

会話ログ分析による ユーザーごとの関心に応じた話題提案手法

武蔵野大学工学部数理工学科佐々木研究室4年 白川桃子
佐々木多希子(武蔵野大学)

- 概要** 【目的】 初対面の人と会話する際に**最適な話題を提案**したい！
- 【流れ】 **日常会話からあなたの興味を抽出**し初対面の会話相手の興味とマッチングする！
- 【結果】 最適な話題を提供することで、**活発に発言し感情豊かな会話が実現**した！

序論・目的

半数以上の人が初対面の人との会話に苦手意識があり
話題のなさが圧倒的に原因となっている。

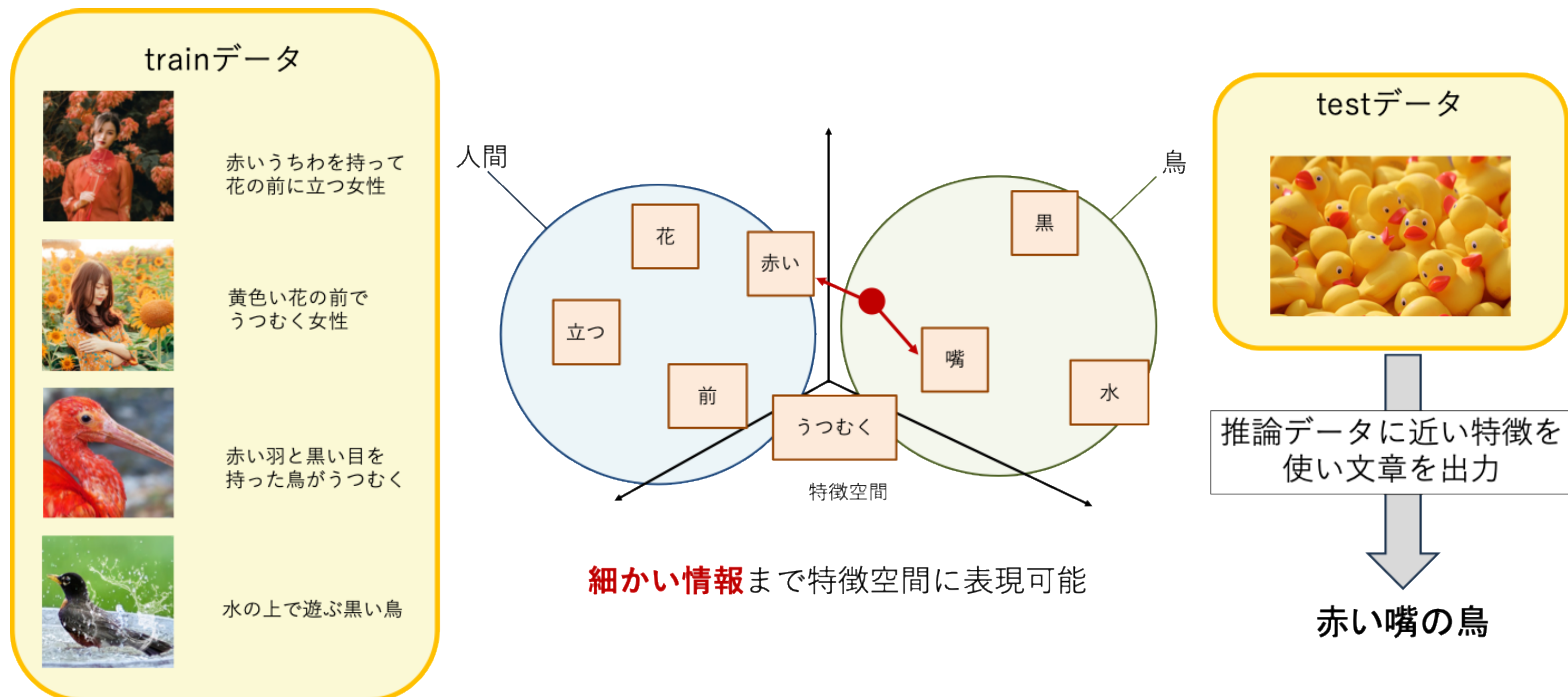
↓ なので本研究では...

