итМО

Структурирование, разметка и обогащение данных. Распознавание эмоций в диалогах

Выполнил:
Смирнов Е. А., гр. P4122

Преподаватель:
Муромцев Д. И.

Предметная область

ИТМО

NLP и мультимодальный анализ речи

Приложения:

- Диалоговые системы и чат-боты
- Психотерапия
- Корпоративные коммуникации
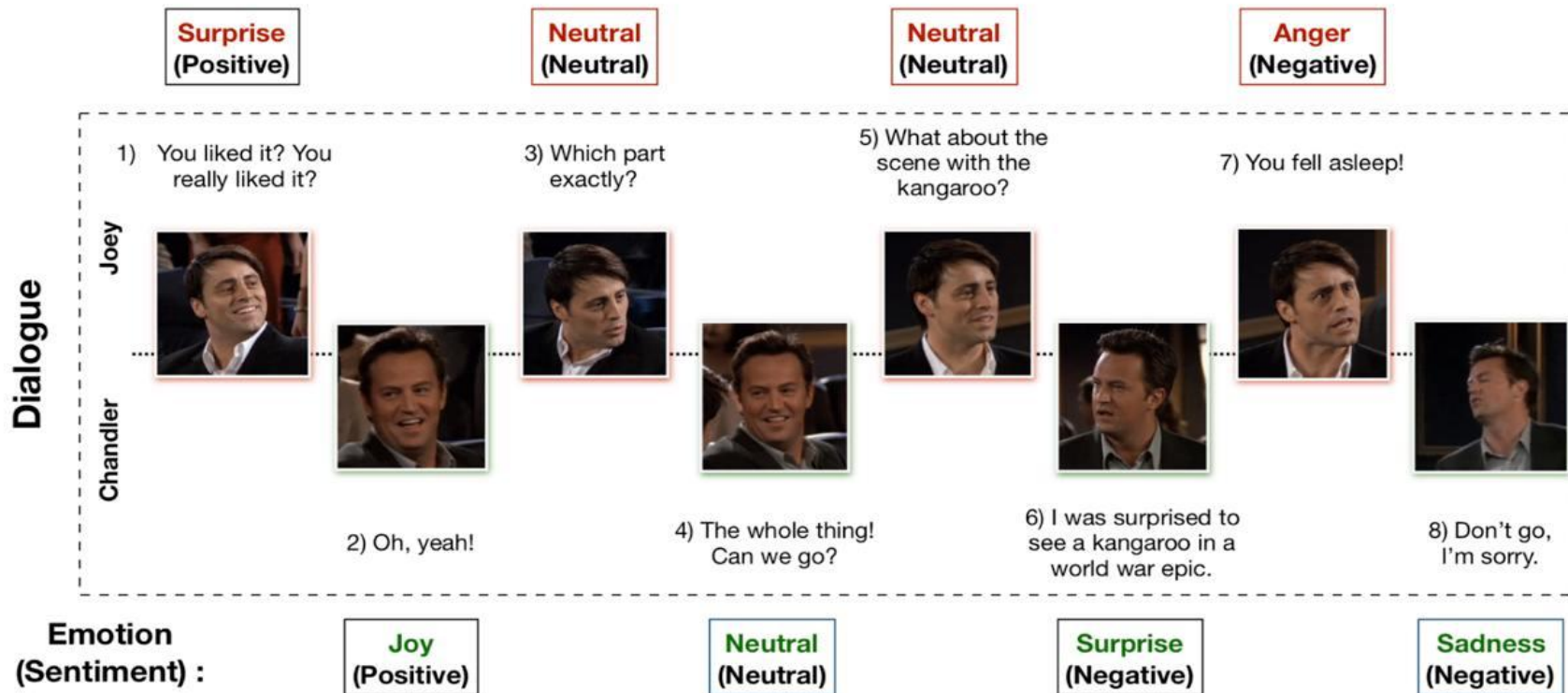- Социальные сети

Постановка задачи

Цель: анализ динамики эмоций в диалоге (текст + аудио)

