

COMPSCI 4NL3: Natural Language Processing Annotation Guidelines

Team 4

Junnan Li
Nawaal Fatima
Rashad Bhuiyan
Sumanya Gulati

02 February 2025

1 Annotation Guidelines

Below are the guidelines developed for the annotation process and to be agreed upon by the annotators. Note that these guidelines have been specifically developed for annotating interview responses from asapsports and should not be directly used to annotate other types of text before considering and adapting to relevant contexts and media. It must also be noted that all examples provided here are related to NBA or WNBA interview responses.

2 Accessing the Dataset

To access the dataset, the annotator is asked to navigate to the **annotations** folder which can be found in the same directory as the annotation guidelines document. Here, you will find 8 CSV files labelled as interview1_annotations.csv, interview2_annotations.csv and so on till interview8_annotations.csv, along with a markdown file titled README.md.

The markdown file contains detailed information about how the dataset was curated and how the steps can be replicated by another user, if desired. To begin annotating, each annotator must open the specific CSV file using MS Excel (recommended) or Google Sheets.

3 Labels and Definitions

The labels are divided into the classes - **focus** and **sentiment**. Focus can be either of the two labels - **individual** or **team** and similarly, sentiment can be either **positive** or **negative**. Focus has to do with whether the response focuses on the *individual (self)* or the *team* and sentiment is based on whether the speaker is expressing *positive* or *negative* emotion.

4 Rules

For annotating focus, think about whether the coach/player emphasizes teamwork or individual contributions. If both are labels are applicable, choose the one that seems more appropriate (has a stronger inference) and if both of carry the same inference, select the label whose instance occurs first .

For annotating sentiment, think about the tone coach/player is trying to convey. If they are excited, hopeful, or showing a growth-mindset this would be considered positive. However, if the response is lacking confidence or conveys disappointment, this would be labelled negative.

When annotating, each annotator is expected to mark the appropriate label for each class with a **1** or a **Yes** in that respective column for each datapoint (or row). Please note that each class must have **only 1** label. This means, there are 4 possible combinations for each datapoint:

- (Team, Positive)
- (Team, Negative)
- (Individual, Positive)
- (Individual, Negative)

5 Case Examples

1. Yeah, we are not down. We're positive. This is a group that believes. We didn't get an opportunity to get a split or win two here on the road. Now Boston held serve. Now we've got to go home and hold serve.

- **Assigned Labels: Positive, Team**

This has been labelled positive as the overall sentiment is hopeful. They mention that they are not down and that this is a group that believes in themselves so they are determined to win. 'We' has been used multiple times when talking about their attitude so this is a sign of a team-focused mentality.

2. Big. The small things, you know, we have to do the small things, and that's part of the game. Those are points that we left on the board, and we didn't shoot free throws well tonight, and we have to be better.

- **Assigned Labels: Negative, Team**

This has been labelled negative as they say mention how they didn't do well which cost them points on the board. It shows disappointment and a desire to improve. It is also team-focused as they are talking about the entire team as a collective and not any single player.

3. Yeah, I think Luka is a special player. He's one of, if not the best player in the world, and he causes a problem. He's able to find guys. Again, creating open opportunities, and we just didn't take advantage of it.

- **Assigned Labels: Positive, Individual**

This has been labelled positive as they are complimenting a player. The comments are mainly focused on a singular player and how they played thus it has been labelled individual.

6 Contact Information

Please contact one of the following members of Team 4 for additional information and issue resolution:

- Junnan Li: lij499@mcmaster.ca
- Nawaal Fatima: fatimn8@mcmaster.ca
- Rashad Bhuiyan: bhuiyr2@mcmaster.ca
- Sumanya Gulati: gulats10@mcmaster.ca