



中国科学院 信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING, CAS



DiffusEmp:

A Diffusion Model-Based Framework with Multi-Grained Control for Empathetic Response Generation

ASCII LAB
Guanqun Bi

◆ 共情 (Empathy)

- 定义：能设身处地体验他人处境，从而达到感受和理解他人情感的能力。
 - 情绪共情：能够对他人的情绪感受产生自发性的替代性情绪体验，以适当情绪回应他人情感状态的能力
 - 认知共情：了解他人观点或精神状态的能力

◆ 共情对话系统

- 定义：在开放域对话回复中融入共情能力，表达理解、感受对方的心情、经历和想法
- 任务形式：条件文本生成
 - 输入：对话历史、情景（可选）、说话人的情绪（可选）
 - 输出：融入共情能力的自然语言回复、情绪分布（可选）

◆ 普通的开放域对话

我知道了



◆ 融入情绪的回复

怎么批评你？
肯定是你又闯祸了！



◆ 共情回复

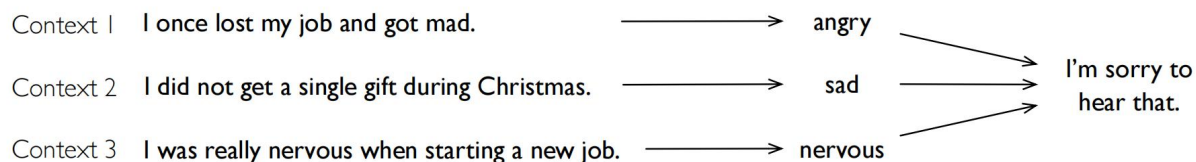
受吧，愿意跟我说说发生了什么吗？



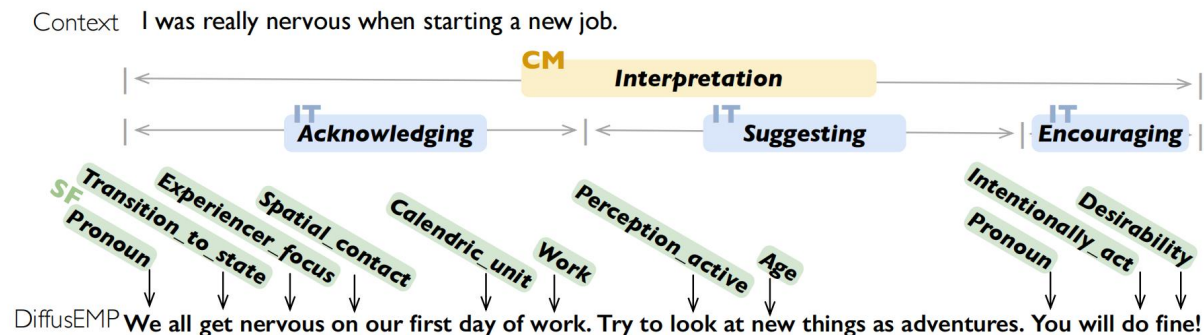
目前的共情对话可以生成一些语言流畅且包含感情的句子，但在话语的共情水平、语义丰富程度、可控性和一致性等方面还亟待提升。

