

A Semi-supervised Framework Dedicated to Customer Service

Gaël Guibon^{1,2}, Luce Lefeuvre², Matthieu Labeau¹, Chloé Clavel¹

- (1) LTCl, Télécom-Paris, Institut Polytechnique de Paris
- (2) Team Plateau IA - Direction Innovation & Recherche - SNCF



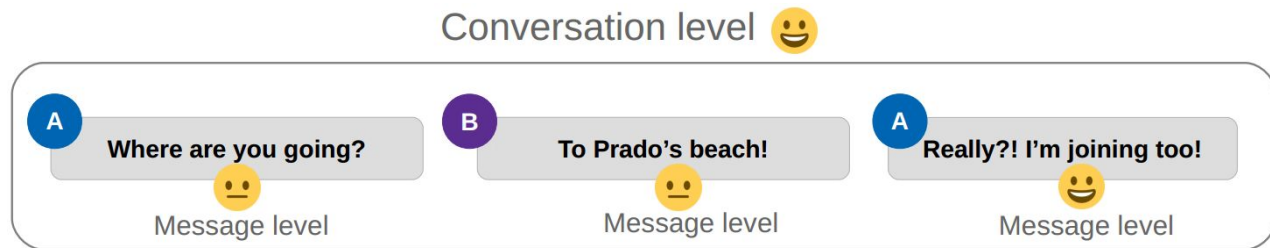
DSAIDIS Chair, 15th june 2022



Data Representation: two-level annotation

Conversation-level labels

Message-level labels



Private Data

Conversation

Label messages Label the conversation

Customer hi
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

Operator Good evening and welcome the customer service. How can I help you?
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

Operator Are you still there?
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

Customer yes
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

Customer i can't book my ticket. doesn't work!
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness Disappointment
Anger Frustration

Customer can you do smth about it?
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

Operator Please give me more details on your
emotion: Surprise Amusement Satisfaction Relief Neutral Fear Sadness
Disappointment Anger Frustration

CONTINUE

Conversation

Label messages Label the conversation

Pick an overall problem status for this chat
Solved

Pick the customer satisfaction
 -3 -2 -1 0 1 2 3




Emergency context?
 yes

FINISH BACK

<https://gguibon.github.io/ezcat/#/>



to annotate customer service conversations

- Customer service conversations 
- Real data from Oui SNCF 
- Two annotators 

Emotions (message-level)

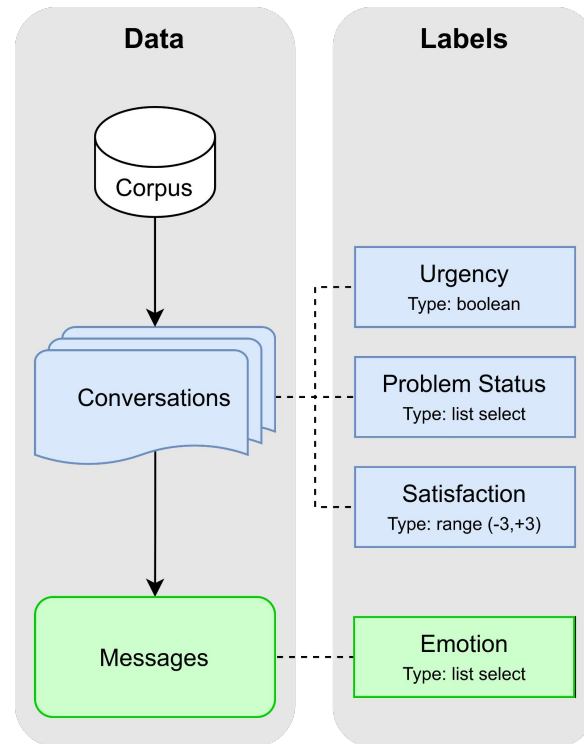
Amusement Satisfaction Relief Fear Sadness Disappointment Frustration Anger Surprise

Customer's Satisfaction (Conversation-level)

-3	Very Unsatisfied
-2	Unsatisfied
-1	Mildly Unsatisfied
0	None
+1	Mildly Satisfied
+2	Partly Satisfied
+3	Fully Satisfied

Problem Resolution Status (Conversation-level)