Назначение:

- Распознавание эмоций в диалоге
- Учет изменения эмоций во времени
- Адаптация под эмоциональное состояние

Пример

Example Dialogue

Статистика MELD

ИТМО

Показатели

Уникальных слов: до 10,643

Средняя длина реплики: ~8 слов

Средняя длительность: ~3.59 сек

Разбиение дата сета

Train: 1039 диалогов, 9989 реплик

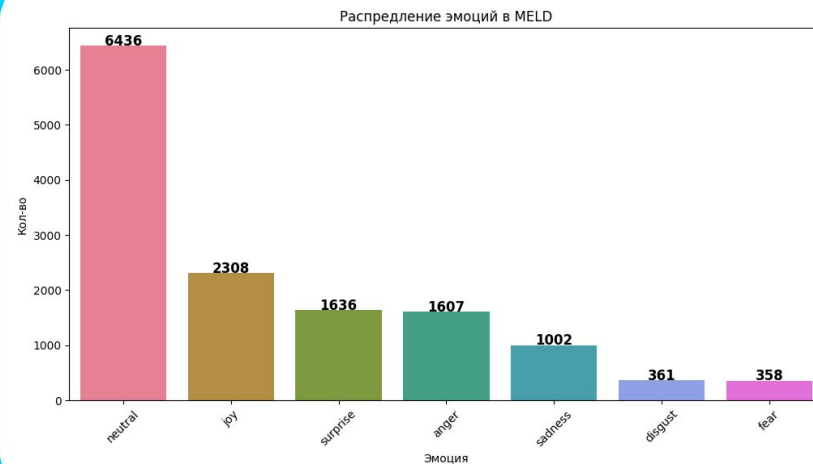Dev: 114 диалогов, 1109 реплик

Test: 280 диалогов, 2610 реплик

Статистика MELD

ИТМО

Эмоции

7 эмоций:

neutral, joy, surprise, anger, sadness, disgust, fear

Распредление эмоций в MELD

Источник данных

ИТМО

Sr No.	Utterance	Speaker	Emotion	Sentiment	Dialogue_ID	Utterance_ID	Season	Episode	Start Time	End Time
1	also I was the point p	Chandler	neutral	neutral	0	0	8	21	00:16:16,059	00:16:21,731
2	You must ve had yo	The Interviewe	neutral	neutral	0	1	8	21	00:16:21,940	00:16:23,442
3	That I did. That I did.	Chandler	neutral	neutral	0	2	8	21	00:16:23,442	00:16:26,389
4	So let s talk a little	The Interviewe	neutral	neutral	0	3	8	21	00:16:26,820	00:16:29,572
5	My duties? All right.	Chandler	surprise	positive	0	4	8	21	00:16:34,452	00:16:40,917
6	Now you ll be headi	The Interviewe	neutral	neutral	0	5	8	21	00:16:41,126	00:16:44,337
7	I see.	Chandler	neutral	neutral	0	6	8	21	00:16:48,800	00:16:51,886

MP4

dia1_utt5.mp4

Основной датасет: **MELD**

Источник: сериал **'Friends'**

Формат данных **csv** + видео
файлы формата **mp4**

Преимущества и недостатки

Преимущества:

Мультимодальность (аудио, видео, текст)

Явные аннотации эмоций

Метки времени

Недостатки:

