ESS2222

Lecture 10 – Machine Learning in Earth Sciences

Hosein Shahnas

University of Toronto, Department of Earth Sciences,
- Land cover classification
- Geological mapping
- Fast magnitude determination of the seismic events
- Climate change and sea level rise
- Weather forecasting
- Inverse problems in the planetary mantle
- The thermal state of the planetary mantle
Review of Lecture 9

\( \text{gini}(D) = 1 - \sum_{c=1}^{C} P_c^2 \)

\( \Delta \text{gini}(S) = \text{gini}(D) - \text{gini}(S) \)
Land cover classification and change analysis of the Twin Cities (Minnesota)

The accurate and timely information describing the nature and extent of land resources and changes over time is important, especially in rapidly growing metropolitan areas.

Fig. 4. Twin Cities Metropolitan Area urban growth from 1986 to 2002 with 2000 MUSA boundary. Rural land cover (agriculture, forest and wetland) that was converted to urban from 1986 to 1991, from 1991 to 1998, and from 1998 to 2002 are highlighted in green, red and yellow, respectively.
Geological mapping

a) Remotely sensed spectral imagery  
b) Geophysical (magnetic and gravity) data  
c) Geodetic (elevation) data  
are useful in a number of Earth science applications such as environmental monitoring and mineral exploration.

Remotely sensed imagery has many applications in Earth science applications such as **environmental monitoring** (Munyati, 2000), **land coverage studies** (Yuan et al., 2005), and **mineral exploration** (Hewson et al., 2006; Sabins, 1999).  
So improving the techniques used in exploration and lithological identification is important for understanding of regional geology.

**As the volume of data increase:**  
1) Manual interpretation cannot maintain the pace with the amount of incoming data and  
2) Manual image interpretation is generally subjective and can be inconsistent among interpreters

**Machine learning techniques** can be employed in geological mapping and interpretation (Harvey & Fotopoulos, 2016) as a rapid approach of geological mapping in contrast to conventional field expedition techniques.
Geological Mapping Using Machine Learning Algorithms

Harvey and Fotopoulos (2016)

Figure 1. Map showing major stratigraphy groups and other major units in the Sudbury region (Ontario Geological Survey, 2011).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Source and Filename</th>
<th>Units</th>
<th>Original Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat 4-5 TM</td>
<td>USGS LT50190282011278EDC00</td>
<td>Spectral Response</td>
<td>30 m × 30 m</td>
</tr>
<tr>
<td>Bands 1-7</td>
<td></td>
<td>16-bit data</td>
<td></td>
</tr>
<tr>
<td>October 2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Elevation Model</td>
<td>SRTM n46_w081_larc_v3</td>
<td>metres</td>
<td>30 m × 30 m</td>
</tr>
<tr>
<td></td>
<td>MNDM ONMAGONL from GDS1036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Magnetic Intensity</td>
<td>OGS; MNDM ONGRAVITY1</td>
<td>nanoTesla</td>
<td>200 m × 200 m</td>
</tr>
<tr>
<td>Bouguer Gravity Anomaly</td>
<td>OGS; MNDM ONGRAVITY1</td>
<td>milliGal</td>
<td>1000 m × 1000 m</td>
</tr>
<tr>
<td>Bedrock Geology</td>
<td>OGS Geopoly from MRD126-REV1</td>
<td>Discrete Geological Units</td>
<td>Resampled to study area density</td>
</tr>
</tbody>
</table>

Table 1. Summary of data, features for classification and validation, and class label inputs. Includes source, units, and original resolution.
Geological Mapping Using Machine Learning Algorithms

All the inputs features:

- Total magnetic intensity
- Elevation
- Gravity
- Spectral images

are used to create a digital signature for each rock-type.