◆ 问题：生成的回复单调乏味

- 仅考虑情绪，控制属性片面
- 仅考虑会话级别，控制粒度粗



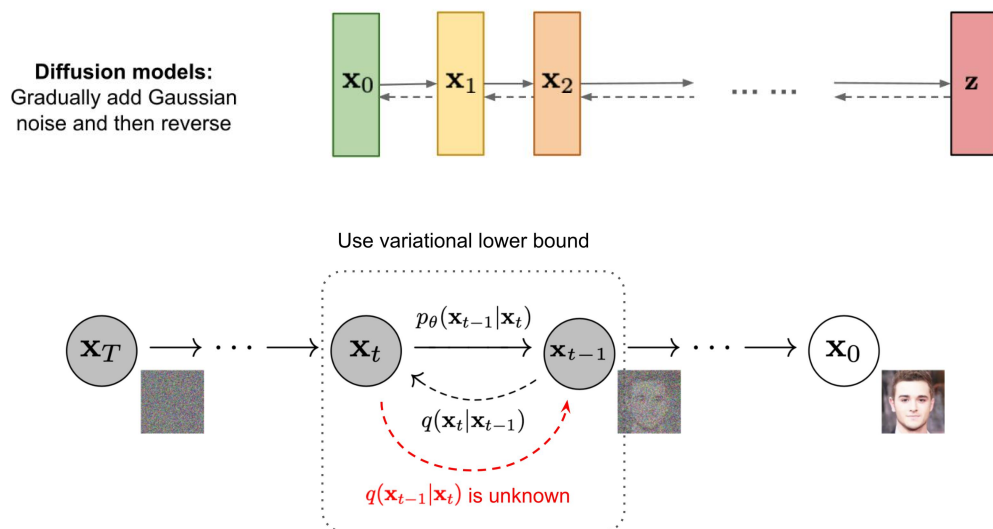
□ 如何优化传统的序列交叉熵生成方式，生成要素丰富且细粒度可控的共情回复？



Background

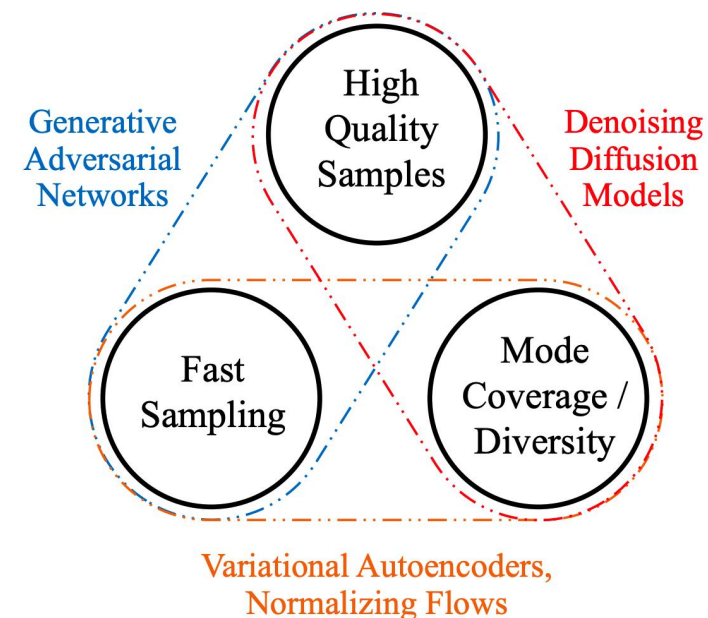
■ 扩散模型 (Diffusion Models)

- Inspired by non-equilibrium thermodynamics (非平衡态热力学)
- 定义了扩散步骤的马尔可夫链，以缓慢地将随机噪声添加到数据中，然后学习逆向扩散过程以从噪声中构造所需的数据样本。



