Review on Applications of Machine Learning for Landslide Monitoring and Prediction

Wasit Limprasert, PhD

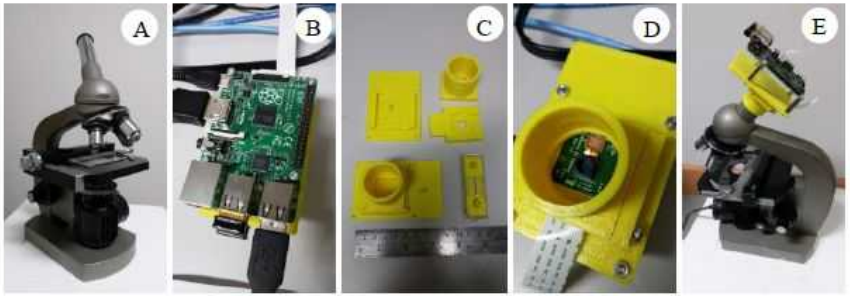
01/06/2017

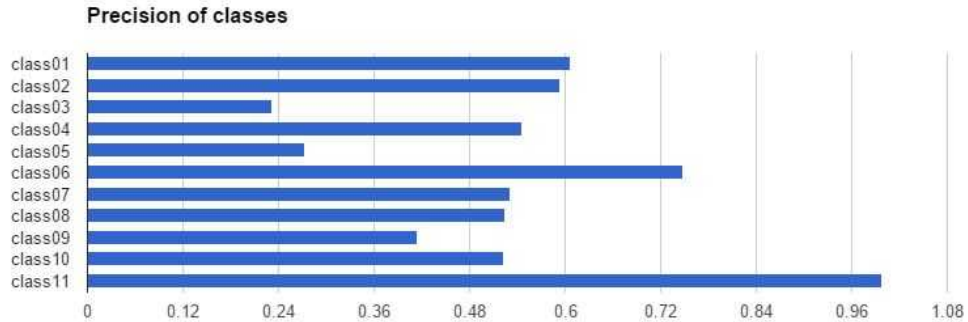
Contents

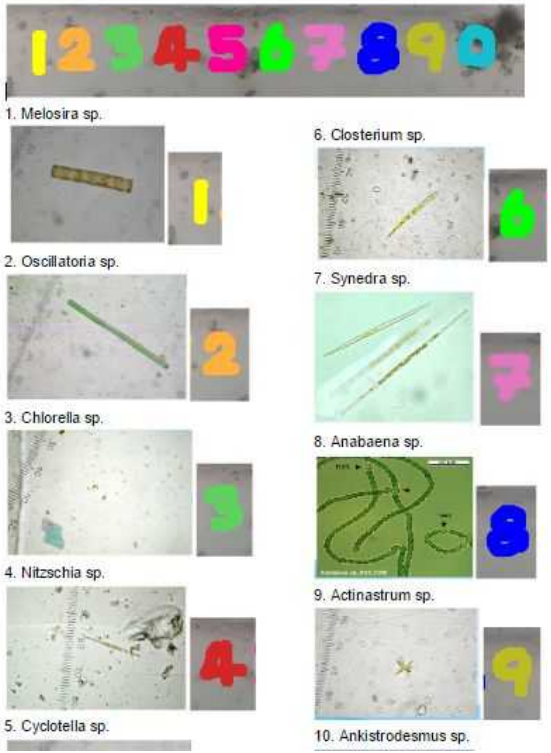
- Our research
- Landslide and Machine learning
- Intro Deee learning

Algae Image Classification using Parallel Randomforest

- To develop a system to detect algae for MWA
- To evaluate the Random Forest classifier







อุปกรณ์เก็บภาพสาหร่าย

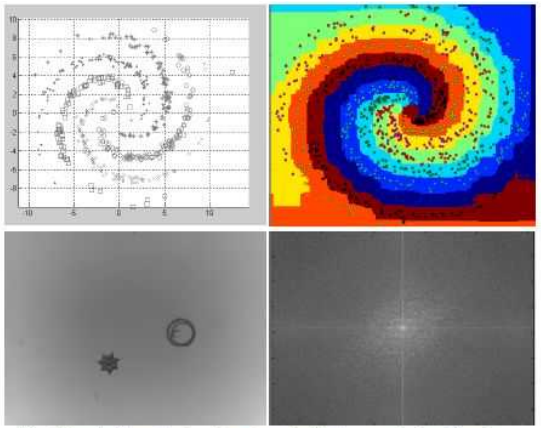
6.2.1 การติดตั้งอุปกรณ์เก็บภาพที่สถานีสูบน้ำสำแล