Figure 2. Rocktype map of the Sudbury Basin and surrounding area. Refer to Table 3 for legend, rocktype descriptions, and proportions within the study area (Ontario Geological Survey, 2011).
<table>
<thead>
<tr>
<th>Legend</th>
<th>% Cover</th>
<th>Rocktype Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.11</td>
<td>Amphibolite, gabbro, diorite, mafic gneisses</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
<td>Basaltic and andesitic flows, tuffs and breccias, chert, iron formation, minor metasedimentary and intrusive rocks</td>
</tr>
<tr>
<td></td>
<td>7.07</td>
<td>Carbonaceous slate</td>
</tr>
<tr>
<td></td>
<td>0.08</td>
<td>Commonly layered biotite gneisses and migmatites; locally includes quartzofeldspathic gneisses, ortho- and paragneisses</td>
</tr>
<tr>
<td></td>
<td>0.44</td>
<td>Conglomerate, sandstone, siltstone, argillite</td>
</tr>
<tr>
<td></td>
<td>0.22</td>
<td>Diorite, quartz diorite, minor tonalite, monzonite, granodiorite, syenite and hypabyssal equivalents</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>Gabbro, anorthosite, ultramafic rocks</td>
</tr>
<tr>
<td></td>
<td>0.82</td>
<td>Granite, alkali granite, granodiorite, quartz feldspar porphyry; minor related volcanic rocks (1.5 to 1.6 Ga)</td>
</tr>
<tr>
<td></td>
<td>13.54</td>
<td>Granophyre</td>
</tr>
<tr>
<td></td>
<td>18.53</td>
<td>Lapilli tuff, breccia, felsic flows and intrusions, minor carbonate and cherty</td>
</tr>
<tr>
<td></td>
<td>2.72</td>
<td>Mafic, intermediate and felsic metavolcanic rocks, intercalated metasedimentary rocks and epiclastic rocks</td>
</tr>
<tr>
<td></td>
<td>10.80</td>
<td>Massive to foliated granodiorite to granite</td>
</tr>
<tr>
<td></td>
<td>0.33</td>
<td>Murray Granite 2388 Ma, Creighton Granite 2333 Ma: granite</td>
</tr>
<tr>
<td></td>
<td>1.64</td>
<td>Nipissing mafic sills (2219 Ma): mafic sills, mafic dikes and related granophyre</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td>Norite, gabbro, granophyre</td>
</tr>
<tr>
<td></td>
<td>7.79</td>
<td>Norite-gabbro, quartz norite, sublayer and offset rocks</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
<td>Quartz sandstone, minor conglomerate, siltstone</td>
</tr>
<tr>
<td></td>
<td>3.50</td>
<td>Quartz-feldspar sandstone, argillite and conglomerate</td>
</tr>
<tr>
<td></td>
<td>0.38</td>
<td>Quartz-feldspar sandstone, sandstone with minor siltstone, calcareous siltstone and conglomerate</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>Rhyolite, rhyodacitic, dacitic and andesitic flows, tuffs and breccias, chert iron formation, minor metasediments and intrusive rocks</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>Sandstone, siltstone, conglomerate, limestone, dolostone</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
<td>Siltstone, argillite, sandstone, conglomerate</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>Siltstone, argillite, wacke, minor sandstone</td>
</tr>
<tr>
<td></td>
<td>2.33</td>
<td>Siltstone, wacke, argillite</td>
</tr>
<tr>
<td></td>
<td>10.70</td>
<td>Tonalite to granodiorite-foliated to gneissic-with minor supracrustal inclusions</td>
</tr>
<tr>
<td></td>
<td>10.40</td>
<td>Tonalite to granodiorite-foliated to massive</td>
</tr>
<tr>
<td></td>
<td>6.67</td>
<td>Wacke, minor siltstone</td>
</tr>
</tbody>
</table>

Table 3. Legend and rock type descriptions for Figure 2. Includes % of how much of the study area each rock type covers. Adapted from Ontario Geological Survey (2011).
Fast magnitude determination using a single seismological station record implementing machine learning techniques

Luis H. Ochoa*, Luis F. Niño, Carlos A. Vargas

*Corresponding author.

Universidad Nacional de Colombia, Carrera 45 No. 26-85 Edificio Manuel Ancizar - Of: 330, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 11 January 2017
Accepted 22 March 2017
Available online 1 April 2017

Keywords:
Earthquake early warning
Support Vector Machine Regression
Earthquake
Rapid response
Local magnitude
Seismic event
Seismology
Bogota
Colombia

ABSTRACT

In this work a Support Vector Machine Regression (SVMR) algorithm is used to calculate local magnitude (ML) using only five seconds of signal after the P wave onset of one three component seismic station. This algorithm was trained with 863 records of historical earthquakes, where the input regression parameters were an exponential function of the waveform envelope estimated by least squares and the maximum value of the observed waveform for each component in a single station. Ten-fold cross validation was applied for a normalized polynomial kernel obtaining the mean absolute error for different exponents and complexity parameters. The local magnitude (ML) could be estimated with 0.19 units of mean absolute error. The proposed algorithm is easy to implement in hardware and may be used directly after the field seismological sensor to generate fast decisions at seismological control centers, increasing the possibility of having an effective reaction.

