

基于支持向量机的非天然地震识别 算法软件

用户操作手册

V1.0

目录

1. 简介	1
2. 运行环境.....	1
3. 安装设置.....	1
4. 软件架构.....	2
5. 界面与操作.....	3
5.1 GUI	3
5.1.1 重新训练.....	4
5.1.2 生成模型.....	6
5.1.3 结果预测.....	7
5.1.4 绘图.....	8
5.2 命令行操作.....	10
6. 说明	11

1. 简介

基于支持向量机的非天然地震识别算法软件（SVM4NneiAS）是一套用于识别天然地震，以及爆破、塌陷、矿震、水库诱发地震等非天然地震的软件。它主要基于机器学习中的支持向量机的算法，利用地震波形提取出的特征值，实现地震事件的分类。其主要作用是协助地震数据处理人员完成地震事件类型的快速判断，为地震速报和科研相关人员提供科学的现实依据。

基于支持向量机的非天然地震识别算法软件（SVM4NneiAS）有独立的图形用户界面（GUI），可方便用户交互操作，易于上手。同时，也可以使用 python 语言进行批处理，最大限度保证用户方便快捷使用。SVM4NneiAS 使用第三方模块较少，安装简便，使用简洁；支持多种操作系统，对电脑性能要求较低。适合各类地震相关从业的用户。

该软件包含的主要功能如下：

- 针对特定研究区域地震特征值提取及重新训练；
- 区域研究预测模型的生成；
- 地震类型的预测；
- 相关的图件绘制。

2. 运行环境

SVM4NneiAS 支持多种操作平台，包括 macOS，Linux 和 Windows 操作系统。运行该程序需要安装 Python 3.x（3.8 已测试可以完美运行，其它版本可自行测试）。

Python 安装方式请参照官方网站：<https://www.python.org/>。

建议使用 anaconda 进行配置安装。Anaconda 安装请参照官方网站：<https://www.anaconda.com/>。

安装完成后，请运行 python 程序，确保安装成功。

```
(nneias) C:\Users\Lin>python --version
Python 3.8.19

(nneias) C:\Users\Lin>python
Python 3.8.19 (default, Mar 20 2024, 19:55:45) [MSC v.1916 64 bit (AMD64)] :: Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> |
```

3. 安装设置

将安装包解压到电脑指定目录，路径中尽量避免中文字符的出现。

建议采用以下安装方法：

（1）安装虚拟环境

```
conda create -n nneias python=3.8
```

```
conda activate nneias
```

(2) 安装运行该程序所需的模块

```
pip install obspy
```

```
pip install scikit-learn
```

或者在主程序目录中输入：

```
pip install -r requirements.txt
```

(3) 运行

进入主程序目录：

```
python run_gui.py
```

若弹出程序主界面，则表示安装成功；否则，请根据 `Trackback` 提示进行其它模块安装，或直接联系作者。

4. 软件架构

软件包含 5 个文件夹：`core`，`data`，`docs`，`out`，`predict_data`。主程序目录下包含 2 个 `python` 主运行程序：`example.py`，`run_gui.py`。分别介绍如下：

- (1) `core`：核心程序包，包含运行该软件的所有程序。
- (2) `data`：输入地震波形文件和配置文件存储区。
- (3) `docs`：说明文档和相关软件图件存储区。
- (4) `out`：输出结果存储区，包括输出 `SAC` 文件、特征值文件、模型、图件等。
- (5) `example.py`：使用命令行运行软件的示例。
- (6) `run_gui.py`：运行 `GUI` 界面的程序。

Tree 图如下：

```

├── core
│   ├── data_process.py # 数据预处理
│   ├── fig.py # 绘图程序
│   ├── find_opt.py # 计算最优特征值
│   ├── nneias.py # 主程序
│   ├── s_gui.py # GUI 程序
│   ├── svm_predict.py # 模型预测
│   ├── svm_train.py # 模型训练
│   └── Train_data.py # 计算特征值
├── data
│   ├── earthquake # 用于训练的天然地震波形数据和震相文件
│   ├── explosion # 用于训练的爆破事件波形数据和震相文件
│   ├── mining # 用于训练的矿震事件波形数据和震相文件
│   ├── reservoir # 用于训练的水库诱发地震波形数据和震相文件
│   └── subsidence # 用于训练的塌陷事件波形数据和震相文件

