

Project Proposal



Ugur Uresin

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 product (problem case) is a digital product that helps doctors quickly identify cases of pneumonia in children.

The digital product is a classification system that

1. Can help flag serious cases
2. Quickly identify healthy cases
3. And, generally, act as a diagnostic aid for doctors

This saves time for doctors and helps to prevent undesired mistakes that can be accidentally done by doctors.

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?

In this project, after careful evaluation and understanding the importance of accurate diagnosis for pneumonia, I have chosen to incorporate three distinct labels for the data:

YES (There is a pneumonia risk): This label is crucial for situations where the evidence within the data clearly indicates a risk of pneumonia. I chose this as a separate category because time is of the essence when it comes to diseases like pneumonia. This swift identification allows healthcare providers to intervene at an earlier stage, potentially improving patient outcomes and ensuring that the necessary medical interventions can be initiated without delay.

NO (Patient has no clear pneumonia risk): It's equally vital to quickly identify patients who don't exhibit signs of pneumonia. By categorizing such patients under the "NO" label, I aim to enhance the healthcare process's efficiency. Doctors often face a deluge of cases daily, and by confidently identifying patients without pneumonia risks, healthcare professionals can prioritize and allocate their time and resources more effectively.

UNC (Unclear): Medical data, especially images, can sometimes be ambiguous due to various factors like the quality of the image, the presence of other unrelated anomalies, or even overlapping symptoms. By creating a separate "UNC" category, I emphasize the importance of a second review or a deeper analysis.

Test Questions & Quality Assurance

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?

Considering the size of this dataset I have provided 4 test questions that seems a feasible number of questions.

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?

ID	% CONTESTED	% MISSED	JUDGMENTS	LAST UPDATED	ENABLED
1881190030	<div><div></div></div>	<div><div></div></div>	2	2 days ago	<input checked="" type="checkbox"/>

It can be concluded that either the instruction or the examples are fuzzy. I'd try to get feedback for the annotators for this instance then might take the following actions:

1. Revising the instruction
2. Adding more explanation specific to the instance
3. Trying to provide more examples so that annotators may feel more confident.

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.)

Contributor Satisfaction ⓘ

Number of participants: 20

3.2 / 5

Overall

3.3 / 5

Instructions Clear

2.9 / 5

Test Questions Fair

2.8 / 5

Ease Of Job

3.7 / 5

Pay

Considering the pneumonia project and the "Contributor Satisfaction" metrics:

1. 20 annotators provided feedback, giving an average satisfaction of 3.2 out of 5.
2. The clarity of instructions received a 3.3 rating. For a medical project like pneumonia diagnosis, these instructions might need further refinement to ensure precision.
3. The lower scores of 2.9 and 2.8 for test question fairness and job ease suggest the need for clearer and more straightforward guidelines and questions.
4. Compensation was the highest-rated aspect at 3.7 out of 5.
5. Overall, while payment is satisfactory, the test questions and instructions may require improvements for the critical context of pneumonia diagnosis.

To improve the results I would suggest the following actions:

1. **Revise Instructions:** Redesign the instruction manual by integrating more illustrative examples and perhaps visual aids to make it clearer. A concise FAQ section can be added to address common queries and challenges faced by annotators.
2. **Refine Test Questions:** Engage in a feedback loop with a few annotators to understand specific pain points in the test questions. Based on this, rephrase ambiguous questions and provide clearer context where needed.
3. **Training Workshops:** Consider hosting training sessions or workshops for annotators, specifically focusing on the intricacies of pneumonia diagnosis.

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?	<p>The dataset I possess is notably limited. It comprises 101 unlabeled entries and 16 that are labeled, summing up to 117 in total. Such a restricted dataset size raises concerns about potential substantial sampling bias affecting our final predictions.</p> <p>From the project's outline and based on past experience, it's evident that the images differ marginally in dimensions and have been captured with varying exposure durations. This could introduce measurement bias into our concluding predictions.</p> <p>For better performance and an unbiased product, the team should:</p> <ol style="list-style-type: none">1. Collate a larger dataset,2. Ensure uniformity in the size and exposure duration of the added data.
Designing for Longevity How might you improve your data labeling job, test questions, or product in the long-term?	<p>For the data labeling job's longevity and relevance, I'd emphasize ongoing feedback from annotators. I'd routinely update test questions with evolving diagnostic trends and, with technological progress, consider AI tools to enhance the labeling process while maintaining accuracy.</p>