Project Proposal

Data Labeling Approach

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML in solving this task?	The project aims to address the challenge of accurately diagnosing pneumonia from chest X-ray images, a task that currently requires expert medical interpretation. By using Machine Learning (ML), we intend to enhance the speed and efficiency of this diagnostic process, potentially aiding in early detection and treatment. ML algorithms can learn from vast datasets of labeled X-rays, enabling them to assist in identifying telltale signs of pneumonia, which is especially valuable in settings with limited access to specialized medical personnel. Note that it's not a complete replacement to a medical professional but a good assistant.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels vs any other option?	For this project, we chose three data labels: 'Healthy', 'Pneumonia', and 'Uncertain'. These labels were selected to provide a clear and straightforward framework for annotating chest X-ray images. 'Healthy' and 'Pneumonia' allow for direct categorization of clear cases, while 'Uncertain' addresses the inherent complexity and ambiguity in some images. This tripartite labeling system balances simplicity for non-expert annotators. Other more complex labeling options were avoided to ensure the task remains accessible to annotators without specialized medical training. In addition to the categorical labels, we implemented a scale annotation ranging from 1 (indicative of 'Healthy') to 6 (suggestive of 'Pneumonia'). This scale provides a nuanced measure of the annotator's confidence in their classification, allowing for a more detailed analysis of each image. It captures the degree of certainty and helps in identifying borderline cases, where signs of pneumonia are not clear-cut. This added layer of granularity is particularly useful for subsequent data analysis, enabling us to fine-tune ML algorithms for more accurate detection and to potentially identify subtle patterns or stages in the progression of pneumonia.

Test Questions & Quality Assurance

<table>
<tr>
<td>

Number of Test Questions

Considering the size of this dataset, how many test questions did you develop to prepare for launching a data annotation job?

</td>
<td>

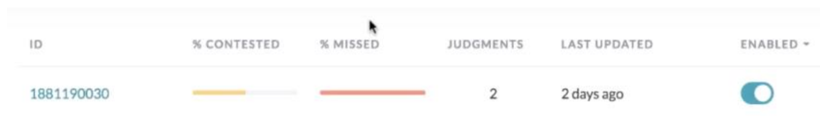
Given the dataset's size and complexity, we developed a set of 9 test questions, carefully crafted to cover all possible cases and the entire scale range. This number strikes a balance between providing sufficient examples for comprehensive understanding and avoiding overwhelming the annotators. These test questions encompass clear-cut examples of 'Healthy' and 'Pneumonia' cases, as well as more ambiguous 'Uncertain' scenarios, ensuring annotators are well-prepared for the variety they will encounter in the actual data. The inclusion of scale-based questions further aids in assessing and improving the annotators' ability to gauge and quantify their certainty in each classification.

</td>
</tr>
<tr>
<td>

Improving a Test Question

Given the following test question which almost 100% of annotators missed, statistics, what steps might you take to improve or redesign this question?

</td>
<td>



Note that we have not actually launched the annotation job in happen, however, the following is based on insights.

To improve or redesign a test question that most annotators missed, the following steps could be taken:

- Analyze the Question: Understand why the question was commonly missed. Was it due to its complexity, ambiguity, or misalignment with the training material?
- Enhance Training Material: Include similar examples in the training material, focusing on the specific features or aspects that made the original question challenging.
- Solicit Feedback: Ask a few annotators for feedback on why they found the question difficult, and use this feedback to make necessary adjustments.
- Iterative Testing: Introduce the revised question in a smaller, controlled group to see if the changes have improved comprehension and accuracy before rolling it out to the entire pool of annotators.

This approach ensures that the test question is effectively assessing the key skills and knowledge required for accurate annotation, thereby enhancing the overall quality of the data labeling process.

</td>
</tr>
</table>

Contributor Satisfaction

Say you've run a test launch and gotten back results from your annotators; the instructions and test questions are rated below 3.5, what areas of your Instruction document would you try to improve (Examples, Test Questions, etc.)

This would just be consistent with the above section improving test questions but in greater details.

In light of the feedback indicating that the instructions and test questions were rated below 3.5, the following areas of the Instruction document would be targeted for improvement, in a manner consistent with the approach for enhancing a test question:

- Analyzing and Clarifying Instructions
- Enhancing Training Material
- Soliciting Feedback on Instructions
- Iterative Improvement and Testing
- Consistency in Approach

This methodical and consistent approach aims to refine both the instructions and test questions in a way that enhances the overall quality and effectiveness of the data labeling process.

Limitations & Improvements

Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	The following could be the potential biases in the data source. Biases in the Data: • Demographic Bias • Equipment and Technique Variability • Age group Bias For this we could take the following actions to correct it. • Diversify the Dataset • Standardize Image Quality • Include Underrepresented Groups
Designing for Longevity How might you improve your data labeling job, test questions, or product in the long-term?	Improving Data Labeling Job, Test Questions, and Product for Longevity: • Continuous Learning and Adaptation: Regularly update the training material and test questions based on the latest medical research and feedback from annotators. • Incorporate AI Assistance: Gradually integrate AI tools to assist annotators in identifying complex patterns, reducing human error. • Expand Data Diversity: Continuously broaden the dataset to include more diverse X-ray images from various demographics and regions. • Iterative Feedback Loop: Establish a systematic feedback mechanism to constantly refine the annotation process and test questions. • Scalability and Flexibility: Design the system to be scalable and adaptable to new types of medical imaging or related tasks.