```

```

|   |—— config.json # 根据震中距截取波形长度的配置文件
|   |—— stations.dat # 台站文件
|—— docs
|   |—— fig
|   |   |—— head.ico
|   |   |—— main_window.png
|   |—— UserGuide.pdf
|—— example.py # 数据批处理程序
|—— out # 输出文件存储文件夹
|—— predict_data # 存储预测波形数据的文件夹
|—— README.md
|—— requirements.txt
|—— run_gui.py # GUI 运行主程序

```

5. 界面与操作

该软件可使用图形用户界面（GUI）进行人机交互，也可以通过命令进行批处理。

5.1 GUI

在主程序目录中执行以下命令：

```
python run_gui.py
```



程序界面主要分为左右两个半区。左半区为用户操作，实现计算和绘图功能；右半区

为数据、图像以及日志的显示窗口。

5.1.1 重新训练

重新训练包括 4 个操作按钮，操作顺序分别为：数据载入、数据预处理、计算所有特征值和计算最佳特征值。

在完成该项功能之前，需要首先归档重新训练的数据以及配置好相关处理参数，具体步骤如下：

(1) 归档数据。在 data 文件夹下分别有 5 个子文件夹，分别为：earthquake（天然地震）、explosion（爆破）、mining（矿震）、reservoir（水库诱发地震）以及 subsidence（塌陷）。请将需要重新训练的地震事件数据 (*.seed) 以及对应的震相文件 (*.phase)，按照分类，放入各个文件夹中。注意 seed 文件和 phase 文件必须一一对应，名称相同。

(2) 编辑 stations.dat 文件。该文件放置于 data 文件夹下，用于存储地震台站经纬度信息。数据结构如下图所示，每一行为每个台站的信息，每一列从左往右分别表示为：台网代码 (Net)、台站代码 (Sta)、纬度 (Lat)、经度 (Lon) 和高程 (Elevation)。

```
GD GZH 23.xxxxx 113.xxxxx xx
GD SHZ 22.xxxxx 114.xxxxx xx
GD SZN 22.xxxxx 114.xxxxx xx
```

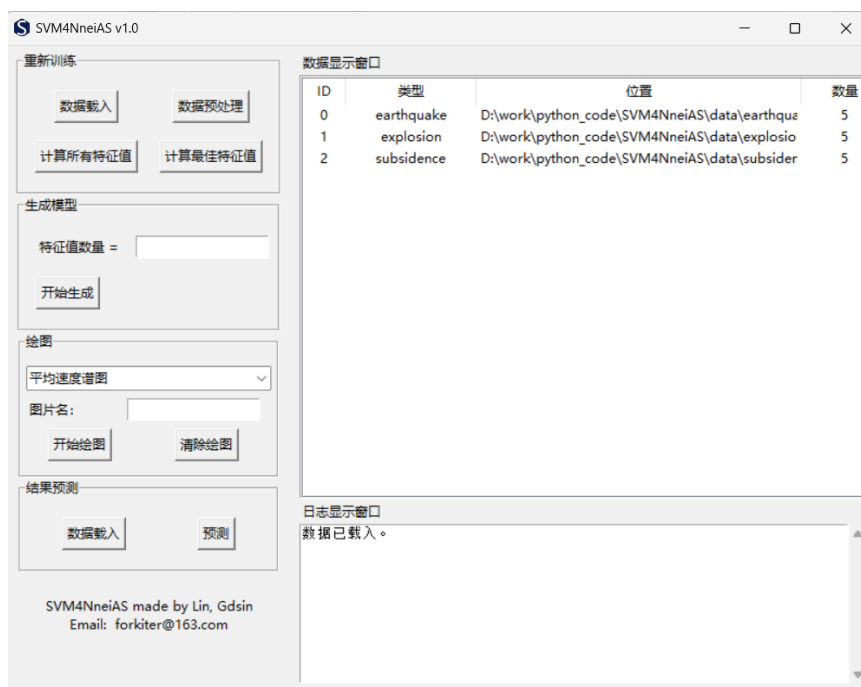
(3) 编辑 config.json 文件。该文件放置于 data 文件夹下，用于配置截取波形的长度。默认文件结构如下图所示。图中“0-25”等的key值表示为震中距范围（公里），[-1,5]等的列表表示该震中距范围内 P 波和 S 波分别截取的长度（秒）。例如：“25-70”：[3, 10]，表示在 25-70km 震中距范围，截取的 P 波数据为 P 波到时之后的 3 秒内数据长度，截取的 S 波数据为 S 波到时之后的 10 秒内数据长度。用户可根据研究区波形特征，设置自己区域的截取范围。注意，图中列表中的-1 表示的是 S-P 的长度距离。

```
{ "0-25": [-1, 5], "25-70": [3, 10], "70-300": [7, 20] }
```

