

U: Votre Apotheke en li
-10 pilu x 100 mg + Cilia
raison gratuite
stème de commande sûr

text document (email)

2x

| \mathbf{X} | | \mathbf{X}_l |
|--------------|---|------------------|
| NEU | | USER123_NEU |
| Votre | | USER123_Votre |
| Apotheke | + | USER123_Apotheke |
| en | | USER123_en |
| ligne | | USER123_ligne |
| Euro | | USER123_Euro |
| ... | | ... |

tokenized, duplicated
bag of words

h

| \mathbf{X}_h |
|----------------|
| 323 |
| 0 |
| 5235 |
| 0 |
| 0 |
| 123 |
| 0 |
| 626 |
| 232 |
| ... |

hashed,
sparse vector

$\mathbf{w}^\top \mathbf{X}_h$

classification