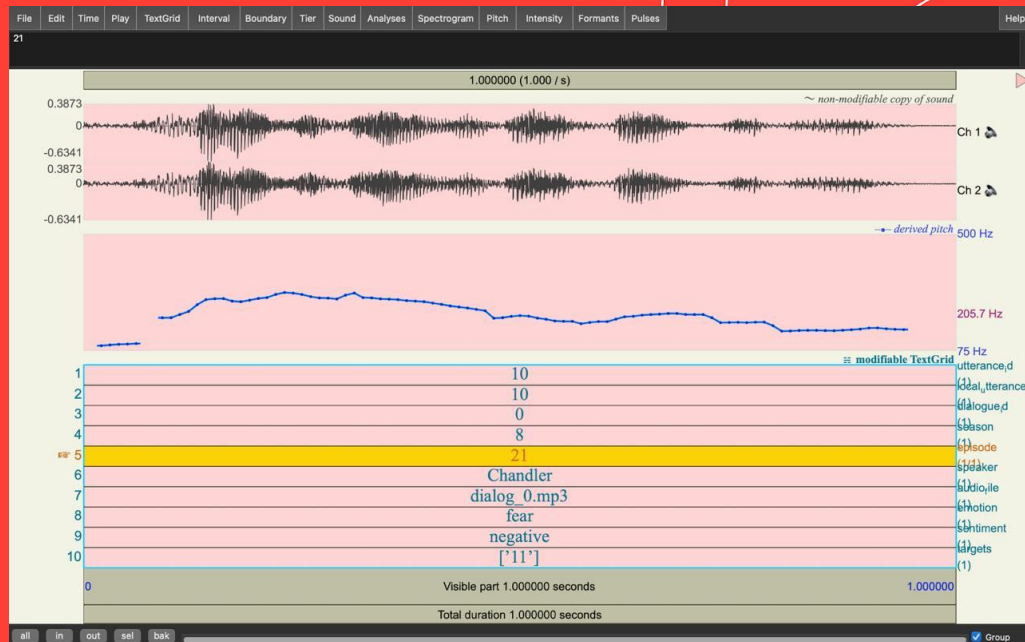Актерская речь (неестественность)

Сценарный контекст

Аннотационная схема

Уровень разметки	Тип разметки	Модальность	Визуализация	Идентификаторы
Диалог	Структурная разметка	Текст, аудио	Таблицы	Диалоговая структура: ● dialog_id ● season ● episode
Участник	Идентификационная разметка	Текст	Таблица	Спикер в диалоге: ● speaker ● dialogue_id
Реплика	Сегментационная + временная разметка	Текст + аудио	ELAN	Реплика участника: ● dialogue_id ● speaker ● utterance ● start_time ● end_time ● audio_file ● targets

ИТМО

Аннотационная схема

ИТМО

Уровень разметки	Тип разметки	Модальность	Визуализация	Идентификаторы
Адресация	Идентификационная разметка	Текст	Граф	Адресация фразы: ● utterance_id ● dialogue_id ● targets
Эмоция	Категориальная разметка	Текст + аудио	Разметка в таблице	Эмоция: ● utterance_id ● эмоция
Сентимент	Категориальная разметка	Текст + аудио	Разметка в таблице	Сентимент: ● utterance_id ● sentiment

Разметка

Разметка с помощью программы Praat

- utterance_id
- local_utterance_id
- dialogue_id
- season
- episode
- speaker
- start_time,end_time
- audio_file
- emotion
- sentiment,targets
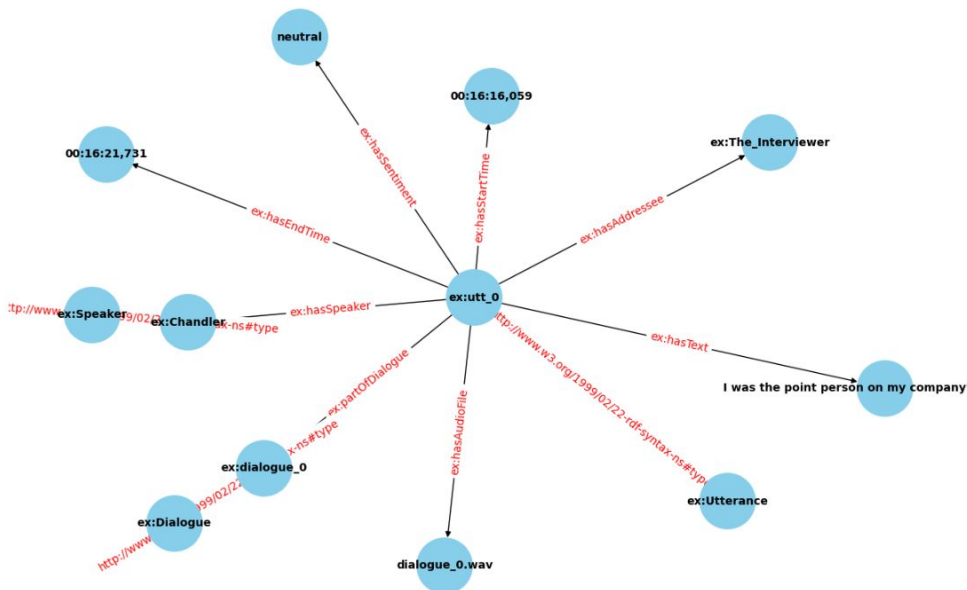
Пример разметки (текст)