在完成以上配置后，再利用 python run_gui.py 命令，打开主程序界面。

1. 数据载入

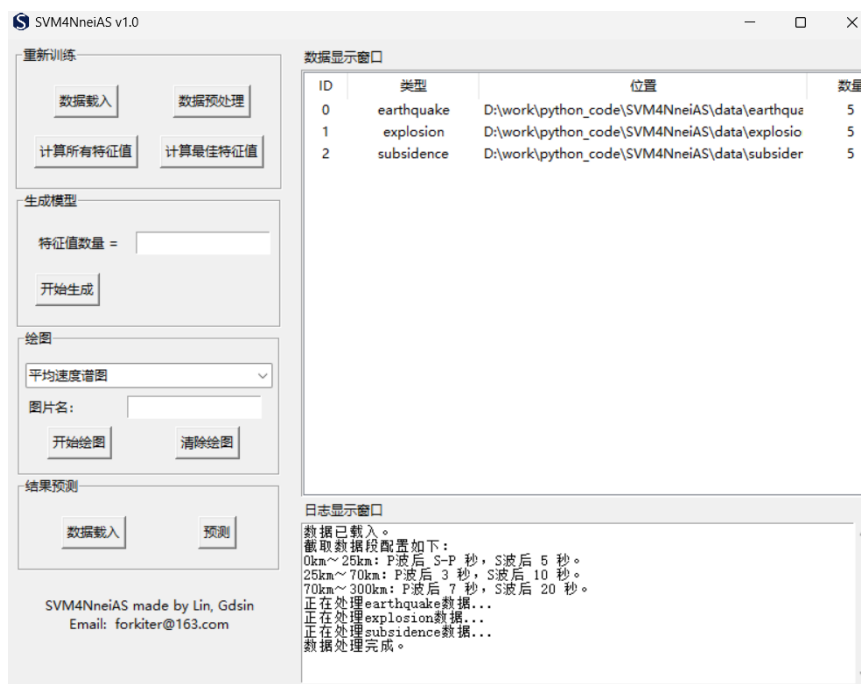
在重新训练功能区点“数据载入”按钮，可将所有 data 中的数据进行载入。“数据显示窗口”会显示所有类型的数据位置目录及数量。



2. 数据预处理

在重新训练功能区点“数据预处理”按钮，软件将自动对载入的数据进行预处理，并将处理后的数据存储在 out/sac_data 文件夹下。该数据为标准的 SAC 格式，方便用户用于其他方面的研究。

在处理过程中，日志显示窗口会显示截取数据段的配置信息以及处理过程。完成后会有“数据处理完成”的提示字样。



3. 计算所有特征值

在重新训练功能区点“计算所有特征值”按钮，软件将自动进行 0.2-15hz 范围内所有频谱特征值的计算。计算结果保存在 out/eigs 文件夹下，以 txt 格式保存，具体命名为“all_