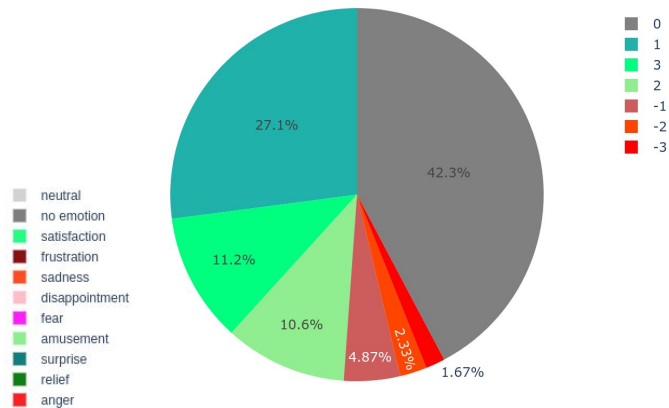
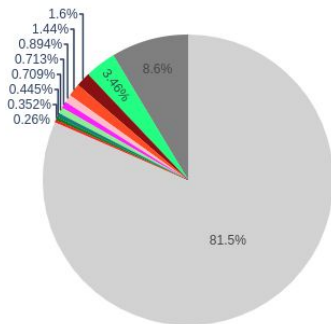
Solved | To be tested | Out of scope | No solution | Aborted





to annotate customer service conversations

- Customer service conversations
- Real data from Oui SNCF
- Two annotators



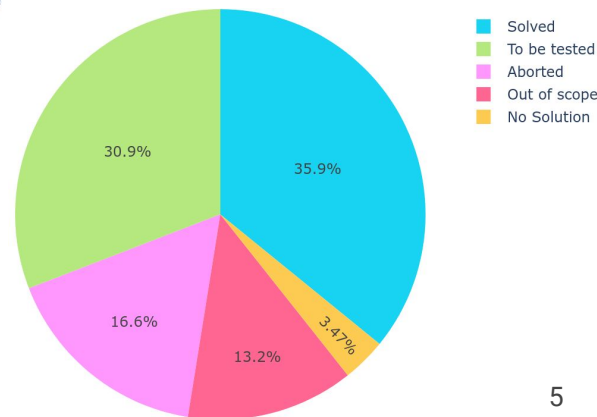
Results

Written Language ⇒ French

Conversations ⇒ **1,500**

Avg messages per conversation ⇒ **13**

Inter-annotator agreement ⇒ **0.65** Kappa score



Related Work

Problem Resolution

- Used the gold values to check the **correlation with sentiment score** (Jain, 2021)
- Some dataset possess a **related information** but do not use it (Chen, 2021)

No work on status of the problem resolution

Customer Satisfaction

- In Samsung Chats using VADER and LSTM (Hutto and Gilbert, 2014; Park et al., 2018)
- Based on tweets (Liu et al., 2021b)
- Using Net Promoter Score (Auguste et al., 2019; Piris and Gay, 2021)
- Satisfaction scale is often 5 (Hutto and Gilbert, 2014; Park et al., 2015) ★★★★★4.7

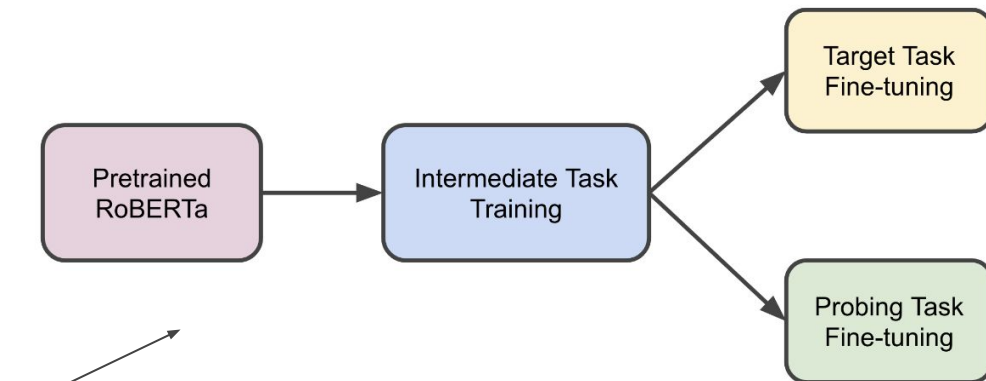
input/macroF1	SVM	CNN	RNN+att
<i>all turns</i>	52.7	53.5	52.7
<i>client turns</i>	51.1	52.3	52.5
<i>agent turns</i>	47.2	48.2	47.6

