

PERSONALIZED DATA-TO-TEXT NEURAL GENERATION

Génération des notes. Style utilisateur

4 juillet 2023

Ben KABONGO

Stage - M1 DAC - Sorbonne Université

- Génération des notes
- Style utilisateur

Génération des notes

- **But** : Générer une note de film étant donné un utilisateur
- **Dataset** : Rotten Tomatoes : description de films, critiques et notes utilisateur
- **Entrée** : Concaténation de la description du film et la critique de l'utilisateur
- **Sortie** : Note du film
- **Modèle** : Prompt Tuning avec T5
- **Taille du train set** : 25 000 exemples
- **Evaluation** : + 600 000 exemples

[illegible]

Génération des notes

Exemples de sortie du modèle

- *< pad > 5.0 < /s >*
- *< pad >< extra_id_0 > of < extra_id_1 > of < extra_id_2 > of < extra_id_3 > of < extra_id_4 > of < extra_id_5 > of < extra_id_6 >< unk > movie > GemmaBovery < unk > /title >< unk > /info >< unk > /title >< unk > /title >< unk > /title >< unk > /title >< unk > /title >< unk > /title >< unk > /title >< extra_id_10 > . < extra_id_13 > . < extra_id_14 > . < extra_id_15 > . < extra_id_16 > . < extra_id_17 > . < extra_id_18 >< extra_id_19 >>< extra_id_20 >>< extra_id_21 >>< extra_id_22 >>< extra_id_23 > || < extra_id_24 >*

- **Modèle : Prompt tuning avec T5**
- **Variantes de T5**
 - Small
 - Base
 - Large
- **Tailles de prompt : 10, 20, 30, 40, 50, 100**

Génération des notes

Métriques

■ Métriques textuelles :

- BLEU
- ROUGE

■ Métriques sur les notes :

- Récupération des notes dans la sortie des modèles avec une expression régulière
- Note mise à -1 sinon
- Nous calculons les métriques en tenant compte de toutes les notes (1) ou en considérant que les sorties convertissables (2)
- Mean Average Error (**MAE**) : $\frac{1}{n} \sum^n |x - \hat{x}|$
- Mean Squared Error (**MSE**) : $\frac{1}{n} \sum^n (x - \hat{x})^2$
- Rooted Mean Squared Error (**RMSE**) : $\sqrt{\frac{1}{n} \sum^n (x - \hat{x})^2}$

	t5-small prompt size=10	t5-small prompt size=50
BLEU	6.18e-08	0.0
ROUGE	0.022	0.024
MAE	437.23	3.70
MSE	8,873,660,548.23	16.13
RMSE	94,200.11	4.02
MAE 2	2,841.77	0.64
MSE 2	58,144,527,612.88	0.74
RMSE 2	241,131.76	0.86

Table 1 – Résultats des métriques pour les variantes de t5

	mean_prediction	muser_prediction	mitem_prediction	opt_bl_prediction	svd_prediction	knn_prediction
metrics						
mae	0.873207	0.813881	0.769587	0.703705	0.596406	0.648000
mse	1.167515	1.019638	0.943279	0.797493	0.597603	0.699307
rmse	1.080516	1.009771	0.971225	0.893025	0.773048	0.836246

Figure 1 – Recommandation - Résultats

Family	Mean	Std
Function words	1.83	2.62
Indexes	2.20	3.17
Letters	1.16	1.84
NER	1.87	1.03
Numbers	2.76	1.52
Punctuation	4.03	7.62
Structural	2.97	2.72
TAG	2.31	2.20

Table 2 – Styles utilisateurs Terreau, Gourru et Velcin 2020

Création du dataset de data-to-text personnalisé

- **STRAP** : méthode de transfert de style par génération de paraphrase avec GPT-2. Krishna, Wieting et Iyyer 2020
- **Entrées du modèles** : datasets de corpus utilisateurs : critiques de Rotten Tomatoes
- **Métriques** : transfer accuracy, semantic similarity, fluency, métrique agrégée, évaluation humaine
- **Dataset de data-to-text personnalisé**
 - Entraîner STRAP sur les critiques de Rotten Tomatoes
 - Inférer des descriptions personnalisées en partant des descriptions non personnalisées de WikiRoto

Références



Krishna, Kalpesh, John Wieting et Mohit Iyyer (2020). “Reformulating Unsupervised Style Transfer as Paraphrase Generation”. In : *Empirical Methods in Natural Language Processing*.



Terreau, Enzo, Antoine Gourru et Julien Velcin (juill. 2020). “Writing Style Author Embedding Evaluation”. In : *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Université de Lyon, Lyon 2, ERIC UR3083, p. 84-93. url : https://doi.org/10.26615/978-954-452-056-4_009.