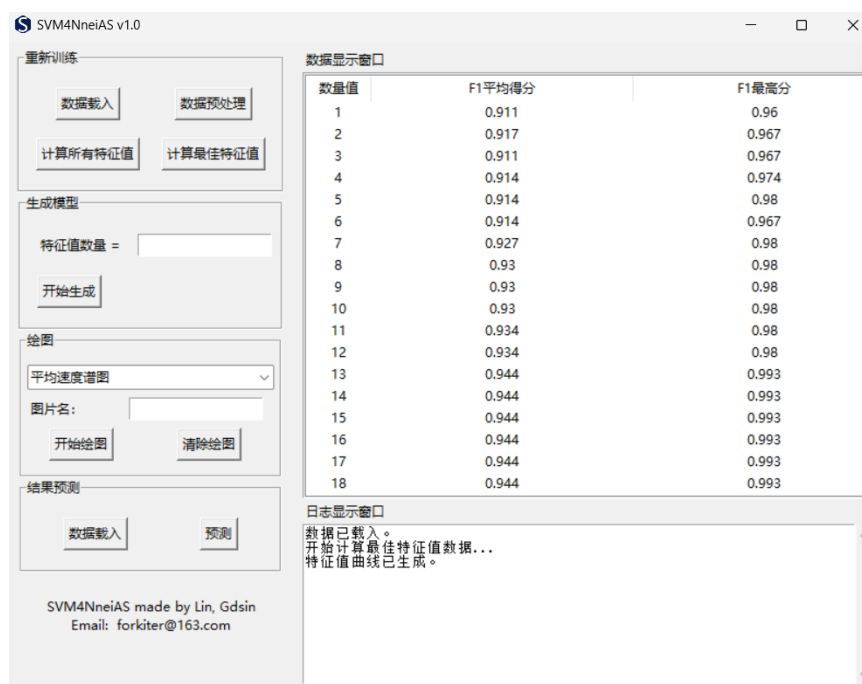
地震类型.txt”。另外，在 out 文件夹下会自动生成一个文件 “eigs_config.json”，该文件为 P 波和 S 波所有中心频率点的排序，该排序是用于下一步选取最优特征值。

在处理过程中，日志显示窗口会显示相关处理过程，完成后会有“所有特征值已计算完毕”的提示字样。

4. 计算最佳特征值

在重新训练功能区点“计算最佳特征值”按钮，软件会按照顺序穷举特征值数量，并计算每个数量值下，模型 F1 得分情况。同时，变换随机种子顺序，分析得分误差。计算结果保存在“eigs”文件夹下的“opt_value.txt”文件中。

在处理过程中，日志显示窗口会显示相关处理过程，完成后会有“特征值曲线已生成”的提示字样。同时，在数据显示窗口将显示每个数量值下 F1 平均得分和 F1 最高分。



5.1.2 生成模型

根据最佳特征值数据，选取需要的特征值数量，并在生成模型功能区下的“特征值数量”框内填入数量值，点击开始生成，生成预测模型。训练集选取 70%的数据，测试集选取 30%的数据。模型生成后会在日志显示窗口显示“模型已生成”的字样，同时在数据显示窗口，显示每种地震类型的精确率、召回率、F1 得分以及测试集数量。

在生成模型后，会在“out”文件夹下生成 3 个文件，分别为“svm_model.pkl”、“model_config.json”和“classification_report.json”。其中，“svm_model.pkl”为生成的预测模型；“model_config.json”存储的为模型参数，用于之后的地震类型预测；“classification_report.json”为测试集报告。



5.1.3 结果预测

该模块是地震类型识别模块，首先需要将地震数据 seed 文件和震相 phase 文件放入“predict_data”文件夹下。seed 文件和 phase 文件必须一一对应。在结果预测功能区点“数据载入”按钮，“数据显示窗口”会显示所有需要预测的数据文件名及文件位置。



