.0	46	5.0
66.7	23	2.0	69	7.0
58.3	24	2.0	93	9.0
50	57	6.0	150	15.0
41.7	43	4.0	193	19.0
33.3	34	3.0	227	23.0
25	63	6.0	290	29.0
16.7	42	4.0	332	33.0
8.3	8	1.0	340	34.0
0	654	66.0	994	100.0

Note that we have lost 654 candidates with the use of the two rules. That may not be acceptable or desirable .. much depends upon the justifications for the two rules.

© 2022 Cognadev Ltd.

5. My Conclusions

5. My Conclusions

Well, it's a simple, straightforward methodology, but because of its binary "match_no-match" format, it lacks any property of 'nearness' or 'closeness', that all other methodologies possess.

Yes, you can adjust the method slightly using weightings, but those weights are being applied to binary decisions, unlike other methods which incorporate more graduated matches i.e. nearness to a target value or range of values - not just match_no-match.

So, for me, binary sum profiling is not recommended except where the logic for such categorical decision-making can be justified.

It is always going to perform badly when just a few attributes are profiled using a large candidate base. However, if using it as part of an autonomous sequential assessment/selection process (*i.e. applicants are screened and profile matched as they apply*), then if looking for say the first 20 candidates to fit the profile, it will work well. But, if the entire applicant database is analyzed on a closing date, it may be that you end up with thousands of candidates and a few attributes, with the unfortunate consequences we see in section 4 26, Example 1.

So, much depends upon how discriminating is a target profile, given large numbers of candidates are to be searched upon. And, whether a target is a simple "greater than" or "less than" bound - or uses a range-target "from x1 to x2", which is equivalent to a bandpass filter-function profile with zero roll-off either side.

Which brings into focus the issue of matching to a fixed constant or unit-weighted 'region-of-interest' -vs- a point-estimate target (a single number or class-category).

The former applies a constant weight (*whether 0 or some positive value*) to any score value that is within (*or outside*) the designated profile target bound or area; the latter method can incorporate directional and non-symmetric nearness functions around a point-estimate or bound..

Finally before any profile solution is deployed, it is always prudent to sense-check one or more target profiles for unexpected outcomes, using actual or realistic simulated data that matches the expected data from applicants, especially if wanting to deploy a target profile into an autonomous screen-out/in system (e.g. *an Applicant Tracking System deployed online*).

© 2022 Cognadev Ltd.

- B -

Binary Sum calculation 6

- C -

conditionals 11

- D -

Definition 6

- E -

Equation_1 6

Example_1 14

Example_2 21

- H -

hybrid profile 10

- L -

Lage_N Weighted Profile Matching 21

- N -

non-compensatory 7

- P -

production-rules 10

- R -

rules 11

- S -

Small_N Unweighted Profile Matching 14

Small_N Weighted Profile Matching 14

- U -

unifinal 7

- W -

weights 10

© 2022 Cognadev Ltd.