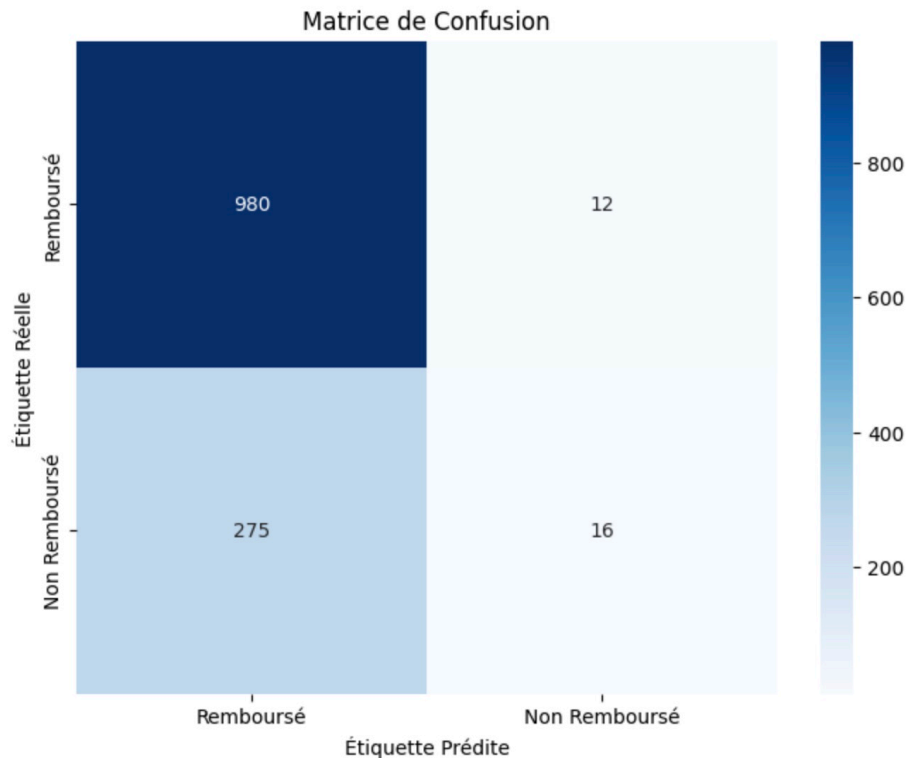


## Rapport TP1

1. Le dataset qu'on a utilisé n'est pas équilibré. En effet on a plus de 8000 valeurs `not.fully.paid = 0`, et environs 1000 valeurs `not.fully.paid = 1`
2. Trainer le modèle sur ces données va nous mener à un résultat pas très satisfaisant, et surtout un problème d'overfitting, car notre modèle détecte les valeurs négatives plus que quelque chose d'autre.



3. Pour ce remédier à ce problème on doit essayer d'équilibrer notre dataset, pour ce faire moi j'ai essayé la méthode d'underSampling, avec un hyperparametre qui dit je supprime la classe majoritaire jusqu'à ce que j'ai 5% de plus par rapport à la classe minoritaire. Sauf que ceci n'a pas complètement résolu mon problème, car après ceci mon modèle est un peu aléatoire.

	precision	recall	f1-score	support
0	0.63	0.61	0.62	339
1	0.56	0.58	0.57	291
accuracy			0.60	630
macro avg	0.59	0.59	0.59	630
weighted avg	0.60	0.60	0.60	630

