

# Model-based quantitative transcription regulations (mqTrans) based on linear regression

Fengfeng Zhou, Meiyu Duan, Xin Feng

Update: 2022-03-03

Contact: [FengfengZhou@gmail.com](mailto:FengfengZhou@gmail.com)

## Command line syntax

```
python main.py <TrainGene.csv> <TrainTF.csv> <TestGene.csv> <TestTF.csv> <Gene.csv> <Output.Dir>
```

### Syntax:

<TrainGene.csv>: gene expression profile of the mRNA features in the training dataset

<TrainTF.csv>: gene expression profile of the TF features in the training dataset

<TestGene.csv>: gene expression profile of the mRNA features in the testing dataset

<TestTF.csv>: gene expression profile of the TF features in the testing dataset

<Gene.csv>: the complete list of all the mRNA features

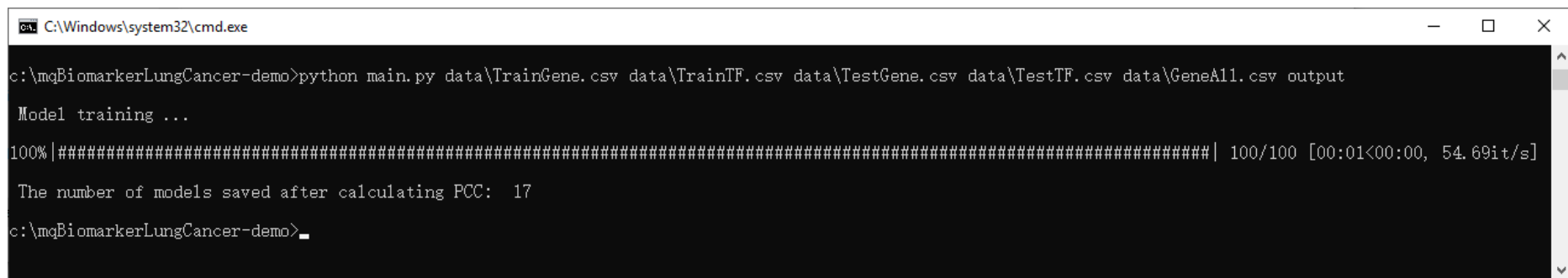
<Output.Dir>: output directory to hold all the result files

The example dataset was retrieved and split from the dataset GSE33356 in the Gene Expression Omnibus (GEO) database.

## Example command line

```
python main.py data\TrainGene.csv data\TrainTF.csv data\TestGene.csv data\TestTF.csv data\GeneAll.csv output
```

The captured screen of the running was displayed here:



```
C:\Windows\system32\cmd.exe

c:\mqBiomarkerLungCancer-demo>python main.py data\TrainGene.csv data\TrainTF.csv data\TestGene.csv data\TestTF.csv data\GeneAll.csv output

Model training ...

100% |#####| 100/100 [00:01<00:00, 54.69it/s]

The number of models saved after calculating PCC: 17

c:\mqBiomarkerLungCancer-demo>
```

## Result files

File name	Description
Matrix_mqTrans.csv	the engineered mqTrans features for the gene expression matrix in the file <TestGene.csv>
mqTrans_Features_PCC.csv	the list of the feature IDs with the mqTrans models, and the PCC values between the predicted and real expression levels of these features
Matrix_mRNA.csv	the transcriptomic features with the mqTrans models for the gene expression matrix in the file <TestGene.csv>
mqTransModels.pkl	the trained mqTrans model for each feature saved in the pickle format
ModelCoef.csv	the coefficients of the TFs for the features with the mqTrans models