VSM Model — wedding gown

Train document set:
d1 → user selected wedding gown
d2 → user ordered on-line rose flowers

Test document set:
d3 → user ~~ordered~~ searched diamond ring
d4 → user selected white wedding gown, online flowers,
     3 carat diamond ring

→ Creating index Vocabulary of words of train document set

$$
E(t) = \begin{cases} 1, & \text{user} \\ 2, & \text{selected} \\ 3, & \text{wedding} \\ 4, & \text{gown} \\ 5, & \text{ordered} \\ 6, & \text{online} \\ 7, & \text{rose} \\ 8, & \text{flowers} \end{cases}
$$

→ Calculating term frequency as
$$ tf(t, d) = \sum_{x \in d} fr(x, t). $$

⇒ $Vd_3 = (1, 0, 0, 0, 0, 0, 0, 0)$
⇒ $Vd_4 = (1, 1, 1, 1, 0, 1, 0, 1)$

⇒ $M_{|D| \times F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$