© 2017 Institute of Seismology, China Earthquake Administration, etc. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Early warning systems employ dense seismological networks to localize and determine the magnitude of the earthquake using at least 3 stations.

The problem with this method: The density of stations in some high seismic risk areas is not enough to make such localization calculations fast.

An alternative solution: The seismological records of previous events recorded at one single station can be used to localize and estimate the magnitude of the event.
Early Earthquake Warning System

Implementation of an early earthquake warning system for the city of Bogota, Colombia

Seismic early warning systems (SEWS) emit an alert, few seconds after the event initiates, from few seconds to a few tens of seconds before the stronger shaking movement arrives.

The main task:
Estimation of magnitude and source location of an earthquake in a short period of time accurately

863 records of historical earthquakes are used in training of a Support Vector Machine Regression (SVMR)–model to calculate (predict) local magnitude (Ml) using only five seconds of signal after the P wave onset of one three component seismic station.

SVMR–model
10-fold cross validation
Polynomial kernel
Early Earthquake Warning System

Historical data:
The seismic catalogue with 2164 seismic events, selected between Jan. 1\textsuperscript{st}, 1998 and Oct. 27\textsuperscript{th}, 2008, located at less than 120 km from the seismic station.

The parameters for Gutenberg-Richter relation change from one seismic region to another which should be taken into account in the training machine leaning models for different regions.

Fig. 3. Local magnitude statistical distribution.

\[ \ln (N) = -2.65 M_l + 12.62 \]
\[ R^2 = 0.99 \]

Fig. 4. Gutenberg–Richter relation for selected seismic events.
Early Earthquake Warning System

SVM-Regression Model

**Tainting data:**
863 historical data from seismic events (2164 events filtered to exclude Ml ≤ 2) and anomalous values

**Noise filtering:**
High-pass filter with a cut-off frequency of 0.075 Hz
Low-pass filter with a cut-off frequency of 150 Hz.

Three-component raw waveforms recorded directly at seismic station.
Climate Change and Sea Level Rise

Features

(Alahmadi & Kolmas, 2015)

Figure 2: Prediction of sea level rise in San Francisco.

Figure 3: Prediction of global sea level rise.
Weather Forecasting using Incremental K-means Clustering

Unsupervised Learning

Table I. Original air-pollution Database

<table>
<thead>
<tr>
<th>Date</th>
<th>CO$_2$</th>
<th>RPM</th>
<th>SO$_2$</th>
<th>NO$_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1/2009</td>
<td>85</td>
<td>183</td>
<td>12</td>
<td>95</td>
</tr>
<tr>
<td>2/1/2009</td>
<td>95</td>
<td>289</td>
<td>14</td>
<td>125</td>
</tr>
<tr>
<td>3/1/2009</td>
<td>112</td>
<td>221</td>
<td>10</td>
<td>101</td>
</tr>
<tr>
<td>4/1/2009</td>
<td>114</td>
<td>191</td>
<td>11</td>
<td>97</td>
</tr>
<tr>
<td>5/1/2009</td>
<td>100</td>
<td>175</td>
<td>11</td>
<td>101</td>
</tr>
<tr>
<td>6/1/2009</td>
<td>78</td>
<td>149</td>
<td>7</td>
<td>93</td>
</tr>
<tr>
<td>..........</td>
<td>......</td>
<td>......</td>
<td>......</td>
<td>........</td>
</tr>
<tr>
<td>1/2/2009</td>
<td>120</td>
<td>197</td>
<td>10</td>
<td>105</td>
</tr>
<tr>
<td>2/2/2009</td>
<td>115</td>
<td>151</td>
<td>10</td>
<td>85</td>
</tr>
<tr>
<td>3/2/2009</td>
<td>98</td>
<td>154</td>
<td>8</td>
<td>96</td>
</tr>
<tr>
<td>..........</td>
<td>......</td>
<td>......</td>
<td>......</td>
<td>........</td>
</tr>
<tr>
<td>4/2/2009</td>
<td>90</td>
<td>195</td>
<td>8</td>
<td>93</td>
</tr>
</tbody>
</table>