Table 2. Models comparison with macroF1 according to the input stream: all turns, only client turns, only agent turns

Related Work

Semi-supervised Learning

- Task-Adaptative Pretraining (Howard and Ruder, 2018)
- Domain-Adaptative Pretraining (Konlea and Jannidis, 2020; Wu et al., 2021)
- Intermediate fine-tuning (Gururangan et al., 2020)
 - with complex tasks (Pruksachatkun et al., 2020)
 - with dummy tasks (Chang and Lu, 2021)



Main Objectives

Detect problematic conversations

Predict the status of the customer's problem

Problem Status detection:

Aborted / **Solved** / **To test** /

Out-of-scope / **No solution**

Obstacles:

- More complex than topic modelling
- Problem compositionality: two or three fold

Predict the customer's satisfaction

Using the conversation context

With fine-grained satisfactions

Customer's Satisfaction du

Customer: -3 -2 -1 0 1 2 3

Obstacles:

- Not always linked to problem's resolution
- Fine-grained...

Only a few available data!

Supervised: DistilCamemBERT Fine-tuning

Direct Fine-tuning

Target Tasks



Speaker Identification:
Operator / **Customer**



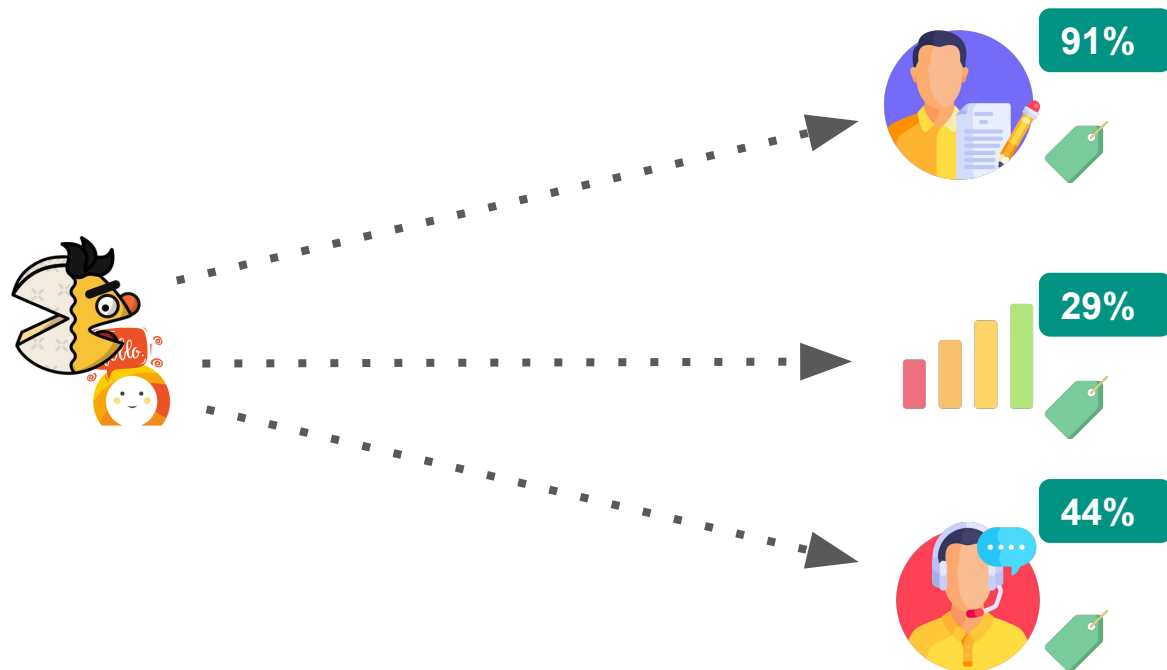
Customer's Satisfaction du
Customer: -3 -2 -1 0 1 2 3



Problem Status detection:
Aborted / **Solved** / **To test** /
Out-of-scope / **No solution**

