

Opinion Mining of Spanish Customer Comments with Non-Expert Annotations on Mechanical Turk

Bart Mellebeek, Francesc Benavent, Jens Grivolla, Joan Codina, Marta R. Costa-Jussà* and Rafael E. Banchs

Barcelona Media Innovation Center

Av. Diagonal 177, planta 9

08018 Barcelona, Spain

{bart.mellebeek|francesc.benavent|jens.grivolla|
joan.codina|marta.ruiz|rafael.banchs}@barcelonamedia.org



Summary:

We compare different HIT design strategies by evaluating the usefulness of resulting MechTurk annotations to train an Opinion Mining System on Spanish Consumer data.

Four Research Questions:

1. What is the quality of the MechTurk annotations?
2. What is the performance of classifiers trained on the data?
3. Language Barriers: is there a large enough worker pool for Spanish?
4. Is it cheaper to use Mturk than Expert annotations, and, if so, how much?

Conclusions:

1. High Inter-Annotator Agreement between MechTurk and Experts.
2. MechTurk annotations outperform initial noisy data and produce results on par with Expert annotations.
3. After introducing a competence test, getting results for Spanish was easy and rapid.
4. For the task at hand, we saved 65% of an Expert Annotator