在结果预测功能区点“预测”按钮，软件会自动判断所有需要预测地震数据的地震类型，并在“数据显示窗口”显示每个数据的预测结果，以及每种类型的预测概率值。在日志显示窗口会显示“预测结果已生成”的字样。

另外，在“predict_data”文件夹下将对应生成预测结果，文件名为“predict_result”。

txt”，结构与数据显示窗口显示的效果一致。



5.1.4 绘图

本程序提供了多种输出图件供用户选择，同时提供了多种输出图片格式。Matplotlib支持 PNG, SVG, SVGZ, PDF, PS, EPS 等栅格或矢量图格式，生成时请在“图片名”一栏直接填入图片名，并加上想要的格式后缀，可直接生成想要的图片类型，并保存在“out”文件夹下的“img”子文件夹中。如果图片名为空，软件默认生成 png 文件。

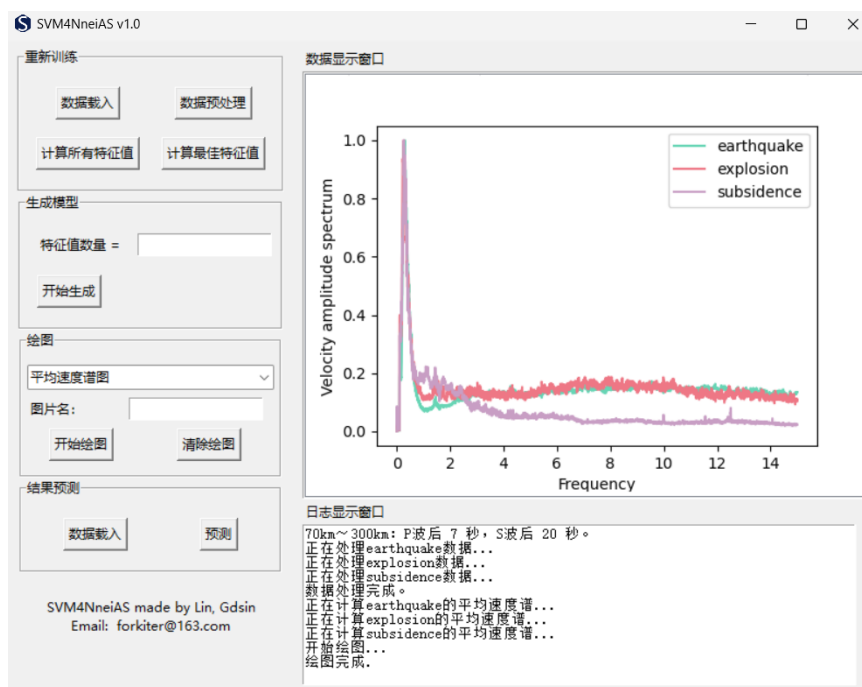


点击下拉框，选择需要输出的图件类型。点击开始绘图，在“数据显示窗口”会自动展示出所需要的图片。点击“清除绘图”按钮，将会清除展示的图片。

本程序提供输出的图件有 6 种类型：平均速度谱图，特征值曲线图，P 波频谱图，S 波频谱图，P/S 谱振幅比图以及 SVM 测试混淆矩阵图。分别介绍如下：

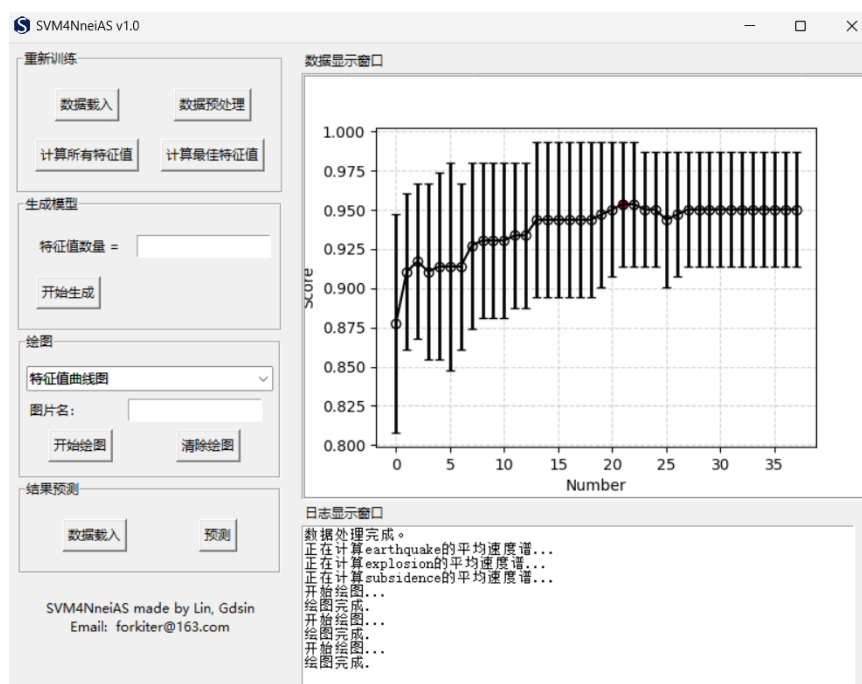
(1) 平均速度谱图

利用截取的 SAC 数据，计算不同类型的地震事件平均速度谱，并生成平均速度谱图，如下图所示。