Supervised: DistilCamemBERT Fine-tuning

Direct Fine-tuning

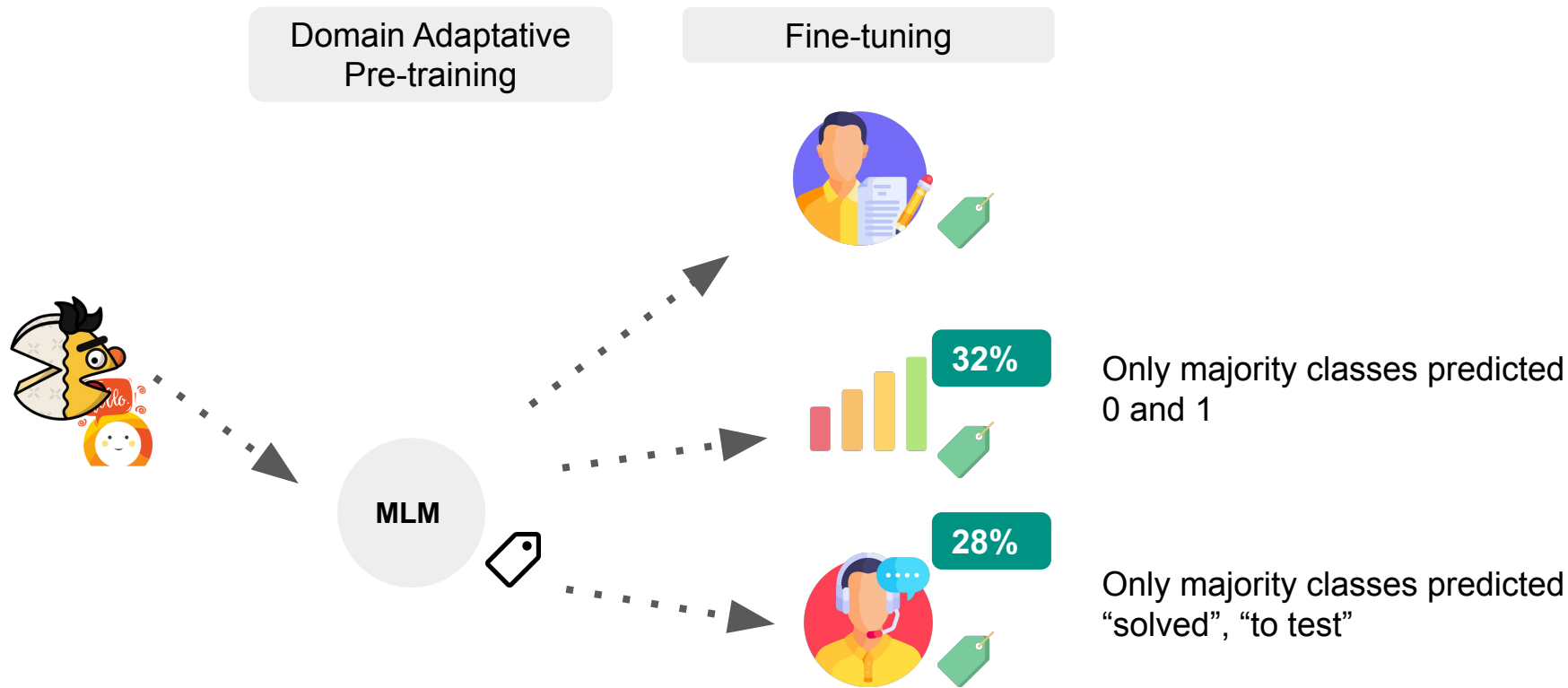


Mainly predicts the majority classes: 0 and 1

“Out-of-scope” and “No solution” are not predicted

Domain Adaptative Pre-training (inductive SSL)