- ① ユーザーの**関心に基づいた話題提供**の実現
- ② ユーザーの**興味関心を定量的に**評価する
- ③ 話題を提供することで初対面時の**会話を盛り上げる**

分析手法

雑談から興味のあることを抽出
会話データが129個しかない

→**trainとtestが不一致でも分類可能な手法を活用**



問題の定式化

Y^S : 学習用の可視クラス
(本研究ではtrainデータの日本語SNLIデータ, 548,014ペア)

Y^u : 予測用の不可視クラス
(本研究ではtestデータの名大会話コーパス, noteのカテゴリ)

$$Y^S \cap Y^u = \emptyset$$
$$Y^S: D^S = \{(x_i^S, y_i^S)\}_{i=1}^N$$
$$Y^u: D^u = \{(x_i^u, y_i^u)\}_{i=1}^M$$

x_i : i 番目のテキスト
 y_i : x_i に対応するラベル
※不可視クラスにおいては, 最もマッチしていると仮定できるラベル

$$\hat{y} = \operatorname{argmax}_{y \in Y} (x, y; \theta) \quad \theta: f \text{ のパラメータ}$$

**Y^S から \hat{y} が最大になるようなマッチングモデルを学習
そのモデルから Y^u の予測を行う**

テキストとラベルから学習

(x_i^S, y_i^S) の例: (いちごを食べます, いちごは食品だ)

↓

[CLS] いちごを食べます [SEP] いちごは食品だ [SEP]

**「いちごを食べます」という文に対して
「いちごは食品だ」は同じような意味を持つか?**

↑ こんな感じの操作を (x_i^S, y_i^S) を変えながら繰り返していく

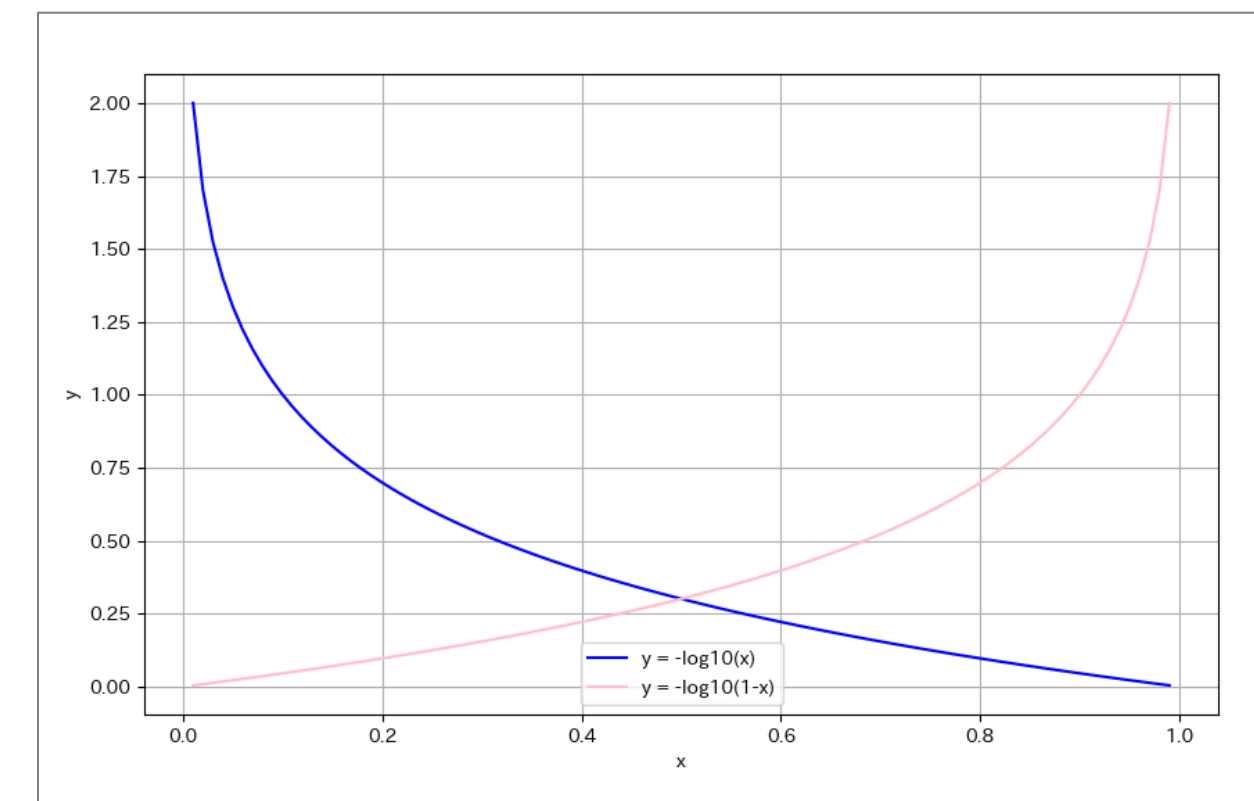
マッチングモデルの構築

線形層(1)を重ね, 損失(2)を計算

$$p_{x,y'} = \sigma(W^T c_{x,y'} + b) \quad (1)$$

$$\mathcal{L} = \begin{cases} -\log(p_{x,y'}) & y' = y \\ -\log(1 - p_{x,y'}) & y' \neq y \end{cases} \quad (2)$$

