Annotation Guidelines

Goal/Task

In this annotation project, we are interested in knowing the sentiment of the reviews (positive, negative, or neutral), and the subject of the review, if applicable. For subjects, we make use of pre-defined labels like **equipment**, **location**, **hygiene**, and **staff**. The goal is to put the sentence/paragraph from the review at hand into the most probable class (determined by you).

Sentiment analysis we are doing uses a simple, one-layer annotation system with three distinct, non-overlapping labels: "**POSITIVE**" for reviews expressing a clear and definite positive view, "**NEGATIVE**" for reviews with a clear negative tone, and "**NEUTRAL**". Notably, our approach uniquely handles mixed sentiment reviews. When a review contains both positive and negative elements, annotators in Label Studio¹ are instructed to select both "**POSITIVE**" and "**NEUTRAL**" labels. This combination categorizes the review as "**NEUTRAL**". It's important to note that the "**NEUTRAL**" label cannot be selected independently in Label Studio.

Topic has also a *one-layer* annotation scheme, including four possible (and potentially overlapping) labels: **EQUIPMENT**, for those reviews related to the fitness equipment in the gym, **LOCATION**, for those reviews related to the geographical location of the gym, **HYGIENE**, for reviews related to the cleanliness of the facilities, and **STAFF**, for reviews related to the staff running the gyms (how friendly they are etc.). Although the "**NOT DETERMINED**" topic is not explicitly selectable in Label Studio, it is assigned to a review if none of the four available topics (Equipment, Hygiene, Staff, Location) are applicable. The "**NOT DETERMINED**" option is assigned when you feel the review label is not present in the given options. The four determined labels can be overlapping, meaning more than one label can be chosen for each review.

Preliminaries

Below, we introduce the labels of this annotation project. It takes around X minutes to read.

A brief explanation of the Subject classes:

- 1. Equipment: This refers to specific or general aspects of the physical machinery installed in the gymnasium. More specifically, it refers to the various tools, machines, devices, and gear used by individuals to engage in physical exercises and activities. Examples of equipment include machines like treadmills, stationary bikes, free weights (dumbbells, barbells), resistance bands, medicine balls, and overall accessories, like weightlifting gloves, lifting belts, etc.
- Location: This refers to the actual geographical location of the fitness facility, as well as its surroundings and ease of access. More specifically, it could include things like neighborhood, nearby public transportation, parking spaces, etc.

¹An open-source data labeling tool for labeling, annotating, and exploring many different data types.

- a. **Hygiene:** This refers to the overall maintenance of a facility's cleanliness and order. It could also be extended to things like how easy it is to find the equipment, the state of bathrooms and showers, lockers, etc.
- b. **Staff:** This refers to the staff who work around the facility. More specifically, here we look at their competence and demeanor, both of which influence the overall gym experience.
- c. **Not determined:** For those cases where the subject of the review is not directly and/or clearly stated.

Few examples review



Labels (Sentiment and Topic)

- 1. Positive, Equipment; Negative, Hygiene
- 2. Negative, Equipment
- 3. Positive, Staff
- 4. Negative, Hygiene
- 5. Positive, Location, Staff, Equipment