

MACHINE LEARNING TOOLBOX: GENERATING TRAINING DATA VECTORS FROM POINT SETS – PROGRAM aaspi_training_data

Contents

| Running aaspi_training_data | 1 |
|---------------------------------------|---|
| Unsupervised Training Data Extraction | 2 |
| Supervised Training Data Extraction | 3 |
| Output file naming convention | 4 |
| Results | 4 |

Running aaspi_training_data

For unsupervised classification, the first step is to generate a subset of input attributes via decimation (i.e. skipping while reading data at a specific interval, such as 5 inlines x 5 crosslines x 5 vertical samples). For supervised classification, after you convert polygon to point sets, the next step is to extract input attributes at those points to generate training data. Go to *Machine Learning Toolbox* menu \rightarrow generate training data (1).

| 📧 aaspi_util GUI - P | ost Stack Utilities (Relea | se Date: 25 March 2019) |
|------------------------------------|--|--|
| Eile Geometric | Attributes Spectral Att | ributes Single Trace Attributes Formation Attributes Volumetric Classification Im |
| Attribute Correlation | n Tools Display Tools | Machine Learning Toolbox Well Log Utilities Other Utilities Set AASPI Default Par |
| SEGY to AASPI format conversion | AASPI to SEGY format conversion f (multiple files) | plot and define polygons convert polygons to point sets generate training data |
| AASPI QC Plotting | - A quick tool to displa | analyze input create m Construct training data by extracting attribute vectors from point data sets |
| AASPI format inpu | t file name (*.H): | perform classification Browse |

| 🔳 aaspi_training_data (Release Date: 25 March 2019) | - | × |
|--|----|--------------|
| j <u>F</u> ile | | <u>H</u> elp |
| Tool to extract training data for either unsupervised (data decimation) or supervised classification (point sets and/or wells) The output will be multiple files corresponding to each attribute PER each point set/well and will be in the form of single trace (1D) data | | |
| List of AASPI attribute files: s\kuwahara_Eugene_energy_ratio_similarity_4_5_2x2x2.H s\kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H s\kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H s\kuwahara_Eugene_k1_2x2x2_5.H s\d_mean_Eugene_k2_2x2x2_5.H s\d_man_K2 s\d_man_K | /5 | |
| Unsupervised Point set | | |
| Specify start value, end value, and interval of data decimation for unsupervised classification | | |

Click "Browse and add to current list" (2) to select input attributes for training data extraction. These attributes should be preconditioned via kuwahara filtering in order to reduce noise and give a more realistic, geological classification, <u>unless</u> you aim to perform fine-scaled reservoir characterization using inversion products (density, impedance, v_p/v_s ratio, etc...) as input attributes.

Choose whether you want dead traces and mute zones to be ignored in extraction and whether to extract the coordinates of data points as spatial attributes (3). If the data is in the time domain, specify a constant velocity to convert time to depth. Give a unique project name and suffix (4).

Unsupervised Training Data Extraction

For unsupervised classification, click on *Unsupervised* tab (5) to define the boundary limits and decimation intervals of vertical axis, crossline, and inline. Vertical axis boundaries can be defined in 3 ways (6):

- 1. Fixed window: specify start and end values for vertical axis (7).
- 2. About a horizon: browse the first horizon (8) and specify the window size above (9) and below (10) that horizon.
- 3. Between two horizons: browse upper (8) and lower (11) horizons, then specify the window size above the upper horizon (9) and below the lower horizon (10).

AASPI support EarthVision grid format horizon file type. If your horizon format is different from the standard EarthVision grid, you need to modify the value in box (12).

Decimation intervals, crossline boundaries, and inline boundaries can be defined in box (13). After all parameters are set, click (14) to extract training data for unsupervised classification.

| Unsupervised Point | |
|--|---------------------------|
| Specify start v successful the start of data decimation for unsupervision of the start of the st | |
| Temporal Operation Window | |
| | Help - Horizon Definition |
| Fixed time window?: | |
| Compute about and between two horizons?: | |
| Compute about a single flattened horizon?: | |
| Start Time in s: 0.2 | |
| End Time in s: 1.4 | |
| Input shallower horizon filename: |] |
| (Choose horizon type below:) View horizon file Convert DOS to Uni | ix |
| Window start wrt shallower horizon in s -0.1 | |
| Input deeper horizon filename: | |
| (Choose horizon type below:) View horizon file Convert DOS to Uni | x |
| Window start wrt deeper horizon in s (vertical axis positive down) | |
| Choose horizon type: gridded (e.g. EarthVision) | |
| Number of header lines to skip: 0 | |
| Total number of columns: 5 | |
| Column number of line_no: 1 | |
| Column number of cdp_no: | |
| Column number of 5 time or depth picks: | |
| znull value (indicates missing pick): -9999999 | |
| Vertical axis of picked surface? Positive Down | |
| Vertical Units of ms Picked Horizons: | |
| | |
| CDP no. Start: 200 13 | |
| CDP no. End: 500 | |
| CDP no. Interval: 10 | |
| Line no. Start: 250 | |
| Line no. End: 350 | |
| Line no. Interval: 10 | |
| Extract unsupervised training data | |

