

MiRduplexSVM instructions

Train - Test MiRduplexSVM

1. Download miRBase data (miRNA.dat) from <http://www.mirbase.org/ftp.shtml> We have test the code on versions 17(included in the download) and 19.
2. Put the .dat file to “MiRduplexSVM/code/input/data” folder
3. Go in folder “MiRduplexSVM/code”, and run script “init.m”. The performed steps are printed in matlab's command window.

NOTES:

- The fold of the cross validation are set in the second cell.
 - Only human and mouse hairpins are selected to train – test the algorithm. You can change this in cell 6, line 76.
4. Run script “runexpCrossVal.m” to optimize hyper parameters employing 5 fold cross validation.
In the **first cell**, the user can set SVMs hyperparameters. Only the polynomial kernel can be used. The default parameters are the ones used in the MiRduplexSVM publication.
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0126151>
Second cell trains the SVMs
In the **third cell** the user should re-set SVMs hyperparameters.
The **forth cell** tests the models which were produced from the second cell.
The **fifth cell** generate figures with several metrics to evaluate performance.
 5. Run script “runexpHoldOut.m” to train and test the final model using a hold out set.
In the **second cell** the user should provide the desired parameters. The default parameters are the ones used in the MiRduplexSVM publication.
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0126151>
The **last cell** generate figures with several metrics to evaluate performance.

Results

.mat file with the actual numbers of the performance metrics can be found in the “MiRduplexSVM/code/Results” folder.

The cumulative distributions of the errors can be found by following the steps below.

- Load a “_CumFreq_10.mat” file
- Duplex errors, (similar to figure 3), are included in the “meanAbsErrorMeanCumRelFreq” double.
- k55, k53, k35, k33 errors (similar to figure 4) are included in the “f5p5pMeanAbsErrorCumRelFreq”, “f5p3pMeanAbsErrorCumRelFreq”, “f3p5pMeanAbsErrorCumRelFreq”, “f3p3pMeanAbsErrorCumRelFreq” doubles, respectively.

Thank you for using MiRduplexSVM