



2.3 Apprentissage de la représentation de mots

Dans cette section, on note V le vocabulaire et $|V|$ sa taille.

2.3.1 Motivation et notations

□ **Techniques de représentation** – Les deux manières principales de représenter des mots sont décrits dans le tableau suivant :

Représentation binaire	Représentation du mot
<ul style="list-style-type: none">- Noté o_w- Approche naïve, pas d'information de similarité	<ul style="list-style-type: none">- Noté e_w- Prend en compte la similarité des mots

□ **Matrice de représentation** – Pour un mot donné w , la matrice de représentation (en anglais *embedding matrix*) E est une matrice qui relie une représentation binaire o_w à sa représentation correspondante e_w de la manière suivante :

$$e_w = E o_w$$

Remarque : l'apprentissage d'une matrice de représentation peut être effectuée en utilisant des modèles probabilistiques de cible/contexte.