

Application of Naive Bayes to Sentiment Analysis of Text

DataSet: Download SataSet from the link: <http://www.cs.cornell.edu/people/pabo/movie-review-data/>

- [polarity dataset v2.0](#) (3.0Mb) (includes [README v2.0](#)): 1000 positive and 1000 negative processed reviews. Introduced in Pang/Lee ACL 2004. Released June 2004.

1. Implement Naive Bayes algorithm from scratch and apply to the Sentiment DataSet (60%).
2. Apply Naive Bayes algorithm using python library (20%)
3. Write a report to show your results and explanation. (20%)

- Submit your homework solution in 2 separate files:
 - 1) Python file (.py or .ipynb)
 - 2) Word file (doc), The Report containing your codes, explanation and achieved results. Write your codes in text format (NOT as images) in the report.
- Write your full name and number of homework assignment in the FILE NAME.
- Do NOT submit files in Zip file format.

Please be informed that your assignment on coding will be checked through Safe Assign in Blackboard. You can submit your assignment 3 times before the deadline. Your assignment will be graded based on mentioned criteria if similarity score is less than 50%

If similarity between 50%-75%, your grade will be penalized 50%.

If similarity is more than 75% then your assignment will be reported to Honor Code committee.

Your assignment will be graded with Zero if any of requirements in the instruction is missing.

Furthermore, any student whose solution may arise a question, will be asked for some explanations as well.