

Grading Criteria for Assignment 3

Total: 100 points + 3 extra points + bonus (up to 10 points)

Any code that crashes on Rod will receive zero.

Running time: maximum 10 minutes for each language. (Very efficient codes might receive a bonus). If, for any language, it takes more than 10 minutes to run main.py, you will receive zero points for that language for both parts A and B.

You should test your code on the provided *dev set* for each language (*language-dev.xml*). The reference precisions below are the results our reference system gets on that same dev set; they should help you estimate how your system is doing. To calculate your grade, we will run your code on a *test set* and compare it to results our reference system and the baseline methods get on the test set.

Part A: 30 points

Classifier + Language	3. KNN ENG	3. KNN SPA	3. KNN CAT	4. SVM ENG	4. SVM SPA	4. SVM CAT
Points	5	5	5	5	5	5
Reference Precision	<u>0.550</u>	<u>0.690</u>	<u>0.705</u>	<u>0.605</u>	<u>0.785</u>	<u>0.805</u>

Language	English	Spanish	Catalan
Baseline Precision	0.535	0.684	0.678

For any language and classifier, if your classifier cannot beat the baseline for that language, you will receive zero for that classifier and language; if your precision is higher than the reference, you will receive the full credit for that language and classifier; otherwise, you will receive a score as follows:

$$\text{Classifier, Language} = 5 - 5 \left(\frac{\text{reference precision} - \text{your precision}}{\text{reference precision} - \text{baseline precision}} \right)$$

Part B: 60 points + bonus (1-10 points)

B.1- Report Feature Extracting methods: 25 points

a	b	c	d	e
5	5	5	5	5

Note: Your experiment report should include the statistics (precision, recall, etc.) you got from each method, your observations, and conclusions (explanations) you obtained.

B.2- Final Classifier: 35 points + bonus (up to 10 points)

Language	English	Spanish	Catalan
Score	13	11	11
Baseline Precision	0.605	0.785	0.805
Reference Precision	0.650	0.810	0.820

If your final classifier does not improve the precision for a language, you will receive zero for that language; if your final classifier precision for a language achieves more than the reference, you will receive the full credit for that language; otherwise, your score will be computed as the following:

$$\text{English score} = 13 - 13 \left(\frac{\text{reference precision} - \text{your precision}}{\text{reference precision} - \text{baseline precision}} \right)$$

$$\text{Spanish and Catalan score} = 11 - 11 \left(\frac{\text{reference precision} - \text{your precision}}{\text{reference precision} - \text{baseline precision}} \right)$$

Also, we will rank the precisions of all the students' submissions. For the top 10 in each language we will give 10-1 bonus points (the first gets 10, second 9, etc.). If you got bonuses in several languages, you will only receive the highest one. To get the bonus, your precision should be higher than the reference precision.

Conclusions: 10 points + 3 extra points.

- Conclusions you draw from this homework, other interesting observations: **10 points**
- Discuss why these languages (English, Spanish and Catalan) differ in WSD tasks: **3 extra points**

Late day policy

10% off for each late day. After three days, the assignment will be given a score of zero.