Acceptability & Complexity evaluation task for Italian at EVALITA 2020

Gerald Cakoni
cakoni.1985088@studenti.uniroma1.it

Abstract

AcCompl-It is task's purpose is to develop and evaluate methods to classify the italian sentences with a scale from 1 to 7 regardles to the acceptability and complexity based on the native speaker judgement. The Acceptability is the judgement based on the correctness of the sentences. On the other hand Complexity is the difficulty to understand and process the sentence. It use neurolinguistic and psycholinguistic theories.

1 Dataset Input

The dataset provided, see table 1 is in TSV format and it has the sentence with some features that are extracted from native speakers evaluation like: number of human scores associated with the sentence, mean value, standard deviation, standard error and the task.

2 Dataset Transformation

To transform the dataset I was based on the mean result of eveluation because this value express better the overall level for each task. As asked from the requirements I transformed the 1-7 likert scale from numbers to Italian word. Then I grouped each evaluation to end up with only 4 choices as the final possible result. The convertion of intervals and numbers to words used to evaluate each result is given in the tables 3 and 2.

3 Prompts

{"id": "1", "text": "Quando gli invitati erano in ritardo lui avvertiva sempre il ristorante.", "choices": ["Scarsa", "Sufficiente", "Buona", "Eccellente"], "label": 3}

Acceptability

To chose the most appropriate promts for the ACCOMPLIT task I tried to be as clear as possible while writing the promt. I was very specific when asking for the response and referred to the information given in the dataset using Jinja template. I started with simple promts and later going further into more complex ones.

- Valuta l'accettabilità della seguente frase: {{text}}. Scegli tra le opzioni {{choices}}.

- Determina l'accettabilità della frase fornita {{text}} e seleziona la migliore opzione tra le scelte {{choices}}.

- Confronta la frase fornita {{text}} con le scelte indicate {{choices}} e seleziona l'opzione che rappresenta meglio la sua accettabilità.

- Considerando il contesto, quale delle opzioni dalla seguente lista: Scarsa, Buona, Ottima o Eccellente, valuta la correttezza della frase {{text}}.

Complexity

- Valuta il livello di complessità della frase {{text}}. Scegli tra le opzioni {{choices}}.

- Considerando la complessità linguistica, la difficoltà della seguente frase {{text}} è Molto Bassa, Moderata, Alta o Massima?

- Confronta la complessità della frase fornita 'text' con le opzioni indicate {{choices}} e seleziona il livello più appropriato.

- Valuta la complessità sintattica e semantica della frase fornita {{text}} e assegna un grado di difficoltà tra 'Molto Bassa', 'Moderata', 'Alta', 'Massima'.

4 Instructions

Download the datasets and add them to the same directory as script.py file. Open the terminal and run the following code: python scripts.py. This will create the dataset in the JSONL asked format.

5 Tables

Table 1: Acceptability and Complexity input data

id	N	mean	sd	se	sentence	task
AC-01-R0001	14	5.285714	1.325987	0.354385	Quando gli invitati erano in ritardo lui avvertiva sempre il ristorante.	acceptability
AC-01-R0002	10	6.7	0.674949	0.213437	I ciclisti hanno pedalato.	acceptability
AC-01-R0003	10	6.8	0.632456	0.2	I malati hanno starnutito.	acceptability
AC-OC-02-R0001	13	3.384615	1.26085	0.349697	Chi dici che l'infermiera voglia chiamare e il dottore visitare?	complexity
AC-OC-02-R0002	15	2.866667	1.302013	0.336178	Sapevo che l'infermiera voleva registrare e Marco visitarlo .	complexity

Table 2: Acceptability converting numbers to words and grouping in 4 choices

Number	1	2	3	4	5	6	7
Mean	1-1.5	1.5-2.5	2.5-3.5	3.5-4.5	4.5-5.5	5.5-6.5	6.5-7
Verbalization	Pessima	Scarsa	Sufficente	Buona	Molto Buona	Ottima	Eccellente
Group Mean	1-2.5		2.5-4.5		4.5-6.5		5-7
Choices	Scarsa		Buona		Ottima		Eccellente

Table 3: Complexity converting numbers to words and grouping in 4 choices

Number	1	2	3	4	5	6	7
Mean	1-1.5	1.5-2.5	2.5-3.5	3.5-4.5	4.5-5.5	5.5-6.5	6.5-7
Verbalization	Molto Bassa	Bassa	Moderata	Media	Alta	Molto Alta	Massima
Group Mean	1-2.5		2.5-4.5		4.5-6.5		5-7
Choices	Molto Bassa		Moderata		Alta		Massima