

Nom:	
Prénom:	
Numéro étudiant:	

note

### Correction

Questions	Réponses
1. On parle d'une approche <i>bag of words</i> (« sac de mots ») quand	<input type="checkbox"/> les mots sont rassemblés par groupes en fonction de leur fréquence <input type="checkbox"/> seuls les mots non grammaticaux sont conservés dans une analyse <input checked="" type="checkbox"/> les mots sont considérés uniquement en fonction de leur fréquence et pas de l'ordre d'apparition
2. Un corpus <i>parallèle</i> est un corpus	<input type="checkbox"/> qui contient des textes répétés <input checked="" type="checkbox"/> qui contient des passages dans plusieurs langues associés entre eux <input type="checkbox"/> qui contient des textes comparables de plusieurs genres littéraires
3. L' <i>exactitude</i> d'un système mesure	<input checked="" type="checkbox"/> le nombre de bonnes réponses divisé par le nombre de réponses <input type="checkbox"/> à quel point le système commet beaucoup d'erreurs de types différents <input type="checkbox"/> le degré de satisfaction des utilisateurs
4. Les systèmes de reconnaissance optique des caractères (OCR)	<input type="checkbox"/> ne fonctionnent pas avec des textes manuscrits <input checked="" type="checkbox"/> améliorent leurs performance avec des modèles de langue <input type="checkbox"/> sont basés sur des règles élaborées depuis des années