

Separated but not Sufficient

Let **pr** = prediction, **hl** = hyperlipidemia, **gd** = gender, **ht** = hypertension ;Our prediction is based on

$$P(\text{Hyperlipidemia} = \text{Yes} | \text{Hypertension}) > 0.4$$

According to our VE algorithm, we have:

$$P(hl = \text{yes} | ht = \text{NO}) = 0.35, P(hl = \text{yes} | ht = \text{NO}, gd = \text{Male}) = 0.38, \\ P(hl = \text{yes} | ht = \text{NO}, gd = \text{Female}) = 0.31$$

$$P(hl = \text{yes} | ht = \text{YES}) = 0.46, P(hl = \text{YES} | ht = \text{YES}, gd = \text{Male}) = 0.5, \\ P(hl = \text{YES} | ht = \text{YES}, gd = \text{Female}) = 0.42$$

$$P(pr | hl, gd)$$

pr	hl	gd	$P(pr hl, gd)$
YES	YES	Male	0.56
YES	YES	Female	0.56
YES	NO	Male	0.44
YES	NO	Female	0.44

We can see from the table that this is a clear separation between prediction and gender since the value of gender doesn't affect the final result of probability at all.

$$P(hl | pr, gd)$$

hl	pr	gd	$P(hl pr, gd)$
YES	YES	Male	0.5
YES	YES	Female	0.42
YES	NO	Male	0.38
YES	NO	Female	0.31

We can clearly see there's no sufficiency here because the prob changes according to the gender.

1 Sufficient but not Separated

Let **rg** = region

Our prediction is based on:

$$P(db = \text{YES} | hl)$$

$$P(pr | db, ht)$$

pr	db	ht	$P(pr db, ht)$
YES	YES	YES	0.5
YES	YES	NO	0.59
YES	NO	YES	0.40
YES	NO	NO	0.22

So we find there's no separation here

$P(db|pr, ht)$

db	pr	ht	$P(db pr, ht)$
YES	YES	YES	0.58
YES	YES	NO	0.58
YES	NO	YES	0.60
YES	NO	NO	0.60