MACHINE LEARNING AT THE CTBTO. TESTING, AND EVALUATION OF THE FALSE EVENTS IDENTIFICATION (FEI) AND VERTICALLY INTEGRATED SEISMIC ANALYSIS (VISA) PROJECTS

Ronan J. Le Bras[1], Stuart Russell[2], Nimar Arora[2], and Vera Miljanovic[1]

Comprehensive Nuclear-Test-Ban Treaty Organization[1] and the University of California at Berkeley[2]

ABSTRACT

Since 2009, an initiative to investigate the potential of machine learning methods to improve automatic data processing at the CTBTO and in particular the recall and accuracy of the automatic bulletins is starting to bear fruit beyond the stage of research and has entered the domain of development and testing with the goal of operational testing for one of the projects (FEI) by the end of 2011. The prospect for FEI is that the tool will comfort analysts in their decision-making process when they make decisions on whether a (mostly smaller) event is real or false, and it is thus an enhancement of the current analysis system. The VISA projects are more ambitious and aim at replacing key components of the processing system. The prototype of the first generation, which aims at replacing the current automatic association tool (GA), is being evaluated on the vDEC collaborative platform of the CTBTO. Results show much improved accuracy using VISA as compared to the SEL3 for the same recall value, or much-improved recall value using VISA as compared to the SEL3 for the same processing accuracy. A consequence is a significant decrease in either the number of false alarms or the number of missed events, depending on the setting of the processing parameters.

FALSE EVENT IDENTIFICATION Results

Testing Method

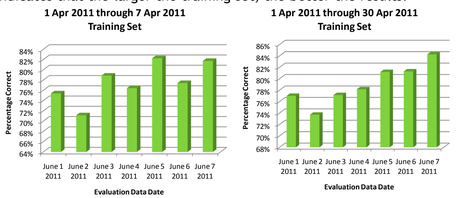
- Four FEI training sets were created and were used to independently evaluate one week in 2011

Start Date	End Date	SEL3 / LEB Events
February 24, 2006	March 05, 2006	1258 / 1161
April 1, 2011	April 7, 2011	1872 / 1504
April 1, 2011	April 30, 2011	4511 / 3357
April 1, 2011	May 31, 2011	13249 / 10639

Training Sets Created for FEI Testing

Results

When the data used for training and evaluations are closely spaced in time, FEI gives very good results: more than 80% of the FEI classifications are correct. Testing of the various classifier sets listed in the above table indicates that the larger the training set, the better the results.



FEI results using two training data sets of different size on the same evaluation data set (June 1-7, 2011)

FALSE EVENT IDENTIFICATION Results (Continued)

When the training and classification sets are from temporally separated time periods, a 25-30% degradation of performance was observed, regardless of the size of the training sets. This degradation is attributable to variance in the composition of the network. When a new station is added to the network, the dynamics of event formation change. Likewise, if a station exists in the training set, but is not in the network of the evaluated data, a similar degradation is observed.

Using a trained classifier (with 2006 or 2011 data) against varying (2006—2011) network composition

- FEI does an excellent job at classifying/categorizing automatic events into either false events or events with a high probability of being real.

- The results presented to analysts as the process currently stands should add confidence to their decisions and help identify obviously wrong associations and false events.

- The next steps on the project include:
  - ➢ Non-intrusive integration (analysts will not see the results) within the operation pipeline, establishment of an operational model based on the above observation that the parameters based on the training set need to be upgraded periodically.
  - ➢ Design of an analyst interface such that the analysts are presented with the FEI results.

NET-VISA Results

Dataset and Evaluation Criteria

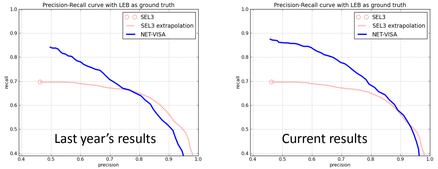
- NET-VISA is trained on 2.5 months of data and evaluated on one week.
- For the purpose of the evaluation, LEB is considered the ground truth.
- The predicted events are matched to a ground truth event within 5 degrees and 50 seconds.
- From the matching, we compute:
  - Recall: Percentage of ground truth events that matched a predicted event.
  - Precision: Percentage of predicted events that matched a ground truth event.
  - Error: Average distance between predicted and matching ground truth events.

m_b range	Number of Events	SEL3		NET-VISA	
		Recall (%)	Error (km)	Recall (%)	Error (km)
0-2	74	64.9	101	86.5	101
2-3	36	50.0	186	77.8	159
3-4	558	66.5	104	86.4	115
>4	164	86.6	70	93.3	78

Recent Improvements

- Associating the ***tx*** phase, which is normally ignored as noise or coda detection, but often contains a real P phase.
- Improved noise model.