◆ 作为生成模型

- 生成质量高
- 模式覆盖率/多样性高

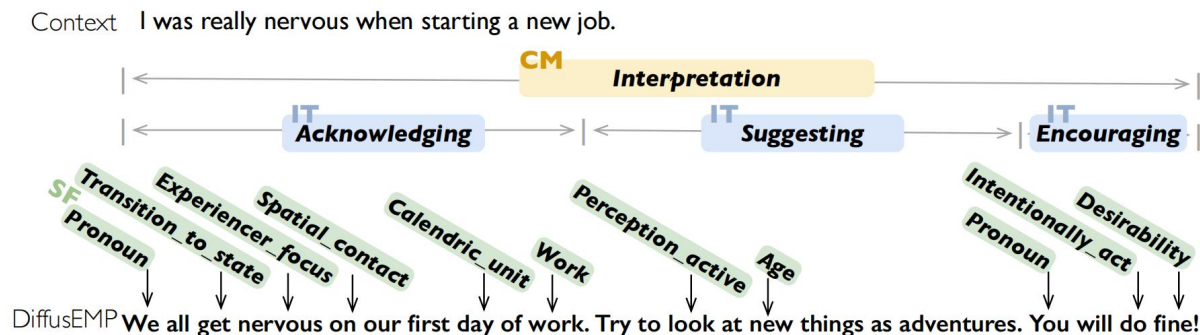


Control Signals

◆ 控制信号类型：

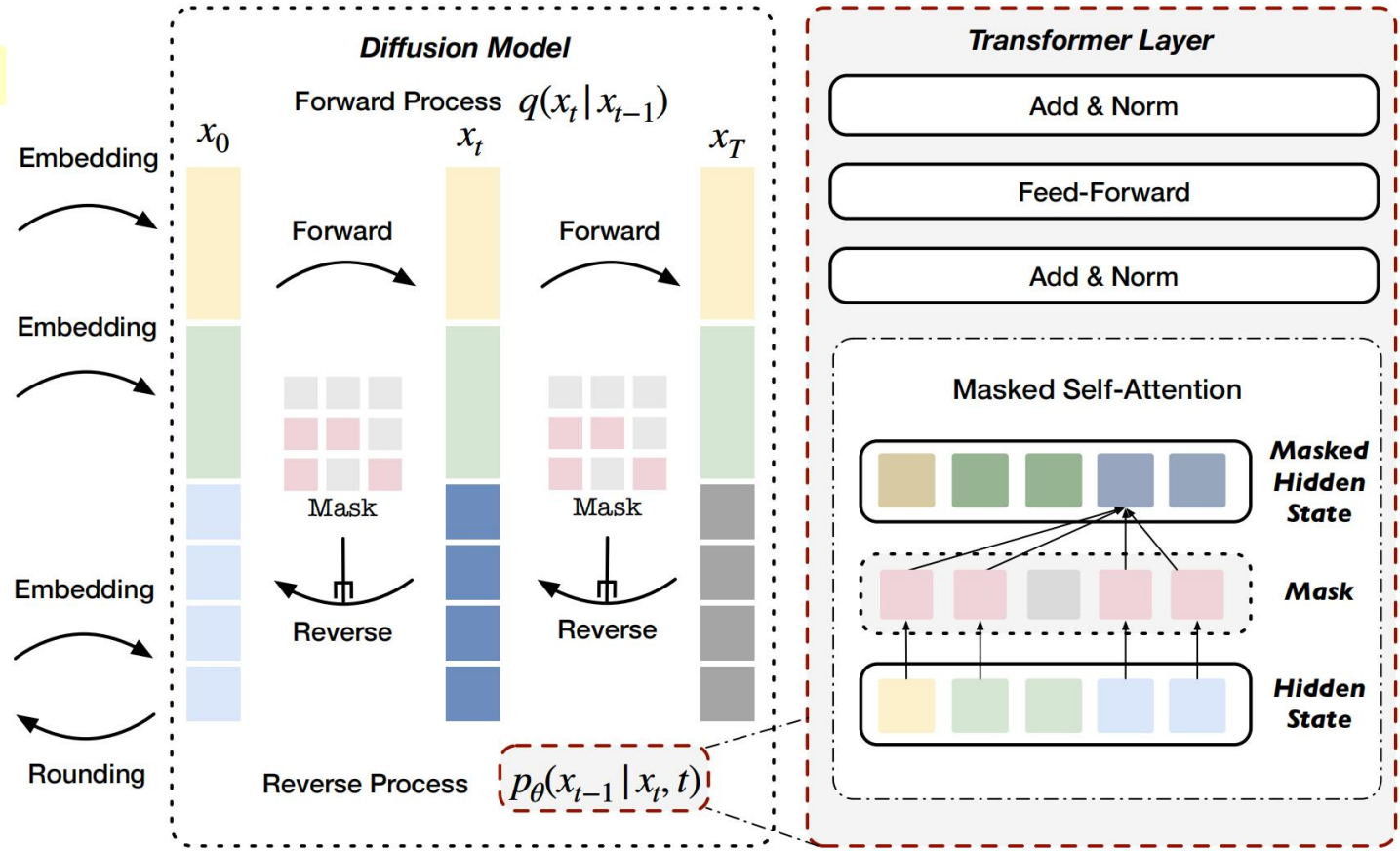
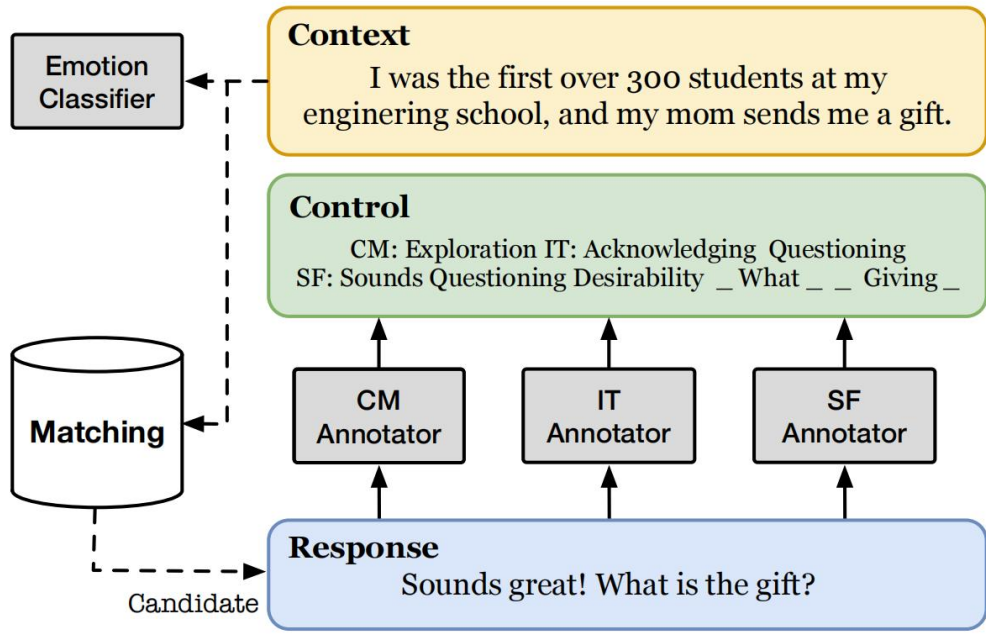
1. **Communication Mechanisms:** 表达共情的机制，conversation-level, 对response的标注
2. **Intention:** 话语的意图，sentence-level
3. **Semantic Frame:** 通用的词语类别信息，来自FrameNet