W, b : 線形層のパラメーター($W \in R^H, b \in R$)
 $c_{x,y'}$: 各文章に対応する隠れベクトル(H 次元)
 $p_{x,y'}$: x と y' のマッチングスコア
 $\sigma(\cdot)$: シグモイド関数



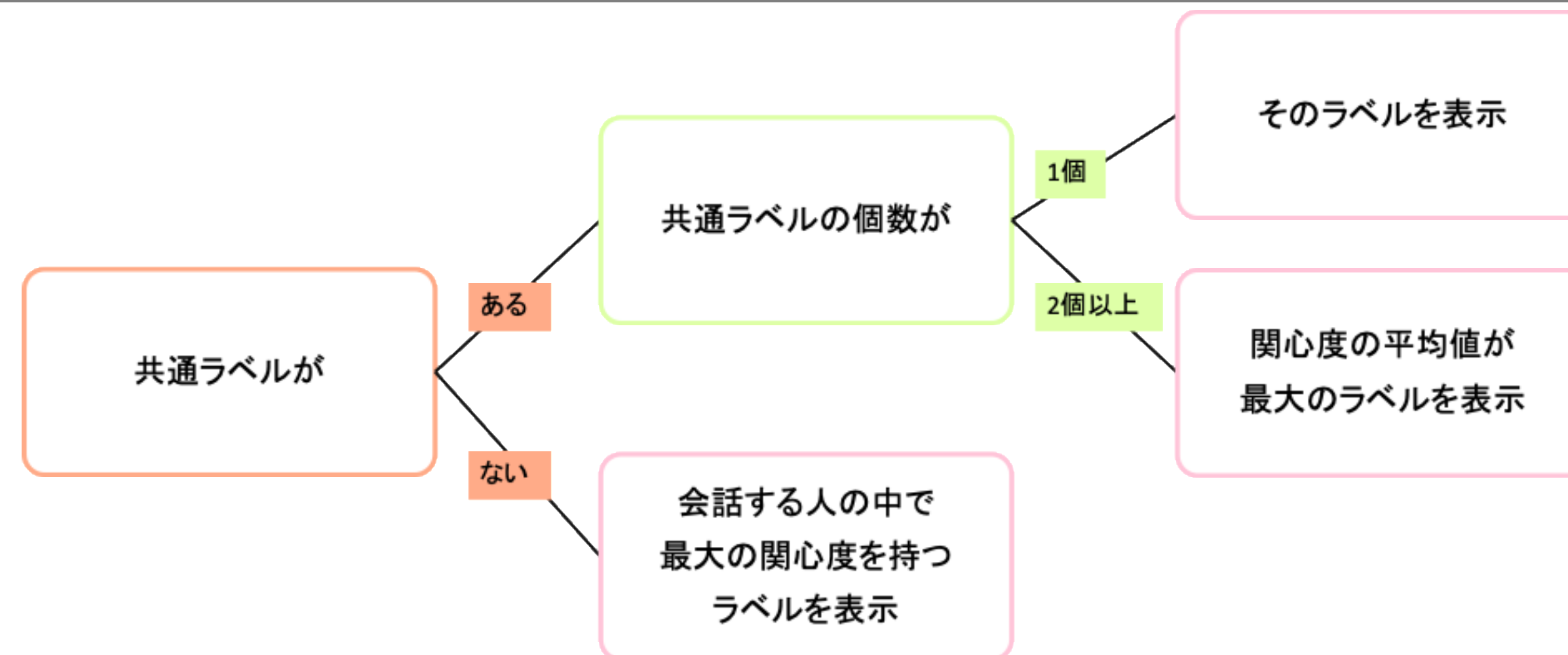
\mathcal{L} が小さくなるようにモデルを構築していく

強化学習/報酬

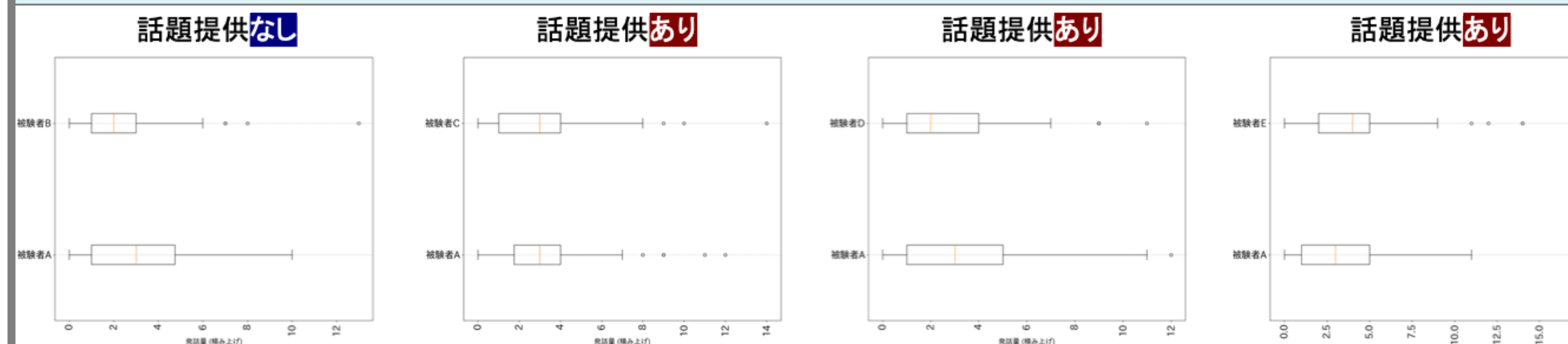
各ステップにてポリシーネットワークの出力に基づきエージェントは
現在のインスタンスを選択するか判断

$\{p_{x,y'} | y' \in Y^u\}$
**会話中の各文章と「この文章は{noteのカテゴリ}に関する会話だ」
は似た意味を持つか?**
↑ テキスト x とのマッチングスコアが最大となるようなラベル y^* を擬似ラベルとする。

話題提供フロー

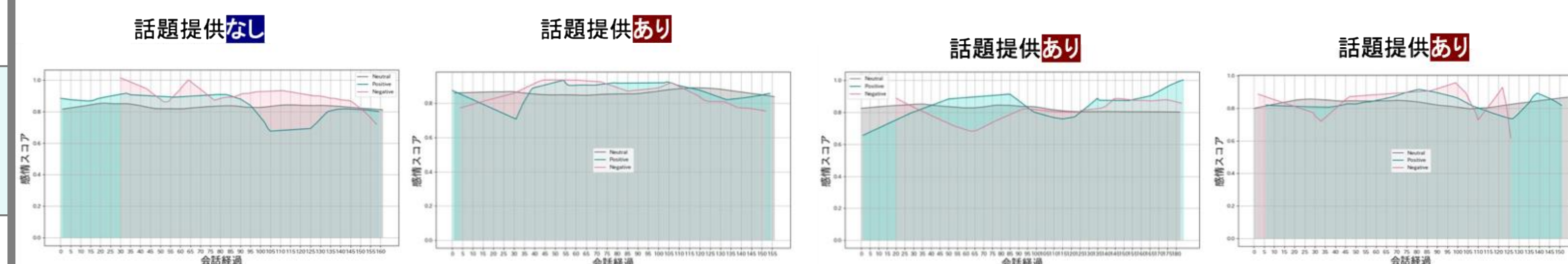


評価 - 初対面会話時の発話量



話題を提供するか否かは**発話量の傾向にも大差がない**
話題提供することでより生き生きと発言できる可能性が示唆

評価 - 会話中の感情推移



話題提供がある方が感情の振れ幅が大きい
感情の起伏が大きく初対面でもインパクトのある会話を実現

参考文献

- ・小町守(2024) “自然言語処理の教科書” 株式会社技術評論社
- ・藤村逸子, 大曾美恵, 大島ディヴィッド義和(2011) “会話コーパスの構築によるコミュニケーション研究” 藤村逸子, 滝沢直宏編『言語研究の技法: データの収集と分析』p.43-72, ひつじ書房
- ・Zhiquan Ye, Yuxia Geng, Jiaoyan Chen, Xiaoxiao Xu, Suhang Zheng, Feng Wang, Jingmin Chen, Jun Zhang, Huajun Chen “Zero-shot Text Classification via Reinforced Self-training” Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 3014–3024 July 5 - 10, 2020. c2020 Association for Computational Linguistics