

Confidence Score - Feature Description

Starting from API V4, Beyond Verbal has added a new feature to its API output called confidence score.

This document explains the provision of a corresponding confidence score to each emotion output & how Beyond Verbal's user community will benefit.

Along with various emotional parameters (Temper, Valence, Arousal) & their group (e.g. low, med, high), Beyond Verbal is now providing a corresponding confidence score (from 55 to 100) for each such output.

The confidence score is a metric that reflects the distributions of likelihoods of the identified output group (e.g. low arousal).

Using this confidence metric, Beyond Verbal will now return the value "ambiguous" when the confidence score of an emotion group does not exceed a predetermined threshold. Accordingly, Beyond Verbal excludes speech segments when they have been deemed to be unanalyzable (e.g. due to audio quality issues like excessive background noise).

By excluding unanalyzable samples, Beyond Verbal will filter out ambiguous & weak speech segments, thereby significantly improving the overall accuracy & performance of our emotion analytics.

Additionally, the confidence score empowers our users to use this parameter during API integration and when customizing their own application logic. For example, on the user's end, the application can exclude any Beyond Verbal output with a confidence score below a specific threshold e.g. 65. This may be helpful in use cases where there is a large voice data set & emotions analytics accuracy is more important than granularity/frequency of analysis.

Beyond Verbal's API output overall accuracy ranges from 70% for more ambiguous speech segments to above 90% for unambiguous & high quality speech segments.

When Beyond Verbal API users setup & adjust their own confidence score filter rules, the following approximate rule of thumb should be helpful to clarify the expected impact on accuracy at varying levels of confidence:

Accuracy increases approximately 4-5% for a 10-point increase of the selection threshold. For example, rejecting samples with confidence score below 65 leads to an increase in Beyond Verbal engine accuracy of 5%.

Therefore, depending on the use case & user's preference for accuracy over total number of analyzed segments, by filtering out segments with lower confidence scores, user can achieve accuracy in the 90% range.