Signal Type	Accuracy	F1	#Classes
CM-ER	79.43	74.46	2
CM-IP	84.04	62.60	2
CM-EX	92.61	72.58	2
IT	87.75	87.71	9
SF	-	86.55	1222



Method

Notation: - - - - -> Inference - - - - -> Training



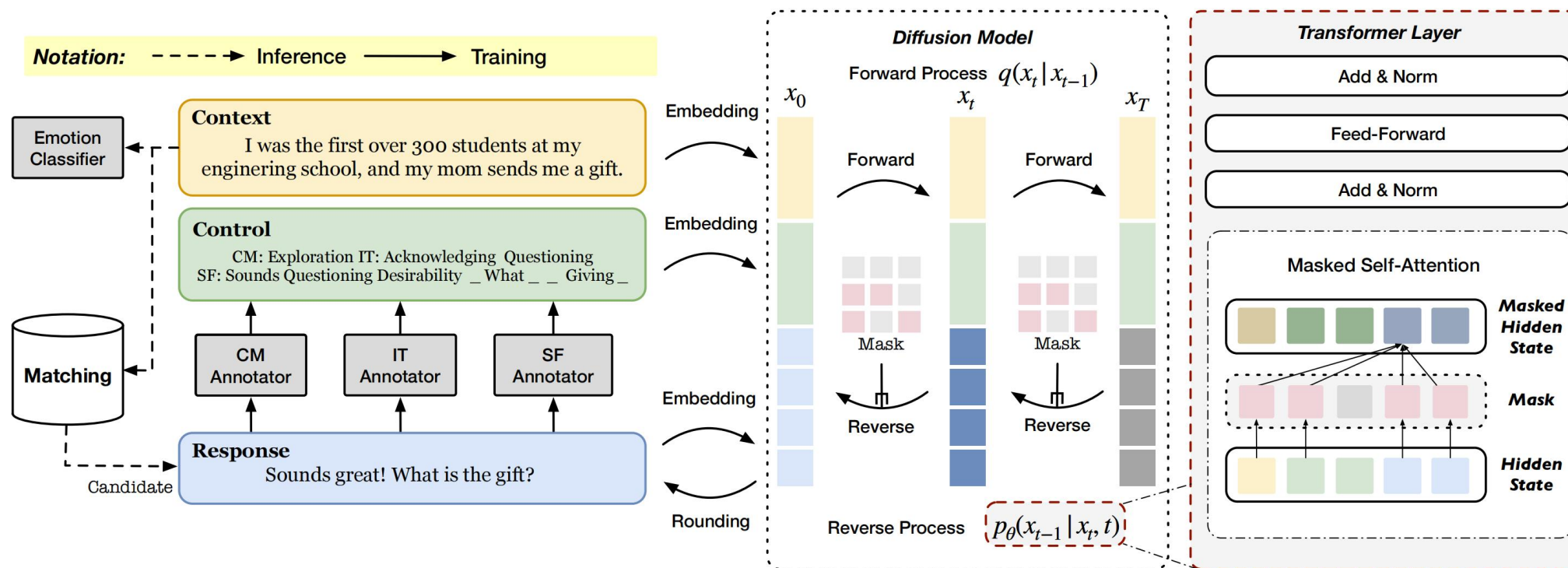
Method

◆ Forward Process

- 逐步添加噪声 $q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I})$

◆ Reverse Process

- 逐渐从噪声中恢复原始的 x_0



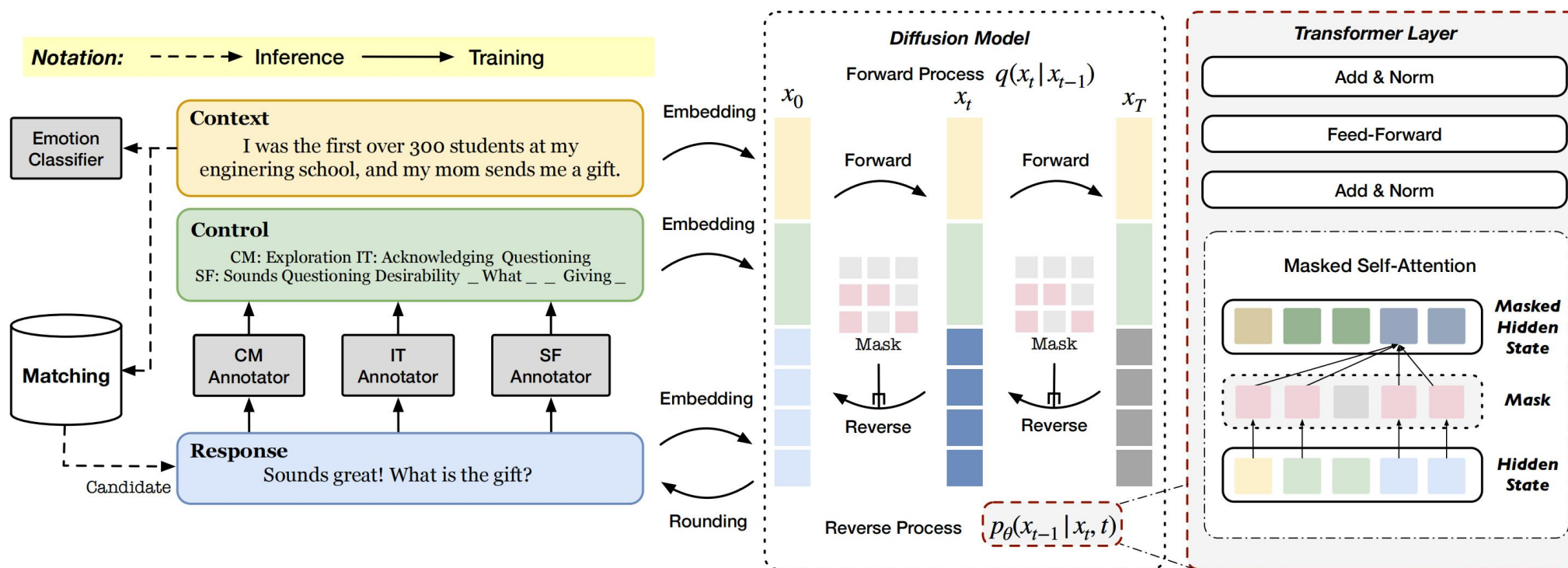
Method

◆ Training

$$\mathcal{L}_{\text{vlb}} = \sum_{t=2}^T \|\mathbf{y}_0 - \tilde{f}_\theta(\mathbf{x}_t, t)\|^2 + \|\text{EMB}(\mathbf{w}^y) - \tilde{f}_\theta(\mathbf{x}_1, 1)\|^2 + \mathcal{R}(\|\mathbf{x}_0\|^2),$$

◆ Inference

- golden label不可用，因此使用情绪增强的匹配方法 $Score = \text{SIM}_{\text{semantic}} + \gamma \text{SIM}_{\text{emotional}}$,



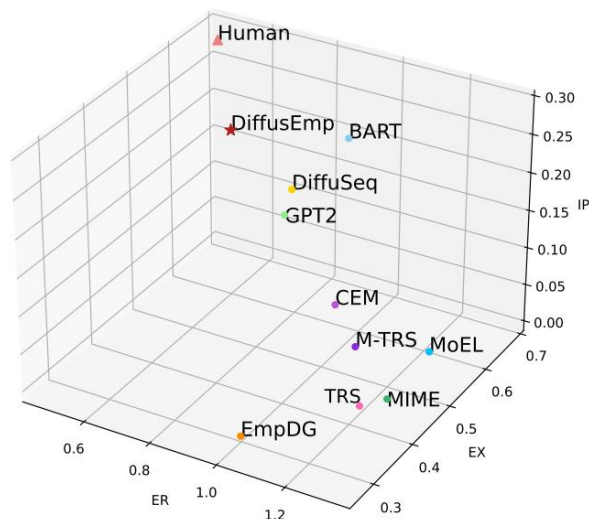
Experiment

Method	#Params	Relevance		Controllability			Informativeness				Length
		BERTScore \uparrow	MIScore \downarrow	ACC-CM \uparrow	ACC-IT \uparrow	F1-SF \uparrow	D1 \uparrow	D2 \uparrow	D4 \uparrow	sBL \downarrow	AvgLen \uparrow
