Understanding of Each column Header in CSV :

Fields Name	Description
Unannotated_id	A unique id for each instance created by hash combining of "Text" and "Root Url"
Text	The text on which annotation will be done
Overall_Sentiment	Overall Sentiment Documents belongs NU= Neutral SP= Strongly Positive WP= Weakly Positive SN= Strongly Negative WN = Weakly Negative
Topic_Wise_Sentiment	Pair of topics and their topic wise Documents NU= Neutral SP= Strongly Positive WP= Weakly Positive SN= Strongly Negative WN = Weakly Negative
Domain	The domain names from where the data belongs to
Validation	True = Manual validation is done after completing annotation by annotator and linguistic/validator False = Manual Validation not done
Root Url	The URL from where the data extracted from
Publish_date	The date when the data was published
Parse_time	Data scraping time
Annotator-1	Annotator basic info with which tag he given to this particular data

Annotator-2	Annotator basic info with which tag he given to this particular data
Annotator-3	Annotator basic info with which tag he given to this particular data
Linguistic-1	Validator/linguistic basic info with which tag he given to this particular data

Data Cardinality :

Number of instances: 10296

avg length of sentences: 329.83

median length of sentences: 285.0

mode length of sentences: 214

max length of sentences: 7204

min length of sentences: 142

Number of words: 3395912

Number of unique words: 164517

Number of unique words after stem: 123374

Kappa score: 0.85