Active Learning for Sibert Study

Lawrence Carin, Levi Kennedy and Xianyang Zhu

Signal Innovations Group, Inc.
Durham, NC

Sponsored by ESTCP
Digital geophysics algorithms require labeled data for training.

- When performing UXO cleanup one may partition a portion of the site for defining training data.
- This portion of site excavated first, with labels revealed for subsequent training.
- No guarantee that such labeled data are representative of the site.
Information-Theoretic Based Learning

- Collect EMI/MAG data and perform geophysical model fitting as is “standard”
- Use information-theoretic measures to define which data would be most informative if the associated labels could be acquired
- These items are excavated first, for the purpose of in situ learning
- Information theory also defines when no more labels are required
- Number of labels required typically small
After acquiring labeled data as above, we may design a classifier. This classifier may then be used to quantify the probability that any of the remaining unexcavated items are UXO. Ordered dig list provided. By construction the labeled data used for classifier design matched to the site under test.
Feature Extraction

- The classifier processes “features” extracted from the EMI and/or magnetometer data.
- We employ dipole-based models for representation of the measured data, and the model parameters fit to the data are employed as features.
- For magnetometer data the features are the dipole strength and the “goodness of fit” between the measured and modeled data.
- For the EMI data the features are in terms of multiple dipole components:
  \[ M = \text{diag}(\beta_1, \beta_2, \beta_3) \]
Classifier Construction

- Assume the real vector $v$ represents the features extracted from the model fit to the measured data.
- Data associated with UXO arbitrarily are assigned label $l = 1$, and non-UXO are labeled $l = 0$.
- The probability that any particular $v$ is associated with UXO is expressed as:

$$p(l = 1 | v, w) = \frac{\exp[y(v; w)]}{1 + \exp[y(v; w)]}$$

$$y(v; w) = \sum_{i=1}^{N_b} w_i K(v, b_i) + w_0$$

- Based on labeled data, wish to infer classifier weights $w$. 
**In Situ** Learning

\[
p(l = 1 | \mathbf{v}, \mathbf{w}) = \frac{\exp[y(\mathbf{v}; \mathbf{w})]}{1 + \exp[y(\mathbf{v}; \mathbf{w})]} \]

\[
y(\mathbf{v}; \mathbf{w}) = \sum_{i=1}^{N_b} w_i K(\mathbf{v}, \mathbf{b}_i) + w_0
\]

- The Fisher information measure is used to quantify which features should be used as basis functions \( \{ \mathbf{b}_i \}_{i=1,N_b} \), also determining \( N_b \)
- Similar measure employed to define which feature vectors \( \mathbf{v}_n \) are most informative for inferring the classifier weights \( \mathbf{w} = \{ w_i \}_{i=1,N_b} \)
- The information-theoretic measures have rigorous mathematical constructions, but their final form is simple
Advantages

- Typically the number of required labeled data (excavation for the purpose of learning) is small

- Acquire the *right* labeled data for the site under test, and typically much less excavation required for acquisition of training data

- Principled means of defining labeled data required for digital geophysics
EMI and magnetometer data collected at the former Camp Sibert, Alabama

Relatively benign test, in that only one relatively large ordnance type of interest

The analysts were evaluated in a blind test, with good performance manifested

Here we re-analyze the data, considering all items (removing none of the items from the analysis)
Data Characteristics

- Results presented for EM61 and magnetometer
- Example EM61 data and model fit:

![Measured EM Data (Sibert)](image1)

![Model Fit](image2)
EM61 Features

EM61 Feature Plots for UXO (blue dots) vs Clutter (red circles)

Lawrence Carin, Levi Kennedy and Xianyang Zhu

Active Learning for Sibert Study
Magnetometer Features
In Situ Learning for EM61

- Basis vector selection
- Of 20 basis vectors, 16 non-UXO and 4 UXO

Stop Here
EM61 Results

Trained via active learning

Trained with data provided by ESTCP

Number of Training Data: 25

Number of Training Data: 134
Joint Processing of EM61 & MAG

![ROC Curve](chart1.png)

- ROC Curve
- C=25
- C=50
- C=100

![Chart 2](chart2.png)

- $P_{disc}$
- $FP_{disc}$

Lawrence Carin, Levi Kennedy and Xianyang Zhu

Active Learning for Sibert Study
Conclusions

▶ Advanced digital geophysics taken to field successfully

▶ Information-theoretic based *in situ* learning provides a principled and effective means of defining labeled data

▶ Labeled data by construction matched to site under test

▶ Moving now to more-challenging sites for demonstration

Research Supported by ESTCP