**LING 572 – HW7**

**Q3**

Table 1: Training and test accuracy

|  |  |  |
| --- | --- | --- |
|  | Training accuracy | Test accuracy |
| Run Mallet directly | 0.9685185185185186 | 0.84 |
| Run q1.sh (one-vs-all) | 0.968518518518519 | 0.833333333333333 |
| Run q2.sh (all-pairs) | 0.968518518518519 | 0.826666666666667 |

**Q4**

Converting a multi-class program into a binary problem produced very little variances in terms of test accuracies (with all-pairs having the lowest accuracy) and in fact, the training accuracies for all three experiments are almost identical.

However, the ability to convert without having to pay a penalty for much lower performance offers us the option to adapt a multi-class problem into a binary problem so that the problem can be addressed by machine learning algorithm that works best on binary classes. An example of such algorithm is the support-vector machines.

**Additional notes**:

1. The shell script q1.sh will call two code files during execution.

* q1.sh will call q1.pl and write\_final\_q1.pl
* q2.sh will call q2.pl and write\_final\_q2.pl

1. All .sh and .pl files must be in the same directory during execution.
2. During execution, temp files and directories will also be created but these will be deleted automatically when the program exits.

*End of HW7 – Joint submission by*

*Scott Mantei*

*Wee Teck Tan*

*Course Name: LING 572*