

話者情報に条件づけられた対話モデルにおける話者情報を抽出する方法の比較分析

安川 浩貴¹、品川 政太郎¹、水上 雅博²、杉山 弘晃²、須藤 克仁¹、中村 哲¹

(1.奈良先端科学技術大学院大学、2.NTTコミュニケーション科学基礎研究所)

1.まとめと展望

ねらい

話者の分散表現を用いた話者らしい応答の生成

課題

従来のモデルでは話者情報として与えられた文に強く条件づけられた返答文を生成してしまう

提案

入力話者情報にVAEで獲得した分散表現を用いることでペルソナ記述文中の単語がそのまま出力されることを抑制

結果

ペルソナ記述文からの条件付けが**緩和**された
 生成文の多様性が**向上**した
 ペルソナ記述文と矛盾した生成文は**あまり見られなかった**

展望

よりよい話者情報の抽出方法や話者情報の反映方法の検討

2.背景

より魅力的な対話システムを作り出すために、話者性を持った多様な応答を生成できるようにしたい

課題

従来モデル | ペルソナ記述文をEncoderへ入力する^[1]**Pros.** 話者の情報を含んだ返答を生成する**Cons.** 話者情報として入力された単語をそのまま出力へ反映してしまう

仮説

モデルが入力に含まれる単語を
 出力に含めるように学習しているのではないか

提案

話者の分散表現という形で入力することで
 話者を表現する単語の単純な出力を抑制する

[1] Song Haoyu et al,BoB: BERT Over BERT for Training Persona-based Dialogue Models from Limited Personalized Data.arXiv preprint arXiv:2106.06169.

3.実験設定

問題設定

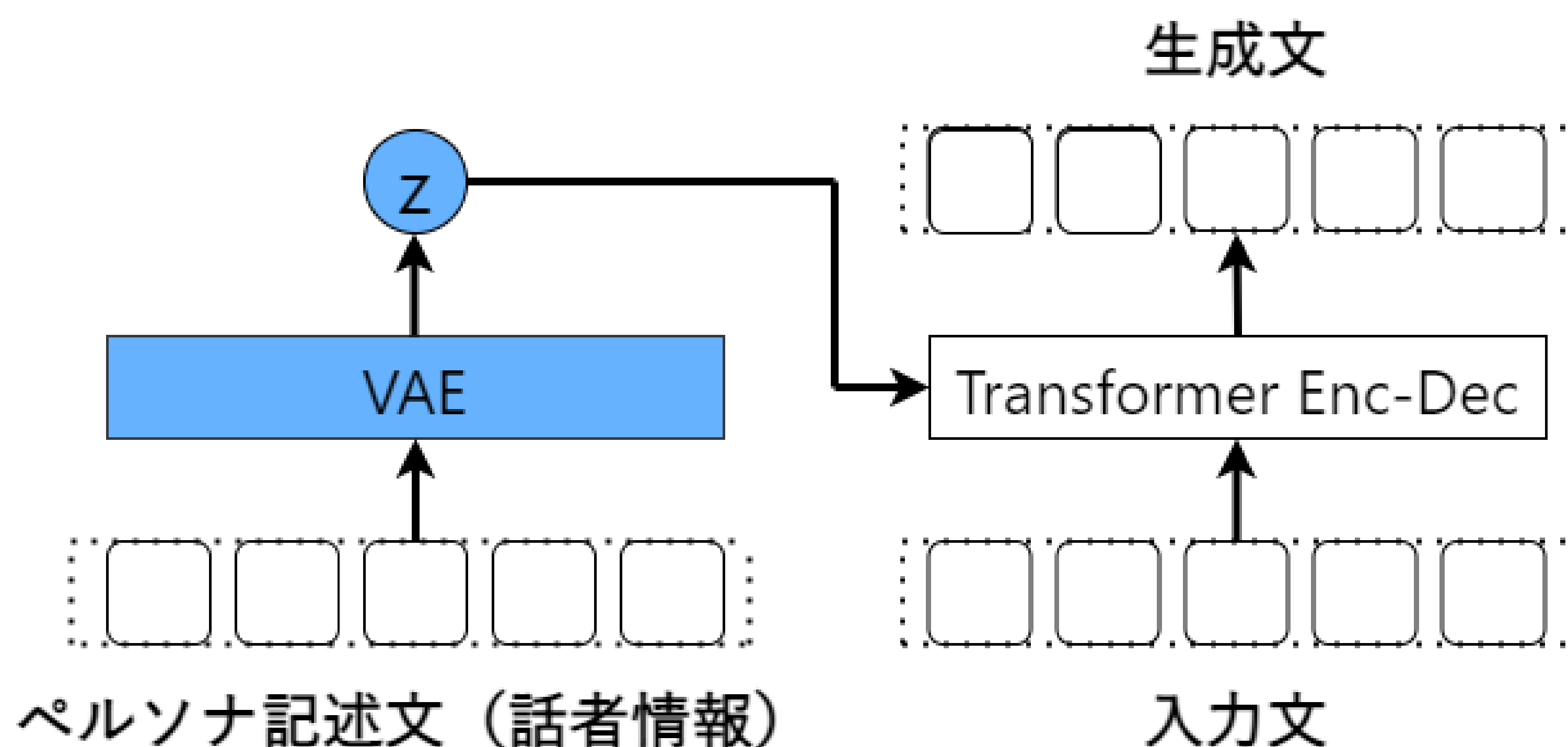
- VAEを用いて話者の情報を記述した文（ペルソナ記述文）から話者の情報を表現した分散表現を獲得する
 - 分散表現をDecode時の各トークンへ加算して事前学習済みTransformerモデルへ入力する
- ⇒話者の情報を反映しつつ、ペルソナ記述文に登場する単語からの条件付けを緩和し、多様な応答を生成できるモデルを目指す

