

# Roberta\_93

## 1. Modèle de detection NLI pré-entraîné

Pour une version en anglais de notre API nous avons utilisé un modèle pré-entraîné issu de AllenNLP[1] entraîné sur les données SNLI qui est un ensemble de données seulement en anglais, représentant les légendes d'images, où ces dernières sont utilisées comme **prémisses**, les **hypothèses** sont créés manuellement par les travailleurs de Mechanical Turk.

## 2. Test de modèle pré-entraîné:

Données tests SNLI:

	sentence1	sentence2
gold_label		
contradiction	100	100
entailment	104	104
neutral	96	96

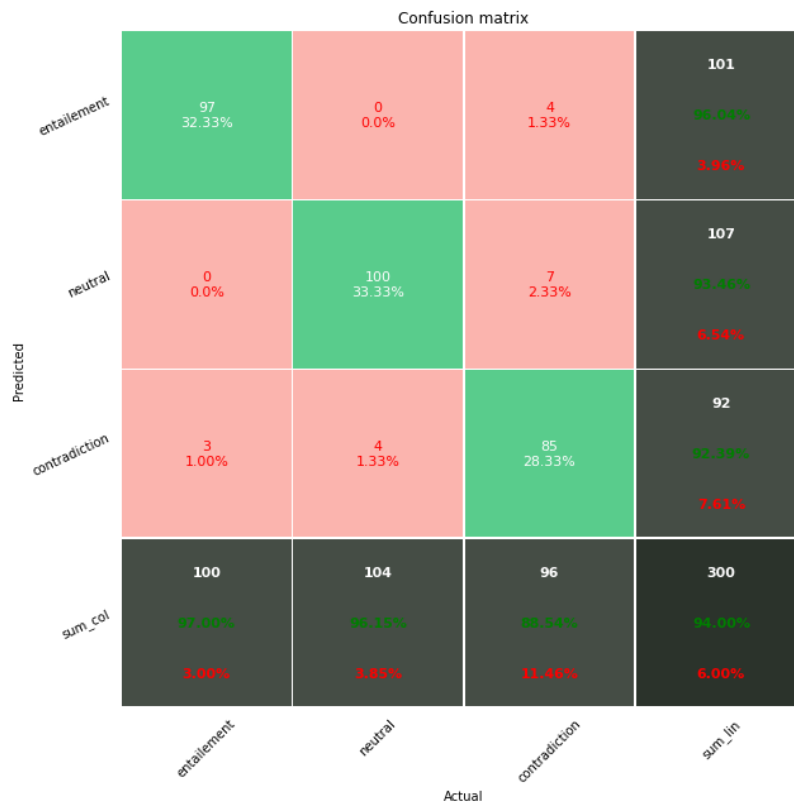
figure1: Échantillon de données test SNLI

Accuracy:

set	accuracy
0 test	94.0

figure2: Precision de prediction sur les données SNLI test

Matrice de confusion:



Test de modèle sur les données contradictory:

Données tests contradictory:

Les données tests sont issues de jeu de données Contradictory, on a pris ce choix parce que ces données seront nouvelles pour le modèle pré-entraîné avec SNLI, cela nous permettra de tester vraiment la fiabilité de modèles et les performances enregistrés par ce modèle.

Les données tests sont comme suit:

	premise	hypothesis
gold_label		
0	172	172
1	165	165
2	163	163

figure2:Données test

Accuracy:

set	accuracy
0 test	83.8

figure3: accuracy de données test contadictory

		Confusion matrix			
Predicted	entailment	148 29.60%	24 4.80%	9 1.80%	181 81.77% 18.23%
	contradiction	19 3.80%	128 25.60%	11 2.20%	158 81.01% 18.99%
	neutral	5 1.00%	13 2.60%	143 28.60%	161 88.82% 11.18%
	sum_col	172 86.05% 13.95%	165 77.58% 22.42%	163 87.73% 12.27%	500 83.80% 16.20%
		entailment	contradiction	neutral	sum_lin
		Actual			

#### Discussion:

Nous remarquons une différence dans les performance de modèle sur les deux tests, tel que nous avons enregistré une baisse de performance de prédiction sur les données test contradictory, cela s'explique de fait que ces données sont nouvelles pour le modèles pré-entraîné, car se sont des données différentes de SNLI, rappelons que SNLI sont des données constituer à partir des légendes d'images, où ces dernières sont utilisées comme **prémisses**, les **hypothèses** sont créés manuellement par les travailleurs de Mechanical Turk.

De la nous constatons que l'issue (genre la manière dont elle se constitue ou leurs types comme snli c'est les légendes d'images, mais contradictoire on sait pas d'où elle proviennent) des données influence sur la qualité de modèle.