Assignment 2
L665/B659: Applying Machine Learning Techniques in CL

March 9, 2017

- Write an analysis of the WSD results in the table below. The analysis should be about half a page long. Answer the following questions:
  - Does the optimization improve results overall? For which words does optimization result in the highest gains? Are there cases where it does not have a positive effect?
  - Which metric and which setting for k are the most robust across words? Add an explanation why there are only odd values for k in the table.
  - Which word is the easiest, which the most difficult? Does optimization in these cases help?
  - Is the test set more difficult or easier than the training set? Explain your answer.
  - For which words do you find the most differences between the fine-grained and the coarse-grained evaluation?
- The analysis should read like a part of the paper, so do not answer the individual questions, but write a cohesive text. Make sure that you use a proper academic style. This should not look out of place in a conference paper.
- This is an individual assignment.
- **DUE DATE**: March 22. Please submit via canvas.

Table I. The best scoring metrics and parameter settings found after 10-fold cross-validation on the training set (see text). The scores are the baseline, the default and optimal settings on the training set (average of 10-fold cross-validation), and the fine-grained, medium and coarse scores on the evaluation set respectively. The scores on the evaluation set were computed by the SENSEVAL coordinators. The average scores are computed over the percentages in this table

word	metric	k	M1-M2-M3	baseline	train.def	train.opt	eval.f	eval.m	eval.c
accident	MVDM	3	0.3-3-3	67.0	81.4	90.2	92.9	95.4	98.1
amaze	IB1-IG	1	1.0-500-0	57.9	99.7	100	97.1	97.1	97.1
band	IGTREE	–	0.5-7-4	73.0	85.4	88.8	88.6	88.6	88.6
behaviour	MVDM-IG	9	0.3-5-5	95.9	94.9	96.7	96.4	96.4	96.4
bet-n	MVDM-IG	1	0.0-5-100	25.5	56.7	71.1	65.7	72.6	75.5
bet-v	IB1-IG	3	0.7-3-3	37.3	64.3	88.6	76.9	77.8	81.2
bitter	MVDM-IG	5	0.5-5-100	30.6	57.6	59.1	65.8	66.4	66.4
bother	MVDM-IG	3	0.2-5-100	45.6	72.8	83.6	85.2	87.1	87.1
brilliant	MVDM-IG	1	0.6-2-100	47.3	57.5	58.8	54.6	62.0	62.0
bury	MVDM-IG	3	0.5-5-100	32.4	35.9	46.2	50.2	51.0	51.7
calculate	IB1-IG	7	0.7-3-3	72.0	79.2	83.2	90.4	90.8	90.8
consume	IGTREE	–	0.7-5-5	37.5	32.9	58.8	37.3	43.8	49.7
derive	MVDM	5	0.0-2-100	42.9	63.9	67.3	65.0	66.1	66.8
excess	MVDM-IG	5	0.5-1-1	29.1	82.6	89.3	84.4	86.3	88.2
float-a	IGTREE	–	0.3-3-3	61.9	57.0	73.5	57.4	57.4	57.4
float-n	MVDM-IG	1	0.8-5-5	41.3	50.8	70.2	64.0	65.3	68.0
float-v	IGTREE	–	0.4-2-100	21.0	34.2	44.0	35.4	40.6	44.1
generous	MVDM	15	0.6-5-100	32.5	44.8	49.3	51.5	51.5	51.5
giant-a	IGTREE	–	1.0-500-0	93.1	92.8	94.1	97.9	99.5	100
giant-n	MVDM-IG	5	0.2-5-100	49.4	77.2	82.6	78.8	85.6	97.5
invade	MB1-IG	3	0.1-10-1	37.5	48.0	62.7	52.7	59.2	62.3
knee	MVDM-IG	5	0.0-5-100	42.8	70.3	81.4	79.3	81.8	84.1
modest	MVDM-IG	9	0.0-5-100	58.8	61.1	67.1	70.7	72.8	75.2
onion	IB1	1	0.8-5-5	92.3	90.0	96.7	80.4	80.4	80.4
promise-n	MVDM-IG	5	0.2-5-100	59.2	63.6	75.3	77.0	83.2	91.2
promise-v	IB1-IG	3	0.5-5-10	67.4	85.6	89.8	86.2	87.1	87.9
sack-n	MVDM-IG	1	0.3-3-3	44.3	75.0	90.8	84.1	84.1	84.1
sack-v	IB1	9	1.0-500-0	98.9	97.8	98.9	97.8	97.8	97.8
sanction	MVDM-IG	1	0.5-3-3	55.2	74.9	87.4	86.3	86.3	86.3
scrap-n	IB1	1	0.4-5-100	37.0	58.3	68.3	68.6	83.3	86.5
scrap-v	IGTREE	–	0.7-3-3	90.0	88.3	91.7	85.5	97.8	97.8
seize	IGTREE	–	0.5-5-100	27.0	57.1	68.0	59.1	59.1	63.7
shake	MVDM-IG	7	0.2-5-100	24.7	71.5	73.3	68.0	68.5	69.4
shirt	IGTREE	–	0.7-5-5	56.9	83.7	91.2	84.4	91.8	96.7
slight	IB1-IG	1	0.3-3-3	66.8	92.7	93.0	93.1	93.3	93.6
wooden	IGTREE	–	0.5-1-1	95.3	97.3	98.4	94.4	94.9	94.9
average				54.1	70.5	78.6	75.1	77.9	79.7