Precision Recall Curve



SEL3 extrapolation results used FEI

NET-VISA Results (Continued)

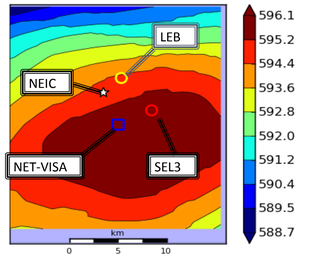
Evaluation with Regional Bulletins as Ground Truth

Since LEB is not complete at lower magnitudes, NET-VISA was also evaluated against regional bulletins, which use many more stations than in the IMS network.

- In the continental United States of 33 events reported by NEIC:
  - LEB got 4 correct out of 4 predicted events
  - NET-VISA got 7 correct out of 35 predicted events
- In Japan out of 1565 events reported by JMA:
  - LEB got 29 correct out of 29 predictions
  - NET-VISA got 33 correct out of 52 predictions
- In Europe out of 101 events reported by PRU
  - LEB got 5 correct out of 10 predictions
  - NET-VISA got 11 correct out of 43 predictions
- In Central Asia out of 101 events reported by NNC
  - LEB got 35 correct out of 74 predictions
  - NET-VISA got 50 correct out of 166 predictions

Evaluation on DPRK Nuclear Explosion (25 May 2009)

In this experiment, NET-VISA was trained on 1 year's worth of data (April 1, 2008 to April 1, 2009)



White star - NEIC, Yellow circle - LEB, Red circle - SEL3, Blue square - NET-VISA.

- NET-VISA event location agreed well with NEIC and LEB (roughly 5 km).
- NET-VISA associated the event with detections in 53 stations versus 39 stations by SEL3.

CONCLUSIONS

- The FEI program has been successfully evaluated on the vDEC platform and is close to operational implementation.
- NET-VISA reduces detection failures by more than a factor of 2 compared with SEL3.
- NET-VISA correctly located the nuclear test conducted by the Democratic People's Republic of Korea on 25th May 2009. This is a verification that the complete prior model, which includes a uniform spatial distribution in addition to the seismicity-dominated prior, is adequate to detect events which do not occur in areas of previous seismicity.
- NET-VISA is currently being tested in the CTBTO vDEC environment for possible deployment in operations.
- The next step in terms of algorithmic development is to develop the SIG-VISA prototype with an extension of the generative model down to waveform level, and include the step of signal detection within the generative model.

REFERENCES

Arora, Nimar S., Stuart J. Russell, Paul Kidwell, and Erik Sudderth, "Global seismic monitoring: A Bayesian approach.", In *Proc. AAAI-11*, San Francisco, 2011.

Arora, Nimar S., Stuart J. Russell, Paul Kidwell, and Erik Sudderth, "Global seismic monitoring as probabilistic inference.", In *Advances in Neural Information Processing Systems 23*, MIT Press, 2011.

Arora, Nimar S.., Russell, S., Kidwell, P., and Sudderth, E., "NET-VISA model and inference improvements", *CTBT:S&T 2011 Conference*, Vienna, Austria, June 8-10, 2011.

Mackey, L., Kleiner, A., Jordan, M. I., 2009, Improved Automated Seismic Event Extraction Using Machine Learning, *American Geophysical Union, Fall Meeting 2009*, abstract #S31B-1714.

Tuma, M. and Igel, C., 2009, Kernel-based machine learning techniques for hydroacoustic signal classification, *CTBTO ISS09 Conference*.

Le Bras, Ronan J., Sheila Vaidya, Jeffrey Schneider, Stuart Russell, and Nimar Arora, "Status of the Machine Learning Efforts at the International Data Centre of the CTBTO." In *Proc. Monitoring Research Review (MRR 2010)*, Orlando, Florida, 2010.

Russell, Stuart, Sheila Vaidya, and Ronan Le Bras, "Machine Learning for Comprehensive Nuclear-Test-Ban Treaty Monitoring.", *CTBTO Spectrum*, 14, 32-35, 2010.

Vaidya, S., Robert Engdahl, Ronan Le Bras, Karl Koch, and Ola Dahlman, 2009, Strategic Initiative in Support of CTBT Data Processing: vDEC (virtual Data Exploitation Centre), *CTBTO ISS09 Conference*

International Seismological Centre, *On-line Bulletin*, http://www.isc.ac.uk, Internatl. Seis. Cent., Thatcham, United Kingdom,

Acknowledgement

We thank Dr. Lassina Zerbo, IDC Director, for allowing us to publish this research and for his support of the machine learning efforts at the IDC, Ronald "Chip" Brogan and Misrak Fisseha for their analyst expertise and efforts in evaluating the FEI software delivered at the IDC.