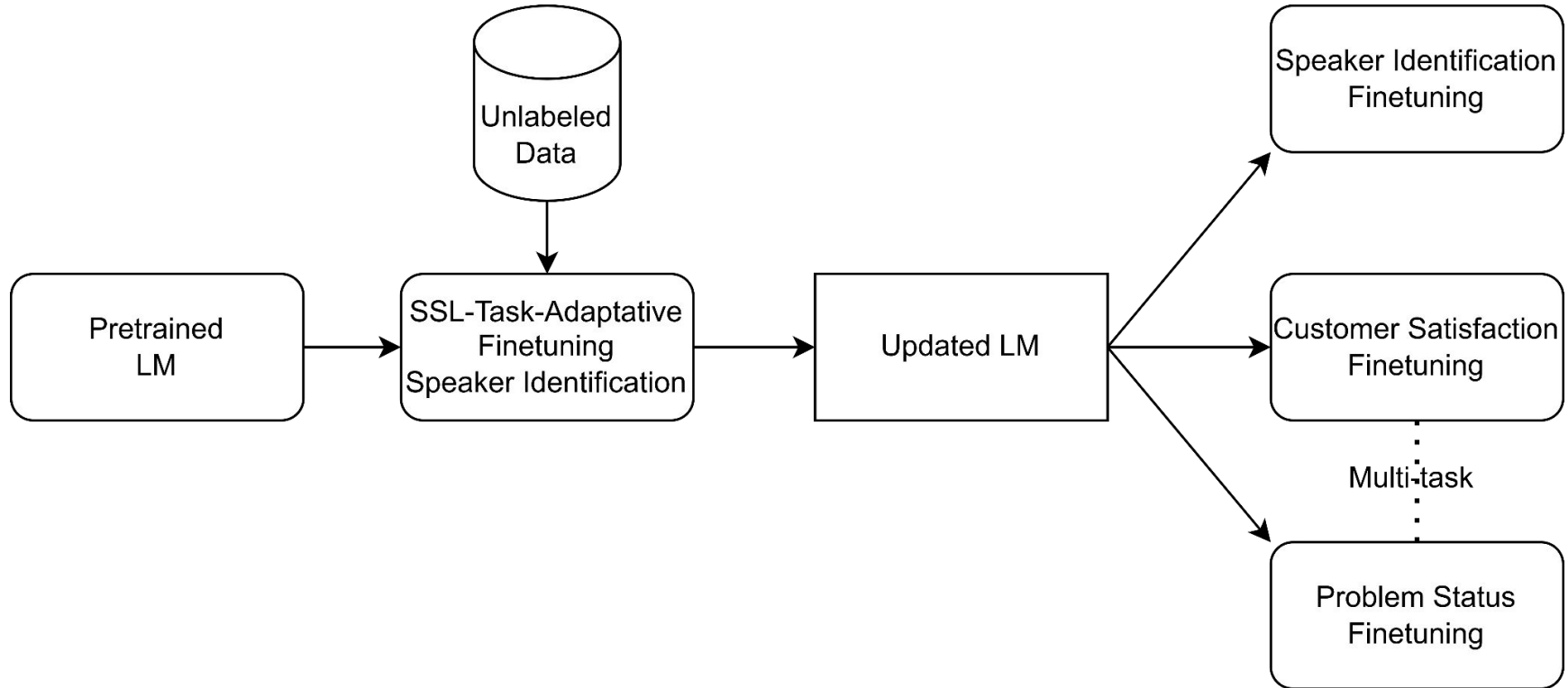
(Konlea and Jannidisa, 2020; Wu et al., 2021)



Only majority classes predicted 0 and 1

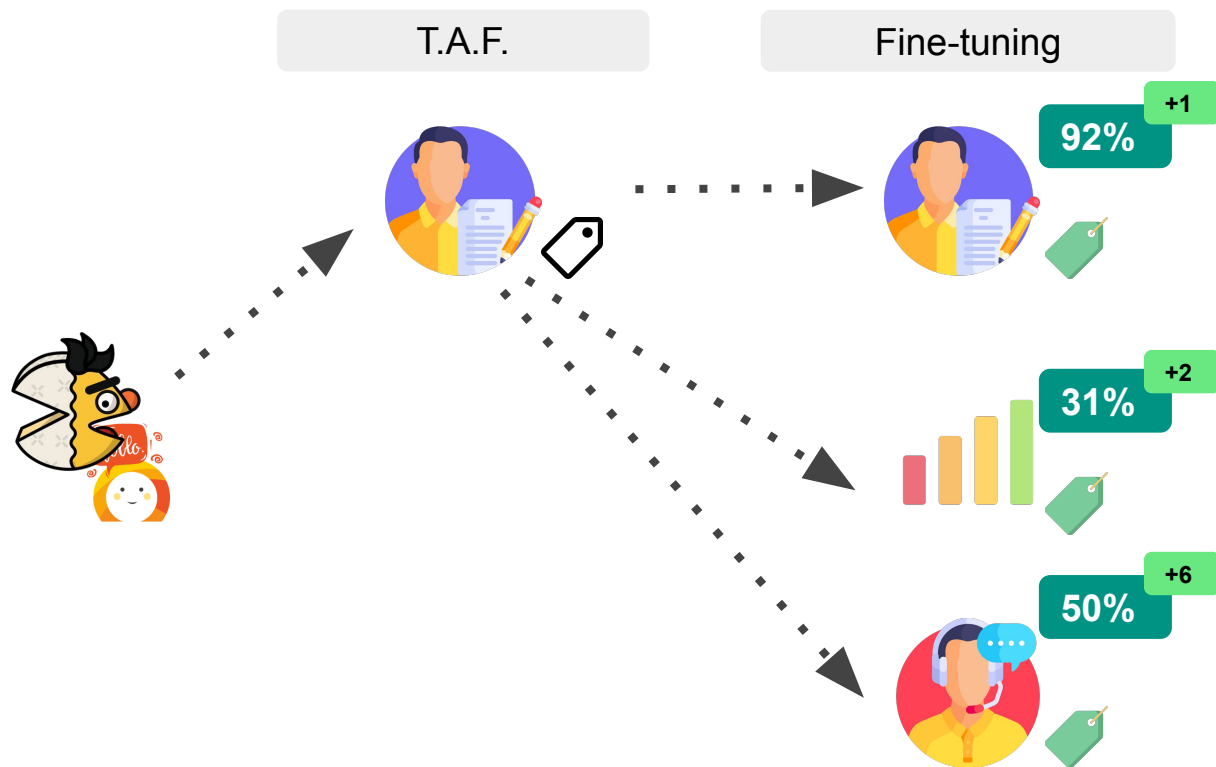
Only majority classes predicted "solved", "to test"