実験設定

- データセット | JPersonaChat
- ペルソナ記述文（5文）と発話文、返答文の組で構成される
- モデル構造 | VAE + Transformer Enc-Dec
- ベースライン | Transformer Enc-Dec

学習手順

- ペルソナ記述文のみを用いてVAEの学習を行う
- 対話文とペルソナ記述文をVAEと事前学習済みTransformerを連結し、同時に学習させる



4.結果

ペルソナ記述文からの条件づけは緩和できたか？

ペルソナ記述文と生成文の一致度合いをBLEUで計測
 →ベースラインと比べて**一致度合いが低下**し従来モデルの過剰な条件付けの緩和ができた
 →テストデータよりも低くなり、条件づけが緩和され過ぎてしまった可能性がある

条件付けの緩和により多様性のある返答文を生成できたか？

生成文の多様性をDistinct-Nで計測
 →ベースラインと比べて**多様な応答を生成**できた

条件付けの緩和により話者情報に反するような返答は生成しなかったか？

ペルソナ記述文と生成文が矛盾しているかを人手評価で計測
 →提案モデルにおいて矛盾していると判定されたものはテストデータ全体の**4%**であった。しかしペルソナ記述文と無関係なものも見られた

考察

分散表現として与えることで条件づけの緩和を行い、多様な返答文を生成できるようになった。
 加えてペルソナ記述文と矛盾する返答も余り見受けられなかった。
 しかしJPersonaChatよりも被覆率が低くなり、無関係な生成文も見られるなど、条件づけの過度な緩和も見られた

モデルへの入力（左下）と出力（右）の例

ペルソナ記述文（一部）	私は香水の匂いが苦手です。	ベースラインモデル	モフモフいいですね。 私は香水の匂いが苦手なので、モフモフできないのが残念です。
対話入力文	もう、すっごいかわいいんですよ。 人懐っこいのでモフモフ出来ます！	提案手法	モフモフしたいです！ 私の実家の猫も人懐っこいですよ。

	BLEU3(↓)	BLEU4(↓)
テストデータの応答文	0.035	0.017
ベースラインの生成文	0.135	0.094
提案モデルの生成文	0.017	0.005

	Distinct-1(↑)	Distinct-2(↑)
ベースラインモデル	0.0290	0.1197
提案モデル	0.0336	0.1415