<i>Transformer-Based Methods</i>											
TRS	15M	0.5717	4598.26	60.98	22.07	15.74	0.42	1.55	4.26	13.63	10.53
MTRS	15M	0.5735	7156.26	60.48	25.77	15.62	0.50	1.89	5.56	11.26	9.92
MoEL	21M	0.5758	14595.61	59.29	26.20	16.51	0.40	1.65	4.62	12.83	11.47
MIME	17M	0.5800	4878.71	61.16	22.00	16.54	0.26	0.87	2.15	14.21	11.12
EmpDG	29M	0.5745	9088.11	61.94	20.06	17.36	0.60	2.54	7.75	11.78	10.11
CEM	17M	0.5713	7635.05	62.28	30.09	14.20	0.54	2.00	4.98	9.13	8.25
<i>Pre-Trained Language Model-Based Methods</i>											
TransferTransfo	117M	0.5634	2138.39	59.70	25.08	18.39	2.81	17.22	36.54	2.68	11.40
BART	140M	0.5977	706.31	60.39	30.69	18.98	2.88	14.12	38.82	2.79	11.09
<i>Diffusion Model-Based Methods</i>											
DiffuSeq	91M	0.5101	715.95	59.23	28.58	17.26	1.79	26.97	88.17	1.29	10.30
DIFFUSEMP	91M	0.5205	626.92	92.36	84.24	52.79	2.84	29.25	73.45	1.09	14.12
<i>References</i>											
DIFFUSEMP (Oracle)	91M	0.7458	615.13	92.38	83.66	51.95	2.84	30.46	89.35	1.11	14.01
Human	-	1.0000	507.97	100.00	100.00	98.40	19.49	43.55	49.02	0.85	13.04

