Specialist English: Assignment 10

Rebecca J. Stones
rebecca.stones82@nbjl.nankai.edu.cn

Date due: 2 January 2019

In this tenth assignment (worth 5% of the final mark), we will look at the Experimental Results section.

I'll scale the marks on this assignment according to $m \mapsto \min(m, 10)$ for Master's students and $m \mapsto \lceil m/1.3 \rceil$ for Ph.D. students.

My marking will be affected by (a) your English writing, (b) your LaTeX typesetting, (c) your mathematical presentation, and (d) your understanding of the underlying computer science. Basically, I will "peer review" your assignments.

Problem 1 The following paper is by authors from Nankai! It is a good example of how to introduce a figure (although there's some quibbles, e.g. "shows" instead of "plots", and "beat" instead of "outperforms"). For this snippet:

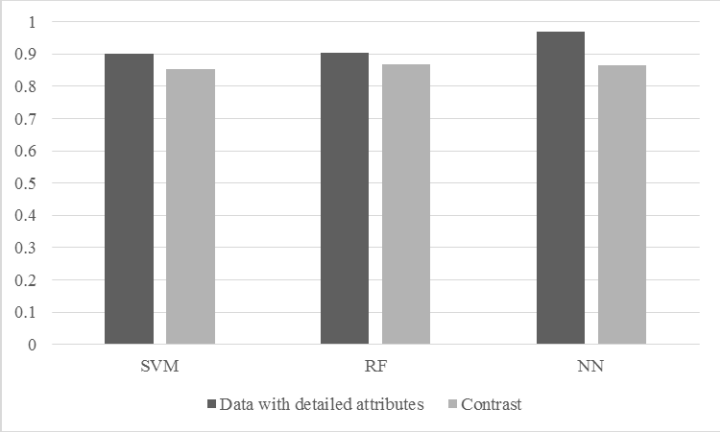
- Identify the sentence which explains what's in Figure 3 in Zhang et al. (2017). [1 mark]
- Identify the sentences, or parts of sentences, which describe observations (i.e., what we see in Figure 3). [2 marks]
- The claims "clearly demonstrates the advantage" and "proves the advantage" are too strongly worded. How can we rephrase this to avoid these problems?

[1 mark]

> Figure 3 shows the accuracy on both two groups of datasets with different ratios of training data in the target domain. The performance of all method improves with more training data in target domain. Under most settings, transfer learning methods perform better than non-transfer learning methods. This fact clearly demonstrates the advantage of transferring knowledge from source to target domain. We can also observe that BLSTM-T and CNN-T beat CDET_J and SCL, which proves the advantage of "domain translation" methods we proposed. Moreover, BLSTM-T performs better than CNN-T, as recurrent neural networks take into account of the emotion of the whole sentence rather than local emotion. Transferring knowledge can bring more gain along with the decreasing of training data in target domain(1/4->1/64). The "domain translation" brings more advantage to the classification accuracy in the case of limited data in the target domain than abundant data in the target domain. In particular, when the training data in target domain is sufficient, neural networks like BLSTM can be better than those transfer learning methods, such as CDET_J.
>
> — Zhang et al., *Weighted Domain Translation …*, SIGIR, 2017.

Problem 2 In the following snippet by Wang et al. (2017):

- How many data points are there in the figure? [1 mark]
- Suggest a more suitable and succinct way to present this data. [1 mark]



(d). F-Measure

Figure 3. Comparison of experiment results between data with detailed attribute and the contrast

— Wang et al., *Machine Learning Methods for Driving Risk Prediction*, EM-GIS, 2017.

Problem 3 For the following two snippets, identify a problem with the numbers. [2 marks]

Table 3. Optimal distribution scheme

UAV	Distribution route	Distribution cost (yuan)	Total distance (km)	Load (kg)
1	0-9-4-2-6-0	160.5635	120.5635	1.83
2	0-7-3-5-0	167.8841	127.8841	1.22
3	0-8-10-1-0	141.8530	101.8530	1.36

Table 4. Comparison between PSO and GA

— Wang et al., *Machine Learning Methods for Driving Risk Prediction*, EM-GIS, 2017.

Table 2: Statistical results of the IGD values of the final populations obtained by three algorithms on the test instances.

instance	MOEA/D		MDM		MDM&DO	
	mean	std.	mean	std.	mean	std.
$D1$	$1.400e-03$	$2.296e-04$	$2.460e-04$	$2.783e-05$	**1.620e-04**	$7.039e-06$
$D2$	$4.253e-04$	$1.259e-04$	$2.396e-04$	$4.498e-05$	**7.025e-05**	$2.404e-05$
$D3$	$2.539e+02$	$1.032e+01$	$1.117e+02$	$4.562e+01$	**1.710e+01**	$1.053e-01$
$D4$	$1.078e+02$	$2.600e+00$	$6.164e+01$	$4.492e+01$	**2.906e+01**	$4.433e+01$
$D5$	$1.112e+00$	$3.860e-02$	$6.780e-01$	$3.389e-01$	**9.630e-02**	$2.300e-03$

— Chen, Zhou, and Dou, *An Evolutionary Algorithm …*, GECCO, 2018.