RPM: Respirable particulate matter

Table VI. Weather forecasting from September, 2009 to June 2010

<table>
<thead>
<tr>
<th>Date</th>
<th>New data inserted into</th>
<th>Weather Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/9/2009</td>
<td>Cluster2</td>
<td>Hot, dry and smogy, dusty, fly ash, smoggy, fog, Mist</td>
</tr>
<tr>
<td>2/9/2009</td>
<td>Cluster3</td>
<td></td>
</tr>
<tr>
<td>3/9/2009</td>
<td>Cluster2</td>
<td>Hot, dry and smogy</td>
</tr>
<tr>
<td>4/9/2009</td>
<td>Cluster2</td>
<td>Hot, dry and smogy, dusty, fly ash, smoggy, fog, Mist</td>
</tr>
<tr>
<td>..........</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28/9/2009</td>
<td>Cluster3</td>
<td>dusty, fly ash, smoggy, fog, Mist</td>
</tr>
<tr>
<td>29/9/2009</td>
<td>Cluster3</td>
<td>dusty, fly ash, smoggy, fog, Mist</td>
</tr>
</tbody>
</table>

Chakraborty et al., 2014
K-Means Clustering

Algorithm:

1) Assign random means (centroids) for K-cluster (orange points): $m^{(1)}_1$, $m^{(1)}_2$, … … $m^{(1)}_k$

2) Assign each data point to the cluster whose mean has the least squared Euclidean distance (the "nearest" mean)

3) Calculate the new mean (centroid) for each cluster

$$m^{(t+1)}_i = \frac{\sum_{j=1}^{n_i} x_j}{n_i} \quad 1 \leq i \leq k$$

4) Iterate until the centroids do not change significantly.
Inverse Problem in Geodynamics

Inverse Problems in Geodynamics Using Machine Learning Algorithms

M. H. Shahnas1,2 DOI, D. A. Yuen2 DOI, and R. N. Pysklywec1 DOI

1Department of Earth Sciences, University of Toronto, Toronto, Ontario, Canada, 2Department of Earth Sciences and Minnesota Supercomputing Institute, University of Minnesota, Twin Cities, Minneapolis, MN, USA

Abstract During the past few decades numerical studies have been widely employed to explore the style of circulation and mixing in the mantle of Earth and other planets. However, in geodynamical studies there are many properties from mineral physics, geochemistry, and petrology in these numerical models. Machine learning, as a computational statistic-related technique and a subfield of artificial intelligence, has rapidly emerged recently in many fields of sciences and engineering. We focus here on the application of supervised machine learning (SML) algorithms in predictions of mantle flow processes. Specifically, we emphasize on estimating mantle properties by employing machine learning techniques in solving an inverse problem. Using snapshots of numerical convection models as training samples, we enable machine learning models to determine the magnitude of the spin transition-induced density anomalies that can cause flow stagnation at midmantle depths. Employing support vector machine algorithms, we show that SML techniques can successfully predict the magnitude of mantle density anomalies and can also be used in characterizing mantle flow patterns. The technique can be extended to more complex geodynamic problems in mantle dynamics by employing deep learning algorithms for putting constraints on properties such as viscosity, elastic parameters, and the nature of thermal and chemical anomalies.
The effect of spin transition on the bulk modulus of \( \text{Pv} \). \( \text{Fe}^{3+} \) Al: Catalli et al. (2011), \( \text{Fe}^{2+} \): (Lundin et al. (2008), Al: Yagi et al. (2004), \( \text{Fe}^{3+} \): Catalli et al. (2010), Mg-Pv: Lundin et al. (2008), PREM: Dziewonski and Anderson (1981) [Catalli et al, 2011].

Iron Spin Transition in the Lower Mantle

(a) density in kg/m\(^3\), (b) thermal expansivity in 1/K, (c) bulk modulus in GPa, in Fp [Wu et al., 2009, Shahnas et al., 2011]
Mid-mantle stagnations are prevalent globally in seismic tomographic inversions, but previous explanations for their existence are not satisfactory.

Iron spin transition in the lower mantle minerals can significantly influence the thermoelastic properties of the mantle material.

