Assessment of the crowdsourcing task "*Sentiment of weather-related tweets*"

Introduction
This task and instruction set was prepared by *Mr. Nirmal Kanagasabai*. Assessment of this task has been done by *Miss Barleen Kaur*.

Instructions
All instructions were written in a legible and presentable manner. The task was interesting. Appropriate examples were also given for every possible case. It would have been better if some prior instructions for capturing sentiments through the linguistic representation of sounds were also set. Like for this tweet: *Mmmmmm warm sunny afternoon on a patio! (@ The Tavern Tech Center).* This can be treated as Positive as well as Neutral and hence causes discrepancy.

Ofcourse, it is difficult while framing instructions to capture every peculiarity pertaining to global set of all possible combinations of tweets. Like there were few incidents where a smiley was made which as per instructions pointed to a positive sentiment, yet after getting the context, the sentiments carried were completely opposite. Like: *I want the sunshine back, please come back :)*

Task Design
The questions were well framed and direct. However, answering the questions after reading the tweets was sometimes tricky. The User Interface was pretty simple and interactive.

Also, I found that keeping two questions for each tweet reduced contributor's chances of correctly answering the questions thereby reducing his accuracy and hence expelling him for a lifetime.

Validation
 In the first iteration of assessment, there was a flaw in setting the test validation questions where the selected responses were in direct conflict with the instruction set. Framing wrong answers for the test questions resulted in an every time low accuracy of the contributor so much so that even after signing up for 9 different accounts and trying to attempt the questions, the accuracy spanned from only 40-60% in most of the cases. Out of those 9 attempts, there was just one attempt where I was able to secure an accuracy level of 80% and the mode changed from test to work finally. On passing the test, I received a screen where I was presented with the questions that I answered incorrectly.

Tweets	Q1_Gold	Q1_User	Q2_Gold	Q2_User
RT @mention: @mention yeah me too, i`m freezing man...don`t let that sunshine fool you FOOOOOOOOOOOOOOOOOOOOOOOL! !!!!!!!!!	negative	negative	cold__freezing	cold
in the house washing clothes on this rainy day #BORED	negative	negative	wet__precipitation	wet

Why the fuck all my teachers tryna go outside today?! Its too hot for dat shit	negative	negative	humid__hot	hot

As per instructions, it asked me to choose from only 4 specified weather extremes (Hot, Cold, Wet, Dry) but in the recommendations, it showed me to include precipitation as well along with 'Wet'. These instances created confusion. The challenge that we face here is that contributor has to bear the consequences of settling for low accuracy just because of the chaos. This is unfair in the sense that even though the contributor can raise flag explaining his viewpoint for a specific question which is in conflict with the set validation questions and get it accepted by the task's creator even then, once expelled, he is expelled for life from doing this task.

Also, the maximum limit of the number of questions that a contributor can answer should have been set to a higher value.

After the test validation questions were re-corrected, I was able to achieve an accuracy of even 80-100%.