The Cost of a False Positive

We all know false positives in machine learning can be costly. And while we also know that high quality data is imperative to the success of your algorithm, in some cases, data quality is even more critical than others. For example, a false positive in an autonomous vehicle or biomedical algorithm could mean life or death, however, in the case of an e-commerce chatbot, it may just result in poor customer service.

Since the weight and severity of a false positive differs across verticals, it's important to define the level of data quality and domain expertise needed to train your algorithm, as a part of your training data strategy.

How to Achieve 99.9% Quality for AI Algorithms

1. **Establish a quality rubric to score the integrity of your data.**
   In your quality rubric, possible error types may range between false positive, false negative, loose tagging, etc., and possible score weights may range from a 10-20% penalty for a non-critical error to a 100% penalty for a critical error. See **Exhibit A** for examples of error types and associated penalties.

2. **Define the domain expertise and training needed for data annotators.**
   This is dependent on a number of factors, including workflow complexity, annotators' level of experience, industry vertical, etc. At Sama, our dedicated digital workforce received customized training and certification, so the annotation team assigned to your project is equipped to handle complex tasks from day one.

sama

3. **Conduct daily sampling to test against the quality rubric.**
Random or 100% sampling can be done to evaluate whether results are above or below the defined quality threshold. If using random sampling, ensure the sample is statistically sound and representative of the rest of the data i.e. 40% sampling for complex workflows and 20% for non-complex workflows.

4. **Perform feedback loops to ensure the work is mastered.**
In addition to sampling to test against the quality rubric, establish feedback loops to ensure the work is mastered. Another solution could be adding a review/verify step to help validate.

Exhibit A

Error Type	Definition	Quality Penalty
Omission - Critical	Cases where subjects that are to be annotated are left out.	100%
Extra Tag - Noncritical	Cases where a subject has unwanted tags.	10%
Wrong Label - Critical	Cases where a label on a subject doesn't match project specifications	100%
Cutting - Noncritical	Cases where the subject is left out of the bounding box on a Loose Tag.	10%
Loose Tagging - Noncritical	Cases where the bounding box is not on the coun-tours of the subject on a Tight Tag	10%
Inconsistent Object ID - Critical	Cases where IDs aren't consistent between frames.	100%