Inductive Semi-Supervised Learning



Task Adaptive Fine-tuning

Leveraging Speaker Information in Customer Services

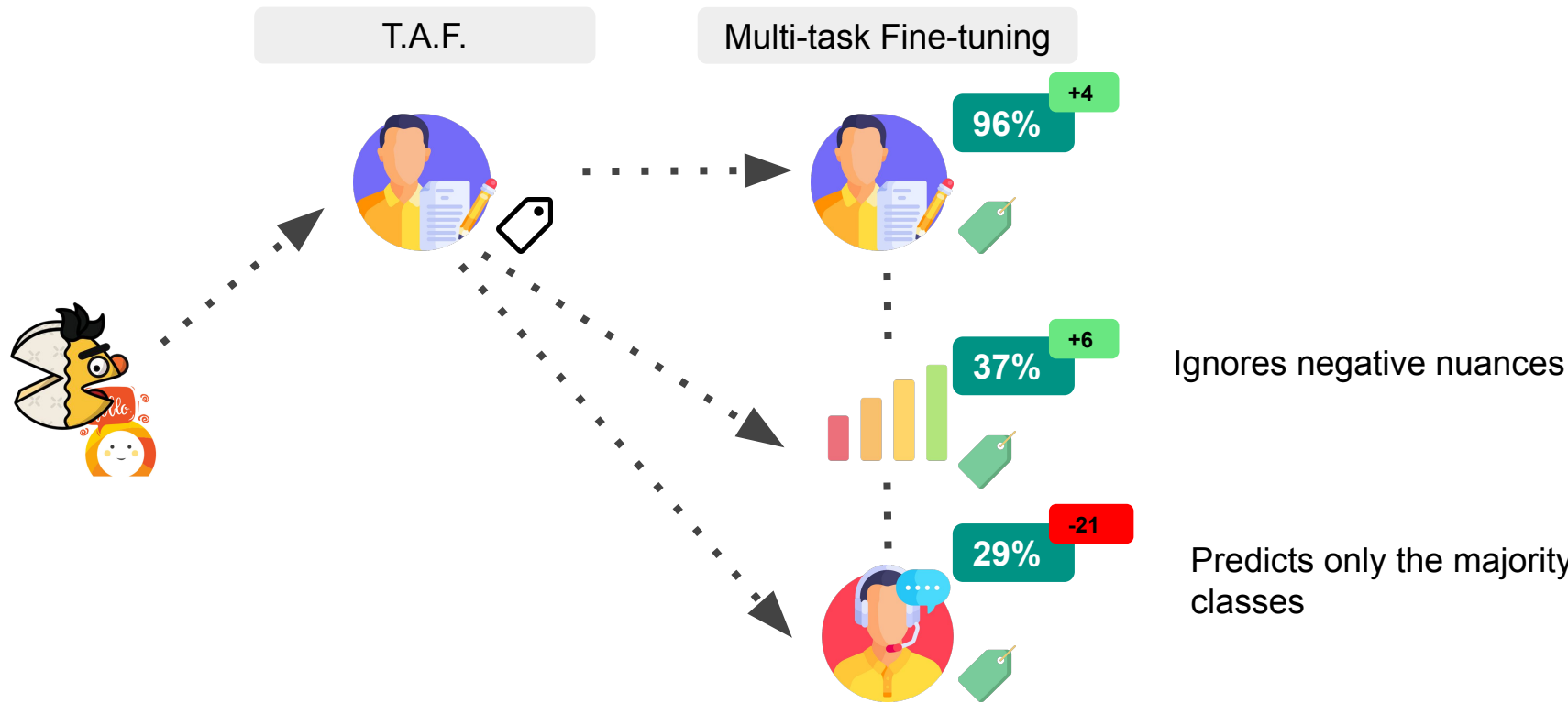


