**Annotation Guidelines**

**General :**

**Annotation Guidelines**

**1. Objective:**

Clearly state the main goal or objective of the annotation task.

**2. Data and Format:**

Provide information about the data to be annotated, its format, and any preprocessing steps that have been taken.

**3. Annotation Categories:**

List and define the specific categories or labels that annotators will assign during the annotation process.

**4. Annotation Instructions:**

**4.1 Language Analysis:**

* Specify the language(s) to be considered during analysis.
* Provide guidance on handling special characters.
* Instruct annotators on considering emoticons in the tweet.

**4.2 Sentiment Classification:**

* Define criteria for assigning each sentiment category (e.g., Positive, Negative, Sarcastic, Neutral).
* Include examples for each sentiment category.

**4.3 Neutrality and Balance:**

* Emphasize the importance of maintaining neutrality during annotation.
* Instruct annotators to focus on language rather than personal views expressed in tweets.

**4.4 Special Cases:**

* Address scenarios such as mixed sentiment, part positive/part negative language, and irrelevant tweets.
* Clearly define criteria for handling these cases.

**4.5 Handling Ambiguities:**

* Instruct annotators on how to deal with uncertain or ambiguous cases.
* Provide a mechanism for annotators to seek clarification on ambiguous instances.

**5. Tools and Resources:**

Specify any tools, resources, or dictionaries that annotators should use during the annotation process.

**6. Quality Control Measures:**

Outline steps for ensuring annotation quality, including regular checks and feedback loops.

**7. Inter-Annotator Agreement:**

Define how inter-annotator agreement will be calculated and the metrics to be used.

**8. Examples:**

Provide illustrative examples for each sentiment category to guide annotators.

**9. Collaboration and Feedback:**

Encourage communication and collaboration among annotators. Establish a feedback mechanism for questions and clarifications.

**10. Conclusion:**

Summarize the guidelines, express gratitude to annotators, and provide any final instructions.

**Medical Domain:**

**Annotation Guidelines for Medical Concept Extraction**

**1. Objective:**

The objective of this annotation task is to extract specific medical concepts and patient information from a dataset within the medical domain.

**2. Data and Format:**

* The dataset consists of medical documents, including patient records and reports.
* Information to be extracted includes diseases, patient names, ages, dates of birth (DOB), and medications.

**3. Annotation Categories:**

Annotators are tasked with extracting the following information:

* **Diseases:** Identifying and labeling specific diseases mentioned in the text.
* **Patient Names:** Extracting the names of patients mentioned in the records.
* **Ages:** Identifying and annotating patient ages.
* **Date of Birth (DOB):** Extracting and annotating the date of birth of patients.
* **Medications:** Identifying and labeling names of medications mentioned in the records.

**4. Annotation Instructions:**

**4.1 Disease Extraction:**

* Annotate diseases mentioned in the text.
* Use standard medical terminology to identify and label diseases.
* Include diseases and related terms such as symptoms or conditions.

**4.2 Patient Names:**

* Extract and annotate names of patients mentioned in the records.
* Pay attention to variations in name formats.

**4.3 Age Extraction:**

* Identify and annotate patient ages mentioned in the records.
* Extract numerical representations of ages.

**4.4 Date of Birth (DOB):**

* Extract and annotate the date of birth of patients.
* Annotate complete DOB in the format YYYY-MM-DD.

**4.5 Medication Extraction:**

* Annotate names of medications mentioned in the records.
* Include both generic and brand names.

**4.6 Handling Ambiguities:**

* If information is unclear or ambiguous, provide a comment indicating uncertainty.
* Annotate only information explicitly present in the text.

**4.7 Special Cases:**

* Address scenarios where information is mentioned in a narrative or non-standard format.
* Clearly define criteria for handling complex cases.

**5. Tools and Resources:**

* No specific tools are required for this annotation task.
* Annotators may refer to medical dictionaries for clarification.

**6. Quality Control Measures:**

* Regular checks will be conducted to ensure consistency across annotators.
* Annotators are encouraged to ask for clarification on ambiguous cases.

**7. Inter-Annotator Agreement:**

* Kappa statistic will be calculated to measure inter-annotator agreement.
* Feedback sessions will be organized to discuss discrepancies and improve consistency.

**8. Examples:**

1. **Disease Extraction:**
   * "The patient was diagnosed with [Disease]."
2. **Patient Names:**
   * "Patient [Name] reported experiencing..."
3. **Age Extraction:**
   * "The patient, aged [Age], presented..."
4. **DOB Extraction:**
   * "DOB: [DOB] - The patient was born on..."
5. **Medication Extraction:**
   * "Prescribed [Medication] for treatment."

**9. Collaboration and Feedback:**

* Annotators are encouraged to communicate and seek clarification on challenging cases.
* Regular feedback sessions will be conducted to address questions and improve annotation consistency.

**10. Conclusion:**

Thank you for contributing to this medical concept extraction task. Consistent and accurate annotations are essential for the success of this project.