Supervised Training Data Extraction

For supervised classification, click on *Point set* tab (15). Browse to the AASPI-formatted point set files (16), which were converted from polygons in the previous step of AASPI machine learning workflow. Box (17) displays detected facies names from the browsed point sets. This list is NOT editable and is for informational purpose only. However, the order of facies names can be changed by "*Move selected facies UP/Down*" buttons (18)

After browsing point sets, click "Extract training data from point sets" (19).



Output file naming convention

Program aaspi_training_data will always generate the following output files:

| Output file description | File name syntax |
|--------------------------|--|
| program log information | training_data_unique_project_name_suffix.log |
| program error/completion | |
| information | training_data_unique_project_name_suffix.err |

where the values in red are defined by the program GUI. The errors we anticipated will be written to the **.err* file and be displayed in a pop-up window upon program termination. These errors, much of the input information, a description of intermediate variables, and any software traceback errors will be contained in the **.log* file.

Results

Unsupervised training data files will have the following naming scheme: training_data_unsupervised_<attribute_file_name>_<unique_project_name>_<suffix>.H training_data_unsupervised_<attribute_file_name>_<unique_project_name>_<suffix>.H@

Supervised training data files will have the following naming scheme: training_data_<point_set_file_name>_<attribute_file_name>.H training_data_<point_set_file_name>_<attribute_file_name>.H@

Below is a sample of a .H unsupervised training data file:

| 1 | |
|-----|---|
| 2 | training_data_unsupervised.exe: Thang@ Mon May 13 01:33:54 2019 |
| 3 | <pre>in="training_data_unsupervised_amp_pos_amp_pos_0.H@"</pre> |
| 4 | unit="" |
| 5 | attribute_fn="amp_pos.H" |
| 6 | limit_type=0 |
| 7 | upper_horizon_fn="" |
| 8 | lower_horizon_fn="" |
| 9 | horizon_type="gridded" |
| 10 | t_start=0.2 |
| 11 | t_end=1.4 |
| 12 | relative_t_start=0 |
| 13 | relative_t_end=0 |
| 14 | skipl=0.04 |
| 15 | min2=300 |
| 16 | max2=500 |
| 17 | skip2=10 |
| 18 | min3=250 |
| 19 | max3=350 |
| 20 | skip3=10 |
| 21 | unique_project_name="amp_pos" |
| 22 | suffix="0" |
| 23 | nl=7161 ol=1 dl=1 labell="Sample" |
| 0.0 | |

The image below is an example of supervised training data file:

🔚 training_data_thang_not_salt_line_451_1_Eugene_5x5x5_d_mean_Eugene_k1_2x2x2_5.H 🗵

```
training_data_point_set.exe:
                                   Thang@
                                             Tue Mar 26 22:57:30 2019
3
             nl=1536 ol=1 dl=1 label1="Sample"
             in="training_data_thang_not_salt_line_451_1_Eugene_5x5x5_d_mean_Eugene_k1_2x2x2_5.H0"
4
            unit=""
5
            point_set_fn="thang_not_salt_line_451_1_Eugene_5x5x5.pnt"
6
7
             attribute_fn="d_mean_Eugene_kl_2x2x2_5.H"
8
            facies_list_fn="training_data_point_set_facies_list_Eugene_5x5x5.txt"
9
            facies_name="Not_Salt"
             facies no=1
            zero_based="n"
             unique_project_name="Eugene"
             suffix="5x5x5"
```

In addition, a list of input attributes for training data extraction is saved to: training_data_attribute_list_<unique_project_name>_<suffix>.txt

🔚 training_data_attribute_list_Eugene_5x5x5.txt 🗵

| 1 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\kuwahara_Eugene_energy_ratio_similarity 4_5_2x2x2.H |
|---|---|
| 2 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\kuwahara_Eugene_glcm_contrast_4_5_2x2x2.H |
| 3 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H |
| 4 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\kuwahara_Eugene_total_energy_4_5_2x2x2.H |
| 5 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\d_mean_Eugene_k1_2x2x2_5.H |
| 6 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\d_mean_Eugene_k2_2x2x2_5.H |
| | |