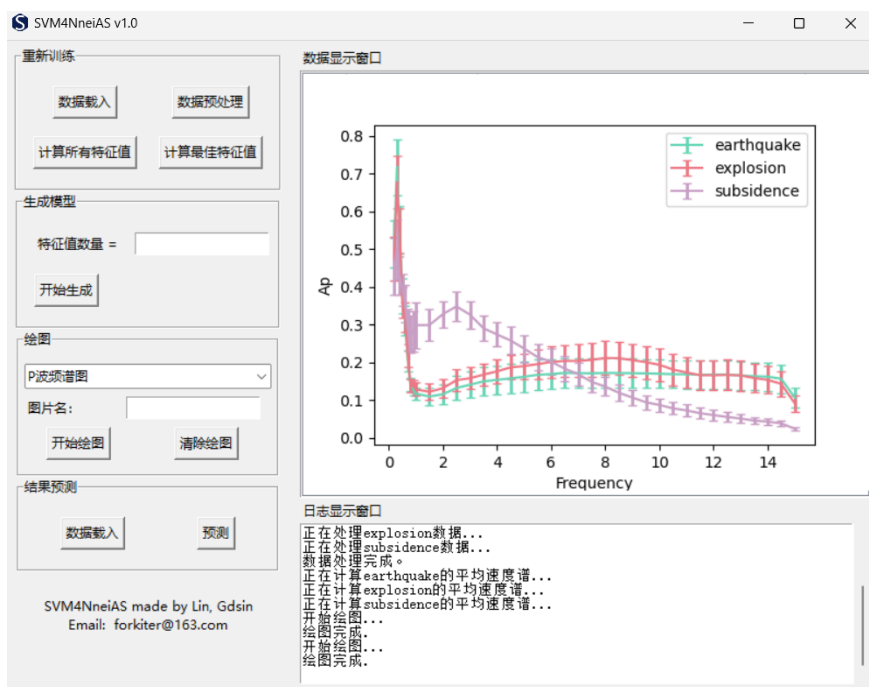
(2) 特征值曲线图

特征值曲线图是用来判断最佳特征值数量的重要依据, 其中最佳特征值数量点会用红色圆点标记出来。如下图所示。



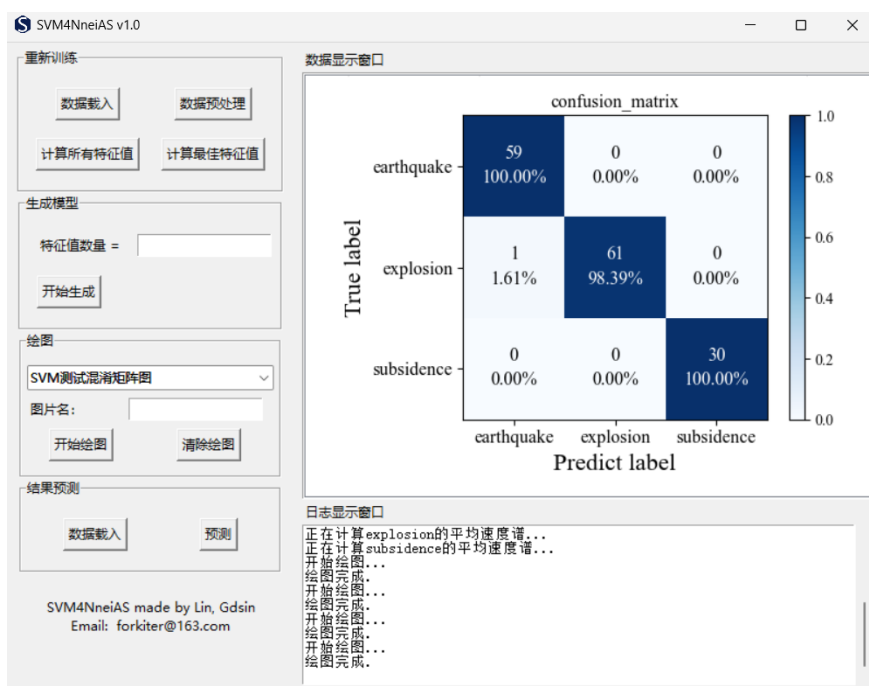
(3) P 波频谱图、S 波频谱图和 P/S 谱振幅比图

计算不同类型地震事件 P 波特征值、S 波特征值以及 P/S 对应的中心频率点的振幅谱值, 误差条是同一类地震事件特征值的标准差。如下图所示。



(4) SVM 测试混淆矩阵图

该图为模型测试数据的混淆矩阵图。左侧为真实标签，下侧为预测标签。利用混淆矩阵图可以判断模型预测效果。如下图所示。



5.2 命令行操作

该软件可通过 python 语言对数据进行批量处理，主要操作命令如下：

```
from core.nneias import NneiA, NPredict
```

```
""" 创建实例 """
nne = NneIa()
print(nne.show_data) # 查看实例是否创建成功，查看载入数据属性
""" 预处理 """
nne.data2sac() # 数据预处理
""" 计算特征值 """
nne.get_all_eigs()
nne.get_opt_eigs()
""" 生成模型 """
nne.get_svm_model()
""" 预测 """
npr = NPredict()
npr.pre()
""" 绘制平均速度谱图 """
nne.mean_fre_fig(gui=False, save_name='mean_spec.png')
""" 绘制特征值曲线图 """
nne.opt_fig(gui=False, save_name='opt_value.png')
""" 绘制 P 波、S 波、P/S 频谱图 """
nne.ps_fre_fig(save_name='p_spec.png', p_type='p', gui=False)
nne.ps_fre_fig(save_name='s_spec.png', p_type='s', gui=False)
nne.ps_fre_fig(save_name='ps_spec.png', p_type='ps', gui=False)
""" 绘制 SVM 混淆矩阵图 """
nne.cm_fig(save_name='confusion_matrix.png', gui=False)
```

用户可通过主程序目录下的 `example.py` 进行测试运行。

6. 说明

基于支持向量机的非天然地震识别算法软件由姜喜姣、林庆西、梁明、朱腾、蒋策和田平独立开发完成，目前版本号为 V1.0，在测试过程中运行良好，但不排除使用过程中出现 BUG 等问题，如发现相关问题，请联系负责人，联系方式如下：

林庆西，Email: forkiter@163.com

欢迎相关开发经验者和作者继续合作开发完善。