Problem 4 Why do I consider the following figure not worthy of publication? [2 marks]

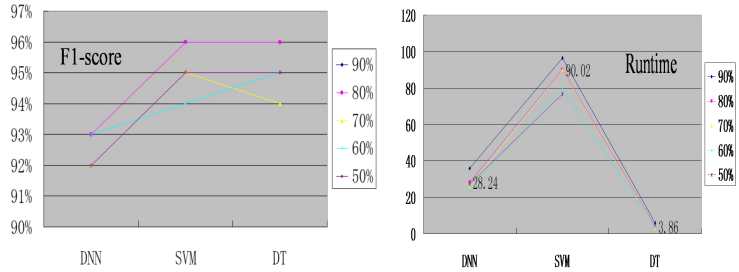


Figure 3: The F1-score and Runtime of the three algorithms

— Ma, Yang, and Zhou, *Using Machine Learning Algorithm …*, ICMSSP, 2018.

Problem 5 Identify two problems with how the following figure is presented. [2 marks]

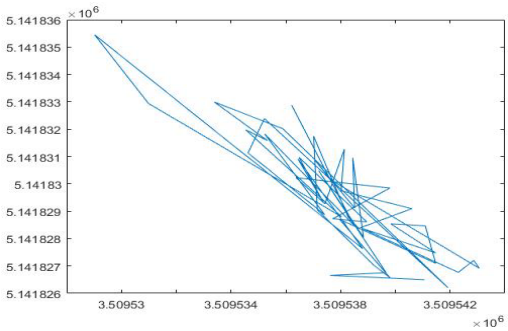


Figure 10: The trajectory of the moving targets in class 6.

— Jiang et al., *A Novel Classification Scheme …*, CSAE, 2018.

Problem 6 What's the problem with the y-axis labels in the following figure? [1 mark]

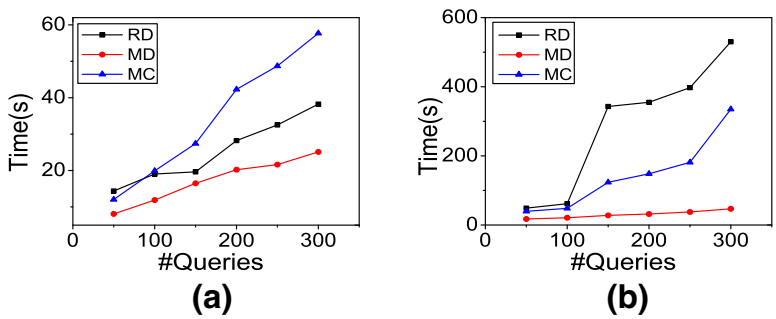
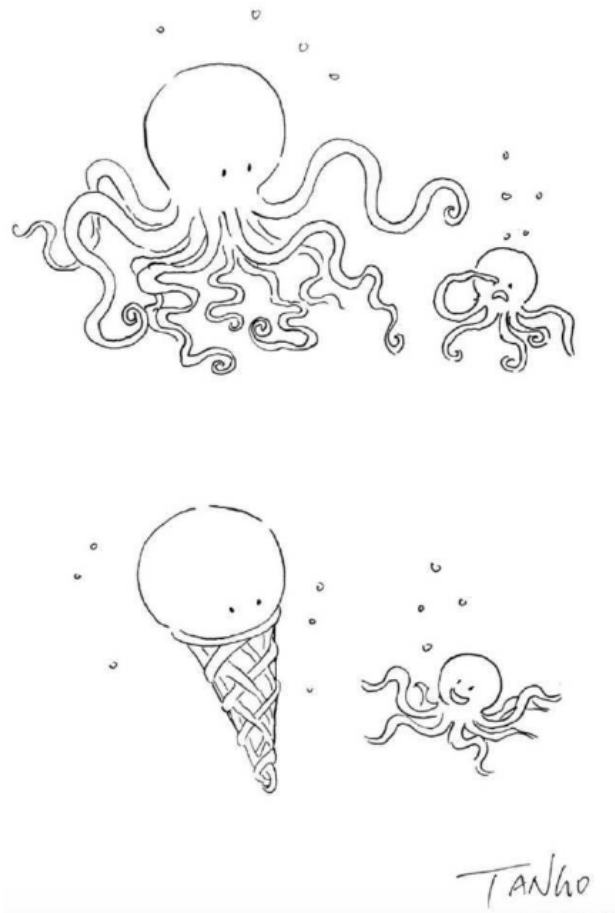


Fig. 10 Sink vertex selection. **a** Varying queries in $\mathcal{Q}^{\mathcal{T}}$. **b** Varying queries in $\mathcal{Q}^{\mathcal{C}}$

— Gao et al., *Path-based holistic detection …*, VLDB Journal, 2018.

Problem 7 Include a funny (wholesome) image in your assignment. Last assignment! Yay!
[0 marks; I can give you a smiley face]

The best answer from last year was by 刘东实 (although I assume it was "borrowed" from the Internet):