For unsupervised classification, a list of all extracted training data files is provided in training_data_unsupervised_extracted_list_<unique_project_name>_<suffix>.txt

```
training_data_unsupervised_extracted_list_amp_pos_0.tt X

training_data_unsupervised_amp_pos_amp_pos_0.H
training_data_unsupervised_conf_amp_pos_0_amp_pos_0.H
training_data_unsupervised_crossline_dip_amp_pos_0_amp_pos_0.H
```

For supervised classification, a list of all extracted training data files is provided in training_data_point_set_extracted_list_<unique_project_name>_<suffix>.txt Point set list and detected facies list are also saved to text files: training_data_point_set_list_<unique_project_name>_<suffix>.txt training_data_point_set_facies_list_<unique_project_name>_<suffix>.txt

| 🔚 trair | ing_data_point_set_extracted_list_Eugene_5x5x5.txt 🔀 |
|----------|---|
| 1 | training data thang not salt line 451 l Eugene 5x5x5 kuwahara Eugene energy ratio similarity 4 5 2x2x2.H |
| 2 | training_data_thang_not_salt_line_451_1_Eugene_5x5x5_kuwahara_Eugene_glcm_contrast_4_5_2x2x2.H |
| 3 | training_data_thang_not_salt_line_451_1_Eugene_5x5x5_kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H |
| 4 | training_data_thang_not_salt_line_451_1_Eugene_5x5x5_kuwahara_Eugene_total_energy_4_5_2x2x2.H |
| 5 | training_data_thang_not_salt_line_451_1_Eugene_5x5x5_d_mean_Eugene_k1_2x2x2_5.H |
| 6 | training_data_thang_not_salt_line_451_1_Eugene_5x5x5_d_mean_Eugene_k2_2x2x2_5.H |
| 7 | training_data_thang_not_salt_line_451_2_Eugene_5x5x5_kuwahara_Eugene_energy_ratio_similarity_4_5_2x2x2.H |
| 8 | training_data_thang_not_salt_line_451_2_Eugene_5x5x5_kuwahara_Eugene_glcm_contrast_4_5_2x2x2.H |
| 9 | training_data_thang_not_salt_line_451_2_Eugene_5x5x5_kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H |
| 10 | <pre>training_data_thang_not_salt_line_451_2_Eugene_5x5x5_kuwahara_Eugene_total_energy_4_5_2x2x2.H</pre> |
| 11 | training_data_thang_not_salt_line_451_2_Eugene_5x5x5_d_mean_Eugene_k1_2x2x2_5.H |
| 12 | training_data_thang_not_salt_line_451_2_Eugene_5x5x5_d_mean_Eugene_k2_2x2x2_5.H |
| 13 | <pre>training_data_thang_not_salt_line_521_1_Eugene_5x5x5_kuwahara_Eugene_energy_ratio_similarity_4_5_2x2x2.H</pre> |
| 14 | <pre>training_data_thang_not_salt_line_521_1_Eugene_5x5x5_kuwahara_Eugene_glcm_contrast_4_5_2x2x2.H</pre> |
| 15 | training_data_thang_not_salt_line_521_1_Eugene_5x5x5_kuwahara_Eugene_glcm_dissimilarity_4_5_2x2x2.H |
| 16 | training_data_thang_not_salt_line_521_1_Eugene_5x5x5_kuwahara_Eugene_total_energy_4_5_2x2x2.H |
| 🗎 train | na data point set list Europe 5x5x5txt 🔀 |
| 1 | b)/ASDI CIT dit/aparts tosting ppp facion/thang not calt line 451 1 Fugane SuFME ppt |
| 2 | p. (Amoriger, gir) adopt besting pum factes (thang not salt line 451 _ Eugene_SXSS, put |
| 3 | D:\AASFI GIT.git\aaspi testing pun facies\thang not salt line 521 1 Eugene 5x5x5.nnt |
| 4 | D:\AASPI GIT.git\aaspi testing pnn facies\thang not salt line 521 2 Eugene 5x5x5.pnt |
| 5 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\thang_not_salt_slice_2000_Eugene_5x5x5.pnt |
| 6 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\thang_salt_line_451_Eugene_5x5x5.pnt |
| 7 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\thang_salt_line_521_Eugene_5x5x5.pnt |
| 8 | D:\AASPI_GIT.git\aaspi_testing_pnn_facies\thang_salt_slice_2000_Eugene_5x5x5.pnt |
| - traini | and data point set facient let Europea ByEvEtet [2] |
| | |
| 2 | pour Jaac |
| | |

These training data files and text-based lists will be used to analyze input and create models for AASPI machine learning workflow in subsequent steps.