Negative nuances have issues

“No Solution” is never predicted

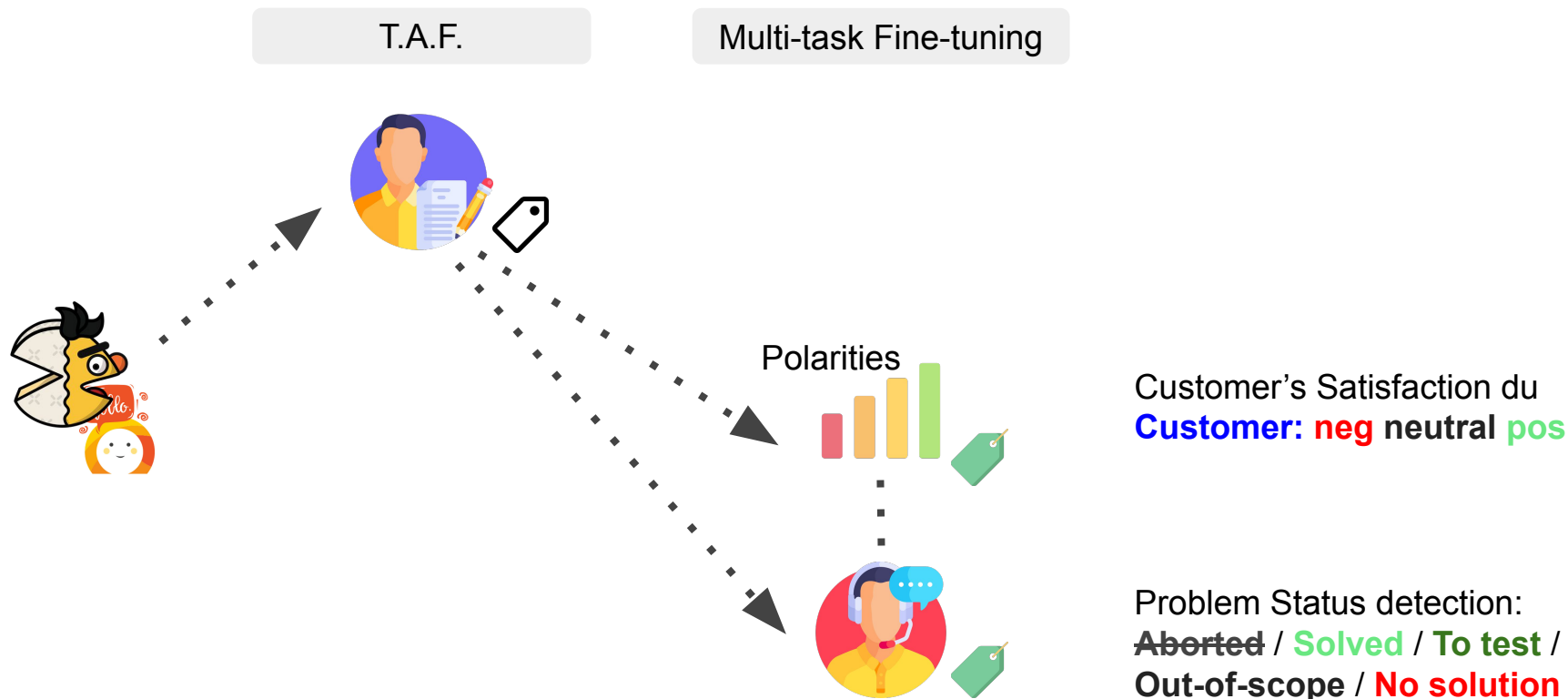
Task Adaptative Fine-tuning with Multi-task Fine-tuning