Analysis

		DIFFUSEMP	w/o SF
Relevance	BERTScore ↑	52.05	51.47
	MIScore ↓	626.92	993.44
Informativeness	Dist-1 ↑	2.84	1.69
	Dist-2 ↑	29.26	22.83
	self-BLEU ↓	1.09	1.31
Length	AvgLen ↑	14.13	13.23

Method	CM		IT		SF
	ACC ↑	F1 ↑	ACC ↑	F1 ↑	F1 ↑
DIFFUSEMP	92.36	90.26	84.24	77.15	52.79
w/o Mask	90.76	87.99	73.80	66.58	49.43
w/o CM	89.34	85.55	83.80	76.38	52.89
w/o IT	92.24	90.21	47.92	41.77	52.63
w/o SF	89.70	86.96	83.12	74.90	22.48



Context	I caught my boyfriend texting his ex.
Golden	Wow. Dump him and beat him up!
MTRS	Oh no! What happened?
MIME	Oh no, did he get hurt?
CEM	What did he do?
TransferTransfo	That is terrible! Was he able to get back to you?
BART	Oh no! Did you confront him about it?
DiffuSeq	Were you hurt?
Candidate A	Ok do ¹ not ² feel ³ bad ⁴ be happy ⁵ and search ⁶ for bad future ⁷ behalf
Control A	EMOTIONAL_REACTION SUGGESTING _ INTENTIONALLY_ACT ¹ NO ² PERCEPTION_EXPERIENCE ³ DESIRABILITY ⁴ _ EMOTION_DIRECTED ⁵ _ SCRUTINY ⁶ _ _ ALTERNATIVES ⁷ _
Response A	Just do ¹ not ² feel ³ bad ⁴ , happy ⁵ to study ⁶ in your future ⁷ .
Candidate B	That could ¹ be embarrassing, do ² you ³ have ⁴ a new ⁵ partner ? ⁶
Control B	EXPLORATION QUESTIONING _ POSSIBILITY ¹ _ _ INTENTIONALLY_ACT ² PRONOUN ³ POSSESSION ⁴ _ AGE ⁵ _ ? ⁶
Response B	That could ¹ be disgusting, do ² you ³ have ⁴ a new ⁵ relationship ? ⁶

投稿经验

- ◆ 杜绝低级错误
- ◆ “如果不生产垃圾，就连垃圾都生产不出来”
- ◆ 多沟通，反复改

- ◆ 敢想
- ◆ 相信并捍卫自己的idea

- 1_introduction_v1.tex
- 1_introduction_v2.tex
- 1_introduction_v3.tex
- 2_related_work_v1.tex
- 2_related_work_v2.tex
- 3_method_v1.tex
- 3_method_v2.tex
- 3_method_v3.tex**
- 4_experiments_v1.tex
- 4_experiments_v2.tex
- 5_discussion_v1.tex
- 5_discussion_v2.tex
- 5_discussion_v3.tex
- 6_conclusion_v1.tex
- 6_conclusion_v2.tex
- 7_limitations_v1.tex
- 7_limitations_v2.tex