

Practical Natural Language Processing

Instructor: Rasika Bhalerao

Assignment 1

Due January 19

Please submit your solutions using Canvas / Gradescope.

Suppose you have trained a classifier to predict the sentiment of movie reviews. Here are the results on a manually annotated test set:

	Ground truth: positive	Ground truth: negative
Predicted: positive	35	10
Predicted: negative	5	30

1. How many movie reviews are there in this test set?
2. How many Type 1 errors are there?
3. How many Type 2 errors are there?
4. What is the precision of your classifier?
5. What is the recall of your classifier?
6. What is the F1 score of your classifier?

If you are not already familiar with these things in Python, I recommend looking over the syntax:

- Data types and operations
 - Numbers
 - Booleans
 - Strings
- Control structures
 - Functions
 - If/else conditionals
 - For loops over indices or over items
 - While loops
- Data structures
 - Lists
 - Sets
 - Dictionaries
- Classes
 - Instantiating objects
 - Attributes
 - Constructors

- Methods
- Reading and writing to text files
- Numpy