Leveraging Speaker Information in Customer Services



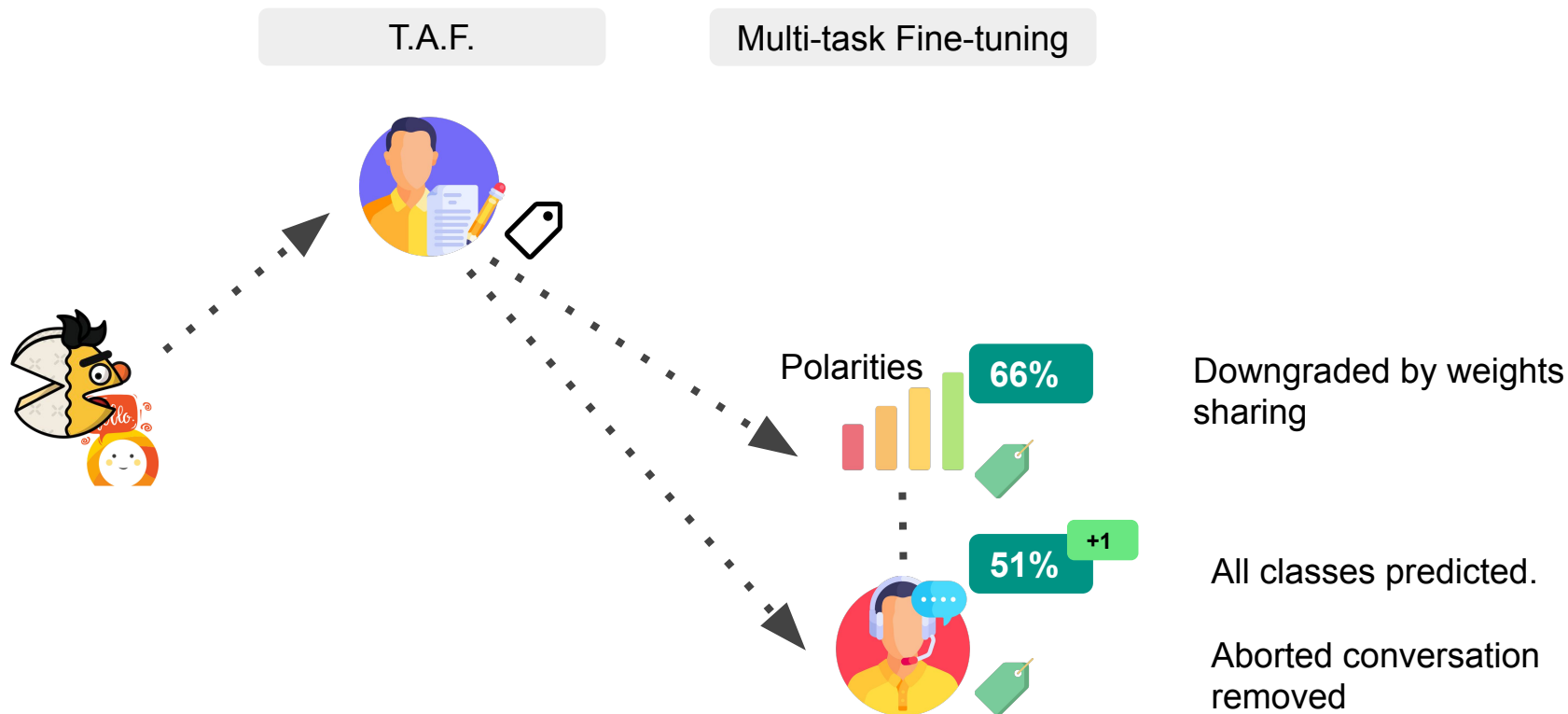
Task Adaptative Fine-tuning with Multi-task Fine-tuning

Leveraging Speaker Information in Customer Services



Task Adaptative Fine-tuning with Multi-task Fine-tuning

Leveraging Speaker Information in Customer Services



Conclusion

- SSL approach derived from Domain Adaptative Pre-training (Konlea and Jannidisa, 2020; Wu et al., 2021) and Intermediate Fine-tuning (Gururangan et al., 2020)
- We show simple intermediate task is sufficient for customer service
- First work on problem status prediction
- Can serve other use cases

But

- Difficult problem status prediction \Rightarrow due to labels' nature
- Still insufficient data