ИТМО

Utterance_id	Local_utterance	Dialogue_id	Season	Episode	Speaker	Utterance	Emotion	Sentiment	Targets
0	0	0	8	21	Chandler	also I was the po	neutral	neutral	['1']
1	1	0	8	21	The Interviewer	You must've had	neutral	neutral	['2']
2	2	0	8	21	Chandler	That I did. That I	neutral	neutral	['3']
3	3	0	8	21	The Interviewer	So let's talk a litt	neutral	neutral	['4']
4	4	0	8	21	Chandler	My duties? All ri	surprise	positive	['5']
5	5	0	8	21	The Interviewer	Now you'll be he	neutral	neutral	['6']
6	6	0	8	21	Chandler	I see.	neutral	neutral	['7']
7	7	0	8	21	The Interviewer	But there'll be pe	neutral	neutral	['8']
8	8	0	8	21	Chandler	Good to know.	neutral	neutral	['9']
9	9	0	8	21	The Interviewer	We can go into	neutral	neutral	['10']
10	10	0	8	21	Chandler	No don't I beg c	fear	negative	['11']
11	11	0	8	21	The Interviewer	All right then, we	neutral	neutral	['12']
12	12	0	8	21	Chandler	Really?!	surprise	positive	['13']
13	13	0	8	21	The Interviewer	Absolutely. You	neutral	neutral	None
14	0	1	9	23	Joey	But then who? T	surprise	negative	['15']
15	1	1	9	23	Rachel	You know? Forg	sadness	negative	['16']
16	2	1	9	23	Joey	No-no-no-no, n	surprise	negative	['17']
17	3	1	9	23	Rachel	No, I-I-I-I don't,	fear	negative	['18']
18	4	1	9	23	Joey	Ok!	neutral	neutral	['19']
19	5	1	9	23	Joey	All right, well...	neutral	neutral	['19']
20	8	1	9	23	Rachel	Yeah, sure!	neutral	neutral	None

Пример разметки (аудио)