Numerical experiments explore how the electronic spin transition in iron modifies the mantle flow, and in particular the fate of sinking cold slabs and rising plumes [Shahnas et al. JGR, 2011; Shahnas et al., G³, 2016; Shahnas et al. GJI, 2017; Shahnas et al. EPSL, 2017; Li et al., 2018].
Tackling the Problem
a) Inverse Problem Approach
b) Machine Learning Approach
\[
\Delta \rho_{STot}(m) = \Delta \rho_{SFp} + \Delta \rho_{\alpha Fp} + \Delta \rho_{KFp} + m\beta \Delta \rho_{KPv}
\]

\[|\Delta \rho_{KPv} | = |\Delta \rho_{KFp} |, \quad \beta = 0.01, \quad 0 \leq m \leq 299.\]

Raw Data

[Wu and Wentzcovitch, 2009; Catalli, 2010; Shahnas et al., 2011, 2017]
Elaborating the Problem

**n samples (images)**

\[ T \]

1
2
3
4

\[ n \]

**2D: \( m \times m \) pixels**

\[ \begin{matrix}
\vdots \\
\text{1D: } m^2
\end{matrix} \]

\[ \begin{array}{cccc}
1 & 2 & 3 & \cdots \\
2 & 3 & 4 & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
\end{array} \]

\( n \) images, one per line

![Schematic of a logistic regression classifier.](image)

**2D to 1D array**
Feature Reduction

Raw Data

Meaningful Features
\[
\Delta V_C = [0.0, 1.0]
\]
\[
\Delta V_H = [0.0, 1.0]
\]

\[T_C < T_{ave}^{up} - 60^\circ\]

\[T_H > T_{ave}^{dn} + 60^\circ\]

Cold material

Hot material
Class Labels in Inverse Problem

Table 1 - Normalized volume fractions (features) representing the degrees of the slab ($\Delta V_C$) and plume ($\Delta V_H$) stagnation.

<table>
<thead>
<tr>
<th>$\Delta V_C$</th>
<th>0.11 ≤ $\Delta V_C$ &lt; 0.20</th>
<th>0.20 ≤ $\Delta V_C$ &lt; 0.26</th>
<th>$\Delta V_C$ ≥ 0.26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree 1</td>
<td>C₁</td>
<td>C₂</td>
<td>C₃</td>
</tr>
<tr>
<td>Degree 4</td>
<td>C₄</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Delta V_H$ ≤ 0.08</th>
<th>0.08 ≤ $\Delta V_H$ &lt; 0.16</th>
<th>0.16 ≤ $\Delta V_H$ &lt; 0.20</th>
<th>$\Delta V_H$ ≥ 0.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree 1</td>
<td>H₁</td>
<td>H₂</td>
<td>H₃</td>
</tr>
<tr>
<td>Degree 4</td>
<td>H₄</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 - a) Class labels based on the amount of the spin dependent density anomaly in $P_v$

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
<th>...</th>
<th>Class 14</th>
<th>Class 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ≤ $m$ ≤ 20</td>
<td>21 ≤ $m$ ≤ 40...</td>
<td>...261 ≤ $m$ ≤ 280</td>
<td>281 ≤ $m$ ≤ 299</td>
<td></td>
</tr>
</tbody>
</table>
The Thermal State of the Planetary Mantle

\[
\theta = \frac{g(f)}{1+f^{-2}} + C(f, Ra_B) \left( \frac{(1+f+f^2)H/3}{Ra_B^\beta} \right)^\alpha
\]

\[
Q = \alpha Ra_B^\beta \gamma
\]

Shahnas & Pysklywec, 2019
The Thermal State of the Planetary Mantle

Figure 7- $R^2$- and $\bar{R}^2$-scores for a) Test model, b) Whole data, for predicting mantle mean temperatures. c) Non-dimensional calculated mean temperatures ($\theta_{Calc}$) versus the predicted values ($\theta_{Pred}$). $R^2$- and $\bar{R}^2$-scores for d) Test model, e) Whole data, for deep learning model specified in the text, for predicting surface mean heat fluxes. f) Non-dimensional calculated surface mean heat fluxes ($Q_{Calc}$) versus the predicted values ($Q_{Pred}$). The non-dimensional mean heat fluxes are scaled to 0-1.