

gold standard

--> annotated test set

10-fold cross validation

taking 1000 tweets
training 800 tweets
test 100 tweets
val 100 tweets

compare to baseline scores

- frequency-baseline
you expect 80% of the tweets to be neutral(?)

- informative baseline
i have a 60% chance that it will rain tomorrow
--> your result need to be higher
otherwise --> why do all the work?

[2]

Step 5: - Pattern in practise (actions)

> ... as a maker of such tools, it has a site of application that has not been intended.

How do you see this, and where will there be talked about certain issues?

< **three levels:**

- fundamental research
without any concrete applications

- IWT* research center (applied)
develop research to improve problems in society (AMiCA)

- industry sponsored efforts

* IWT, Agentschap voor Innovatie door Wetenschap en Technologie

[2]



Fig. 2. regress

[3]