ИТМО

Utterance_id	Local_utterance	Dialogue_id	Season	Episode	Speaker	Start_time	End_time	Audio_file	Emotion	Sentiment	Targets
0	0	0	8	21	Chandler	00:16:16,059	00:16:21,731	dialog_0.mp3	neutral	neutral	['1']
1	1	0	8	21	The Interviewer	00:16:21,940	00:16:23,442	dialog_0.mp3	neutral	neutral	['2']
2	2	0	8	21	Chandler	00:16:23,442	00:16:26,389	dialog_0.mp3	neutral	neutral	['3']
3	3	0	8	21	The Interviewer	00:16:26,820	00:16:29,572	dialog_0.mp3	neutral	neutral	['4']
4	4	0	8	21	Chandler	00:16:34,452	00:16:40,917	dialog_0.mp3	surprise	positive	['5']
5	5	0	8	21	The Interviewer	00:16:41,126	00:16:44,337	dialog_0.mp3	neutral	neutral	['6']
6	6	0	8	21	Chandler	00:16:48,800	00:16:51,886	dialog_0.mp3	neutral	neutral	['7']
7	7	0	8	21	The Interviewer	00:16:48,800	00:16:54,514	dialog_0.mp3	neutral	neutral	['8']
8	8	0	8	21	Chandler	00:16:59,477	00:17:00,478	dialog_0.mp3	neutral	neutral	['9']
9	9	0	8	21	The Interviewer	00:17:00,478	00:17:02,719	dialog_0.mp3	neutral	neutral	['10']
10	10	0	8	21	Chandler	00:17:02,856	00:17:04,858	dialog_0.mp3	fear	negative	['11']
11	11	0	8	21	The Interviewer	00:17:05,025	00:17:13,324	dialog_0.mp3	neutral	neutral	['12']
12	12	0	8	21	Chandler	00:17:13,491	00:17:16,536	dialog_0.mp3	surprise	positive	['13']
13	13	0	8	21	The Interviewer	00:17:17,579	00:17:20,707	dialog_0.mp3	neutral	neutral	None
14	0	1	9	23	Joey	00:36:40,364	00:36:42,824	dialog_1.mp3	surprise	negative	['15']
15	1	1	9	23	Rachel	00:36:44,368	00:36:46,578	dialog_1.mp3	sadness	negative	['16']
16	2	1	9	23	Joey	00:36:44,368	00:36:49,122	dialog_1.mp3	surprise	negative	['17']
17	3	1	9	23	Rachel	00:36:49,290	00:36:51,791	dialog_1.mp3	fear	negative	['18']
18	4	1	9	23	Joey	00:36:52,376	00:36:53,543	dialog_1.mp3	neutral	neutral	['19']
19	5	1	9	23	Joey	00:36:53,545	00:36:55,000	dialog_1.mp3	neutral	neutral	['19']
20	8	1	9	23	Rachel	00:36:59,425	00:37:01,439	dialog_1.mp3	neutral	neutral	None

RDF-граф (rdflib)

ИТМО

RDF-граф: пример связи utterance-сущностей

Классы (RDF Types):

- Dialogue — диалог
- Utterance — реплика
- Speaker — спикер
- Emotion — эмоция
- Sentiment — сентимент

Свойства (RDF predicates):

- hasSpeaker — реплику произнёс спикер
- hasText — текст реплики
- hasStartTime, hasEndTime — временные метки
- hasAudioFile — путь к аудиофайлу
- hasEmotion, hasSentiment
- hasAddressee — адресат
- partOfDialogue — принадлежит диалогу

Примеры запросов к графу знаний   ИТМО

 # 1. Найти все реплики Чендлера

```
query1 = """
PREFIX ex: <http://example.org/emotion/>
SELECT ?utterance ?text
WHERE {
  ?utterance a ex:Utterance ;
        ex:hasSpeaker ex:Chandler ;
        ex:hasText ?text .
} """
```

 # 2. Найти все utterance'ы с эмоцией surprise

```
query2 = """
PREFIX ex: <http://example.org/emotion/>
SELECT ?utterance ?emotion ?speaker
WHERE {
  ?utterance a ex:Utterance ;
        ex:hasEmotion "surprise" ;
        ex:hasSpeaker ?speaker .
} """
```

```
1 q = """
2 PREFIX ex: <http://example.org/emotion#>
3
4 SELECT ?utterance ?text
5 WHERE {
6   ?utterance ex:emotion "surprise" .
7   ?utterance ex:utteranceText ?text .
8 }
9 """
10
11 for row in g.query(q):
12     print(f"Utterance URI: {row.utterance}, Text: {row.text}")
```

```
Utterance URI: http://example.org/emotion#utterance_4, Text: My duties? All right.
Utterance URI: http://example.org/emotion#utterance_12, Text: Really?!
Utterance URI: http://example.org/emotion#utterance_14, Text: But then who? The waitress I went out with last month?
Utterance URI: http://example.org/emotion#utterance_16, Text: No-no-no-no, no! Who, who were you talking about?
```

Спасибо
за внимание!

iT'sMOre than a
UNIVERSITY