



LOG3430 -Méthodes de test et de validation du logiciel

TP3 – Test d’interactions

Groupe 1

Équipe 7

Hugo Lachieze-Rey (1934177)

Dimitry Kamga (1898357)

Remis à :

Hanane Ikhelef

Hiver 2022

Tailles des matrices MCA

Selon le type d'interaction, on obtient les 3 matrices ci-dessous :

- $CA(N, t, k, v) = CA(8, 2, 4, \{2, 2, 2, 4\}) \rightarrow$ interactions doubles ;
- $CA(N, t, k, v) = CA(16, 3, 4, \{2, 2, 2, 4\}) \rightarrow$ interactions triples ;
- $CA(N, t, k, v) = CA(32, 4, 4, \{2, 2, 2, 4\}) \rightarrow$ interactions quadruples ;

Analyse des résultats

Pour les interactions doubles, nous obtenons les résultats suivants :

Cas	log_prob	log_combine	vocab	clean_opt	Accuracy	Precision	Recall
1	TRUE	FALSE	1	1	0.51	0.4	0.57
2	FALSE	TRUE	1	0	0.82	0.74	0.83
3	TRUE	TRUE	2	1	0.51	0.4	0.57
4	FALSE	FALSE	2	0	0.65	0.53	0.83
5	TRUE	TRUE	3	0	0.53	0.41	0.49
6	FALSE	FALSE	3	1	0.45	0.41	1
7	TRUE	TRUE	4	0	0.53	0.41	0.49
8	FALSE	FALSE	4	1	0.41	0.4	1

Pour l'interaction triple, nous obtenons les résultats suivants :

Cas	log_prob	log_combine	vocab	clean_opt	Accuracy	Precision	Recall
1	TRUE	TRUE	1	0	0.52	0.39	0.48
2	TRUE	FALSE	1	1	0.51	0.4	0.57
3	FALSE	TRUE	1	1	0.79	0.69	0.81
4	FALSE	FALSE	1	0	0.74	0.64	0.78
5	TRUE	TRUE	2	1	0.51	0.4	0.57
6	TRUE	FALSE	2	0	0.52	0.4	0.52
7	FALSE	TRUE	2	0	0.76	0.63	0.87
8	FALSE	FALSE	2	1	0.51	0.44	1
9	TRUE	TRUE	3	0	0.53	0.4	0.49
10	TRUE	FALSE	3	1	0.5	0.39	0.56
11	FALSE	TRUE	3	1	0.59	0.48	0.99
12	FALSE	FALSE	3	0	0.62	0.51	0.83
13	TRUE	TRUE	4	0	0.53	0.41	0.49
14	TRUE	FALSE	4	1	0.49	0.39	0.55
15	FALSE	TRUE	4	1	0.49	0.43	0.99
16	FALSE	FALSE	4	0	0.56	0.46	0.87

Pour les interactions quadruples, nous obtenons les résultats suivants :

Cas	log_prob	log_combine	vocab	clean_opt	Accuracy	Precision	Recall
1	TRUE	TRUE	1	0	0.52	0.39	0.48
2	TRUE	TRUE	1	1	0.51	0.4	0.57
3	TRUE	FALSE	1	0	0.52	0.39	0.48
4	TRUE	FALSE	1	1	0.51	0.4	0.57
5	FALSE	TRUE	1	0	0.82	0.74	0.83
6	FALSE	TRUE	1	1	0.79	0.69	0.81
7	FALSE	FALSE	1	0	0.74	0.64	0.78
8	FALSE	FALSE	1	1	0.64	0.52	0.75
9	TRUE	TRUE	2	0	0.52	0.4	0.52
10	TRUE	TRUE	2	1	0.51	0.4	0.57
11	TRUE	FALSE	2	0	0.52	0.4	0.52
12	TRUE	FALSE	2	1	0.51	0.4	0.57
13	FALSE	TRUE	2	0	0.76	0.63	0.87
14	FALSE	TRUE	2	1	0.7	0.57	1
15	FALSE	FALSE	2	0	0.65	0.53	0.83
16	FALSE	FALSE	2	1	0.51	0.44	1
17	TRUE	TRUE	3	0	0.53	0.4	0.49
18	TRUE	TRUE	3	1	0.5	0.39	0.56
19	TRUE	FALSE	3	0	0.53	0.4	0.49
20	TRUE	FALSE	3	1	0.5	0.39	0.56
21	FALSE	TRUE	3	0	0.69	0.56	0.86
22	FALSE	TRUE	3	1	0.59	0.48	0.99
23	FALSE	FALSE	3	0	0.62	0.51	0.83
24	FALSE	FALSE	3	1	0.45	0.41	1
25	TRUE	TRUE	4	0	0.53	0.41	0.49
26	TRUE	TRUE	4	1	0.49	0.39	0.55
27	TRUE	FALSE	4	0	0.53	0.41	0.49
28	TRUE	FALSE	4	1	0.49	0.39	0.55
29	FALSE	TRUE	4	0	0.59	0.48	0.88
30	FALSE	TRUE	4	1	0.49	0.43	0.99
31	FALSE	FALSE	4	0	0.56	0.46	0.87
32	FALSE	FALSE	4	1	0.41	0.4	1

En premier lieu, on remarque que le calcul de probabilité (formule avec ou sans logarithme) et le calcul de la combinaison des probabilités n'influencent pas nos métriques (*accuracy*, *precision* et *recall*). On pouvait s'y attendre car les deux méthodes de calculs sont similaires.

Ensuite, on peut remarquer que les deux autres paramètres, soit la création du vocabulaire (*vocab*) et le nettoyage de texte (*clean_opt*) exercent une influence sur nos métriques.

On peut remarquer que le paramètre *clean_opt* a une influence sur les résultats car il permet de faire le stemming (ou non) et d'enlever les « stop words ». En effet, on remarque que les métriques de performances varient positivement lorsqu'on n'effectue pas le stemming.

Lors de la création du vocabulaire, on peut choisir de faire varier la fréquence des mots minimale pour ajouter un mot dans le dictionnaire. On peut remarquer lorsque la fréquence des mots minimale augmente, l'*accuracy* et la précision du système diminuent. Cependant, on peut voir que le *recall* du système a tendance à augmenter légèrement.

En se basant sur les interactions triples, on peut voir que nos meilleurs résultats sont obtenus avec le cas de test 5 soit lorsque la fréquence des mots minimale est égale à 1, que nous effectuons le stemming et que nous enlevons les « stop words » (*clean_opt* = 0). La méthode de calcul de probabilité ou la combinaison de probabilité n'exercent aucune influence sur nos résultats. Un autre résultat intéressant est le cas de test 6 qui est similaire au cas de test 5 mais le paramètre *clean_opt* vaut 1 donc nous n'effectuons pas le stemming et nous enlevons les « stop words ».

Une amélioration du système pourrait être de faire le stemming et de ne pas enlever les « stop words ». Ainsi, la performance du système serait améliorée.

Défauts :

Les défauts remarqués cependant concerne le calcul de la probabilité. On a remarqué que les variables *p_subject_spam*, *p_subject_ham*, *p_body_spam*, *p_body_ham* sont délicates à manipuler. En effet ces valeurs sont la plupart du temps proche de 0. Donc le logarithme de ces valeurs tendra vers $-\infty$ et par conséquent élèvera des exceptions pendant l'évaluation des emails. Nous avons appliqué `math.log10()` pour nous permettre d'avoir des valeurs plus précises lorsque nous évaluons une probabilité proche de 0.