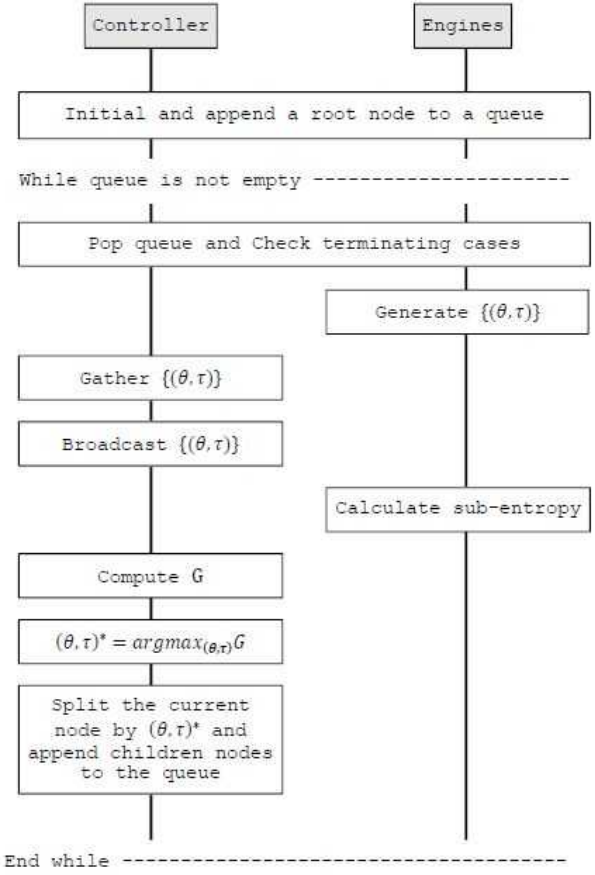
รูปที่ 5: ลักษณะการดำเนินการซ่อมบำรุง

รูปที่ 6: ส่วนประกอบของระบบอิเล็กทรอนิกส์

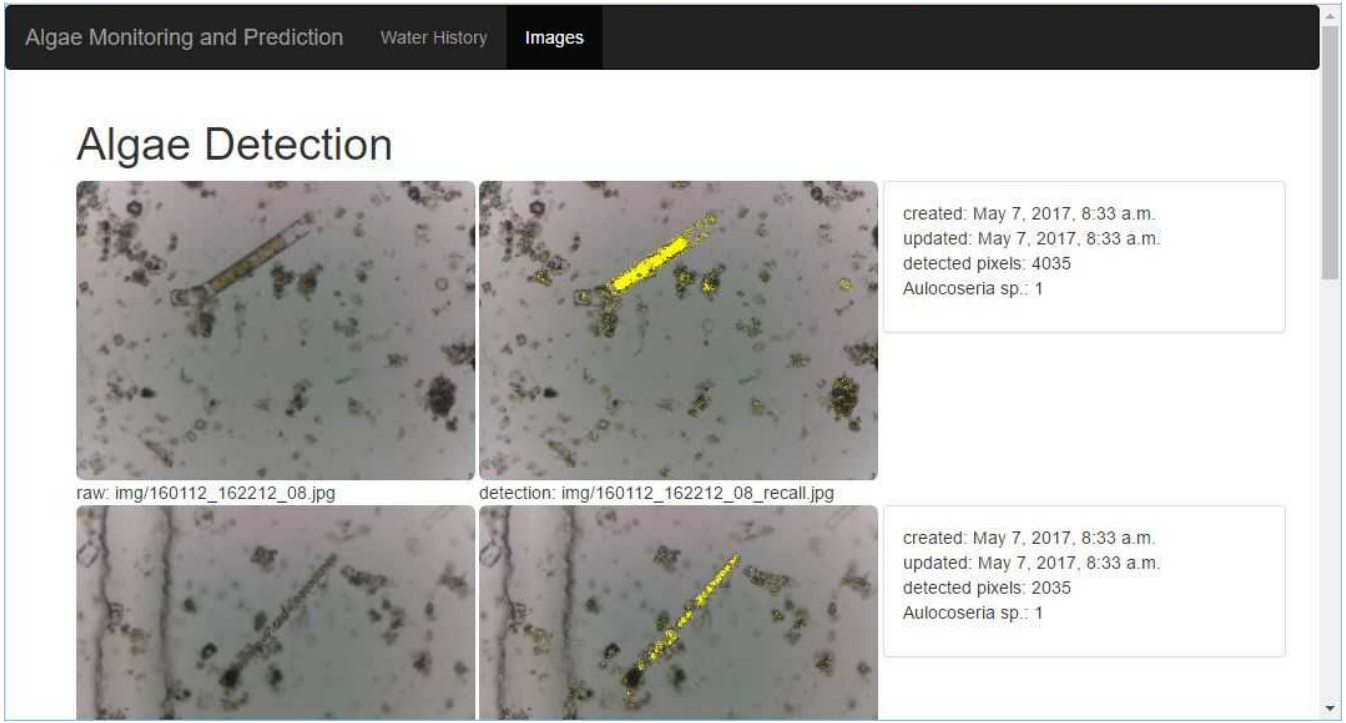
Parallel Randomforest

- To accelerate training speed of Random Forest
- To study IPython Cluster



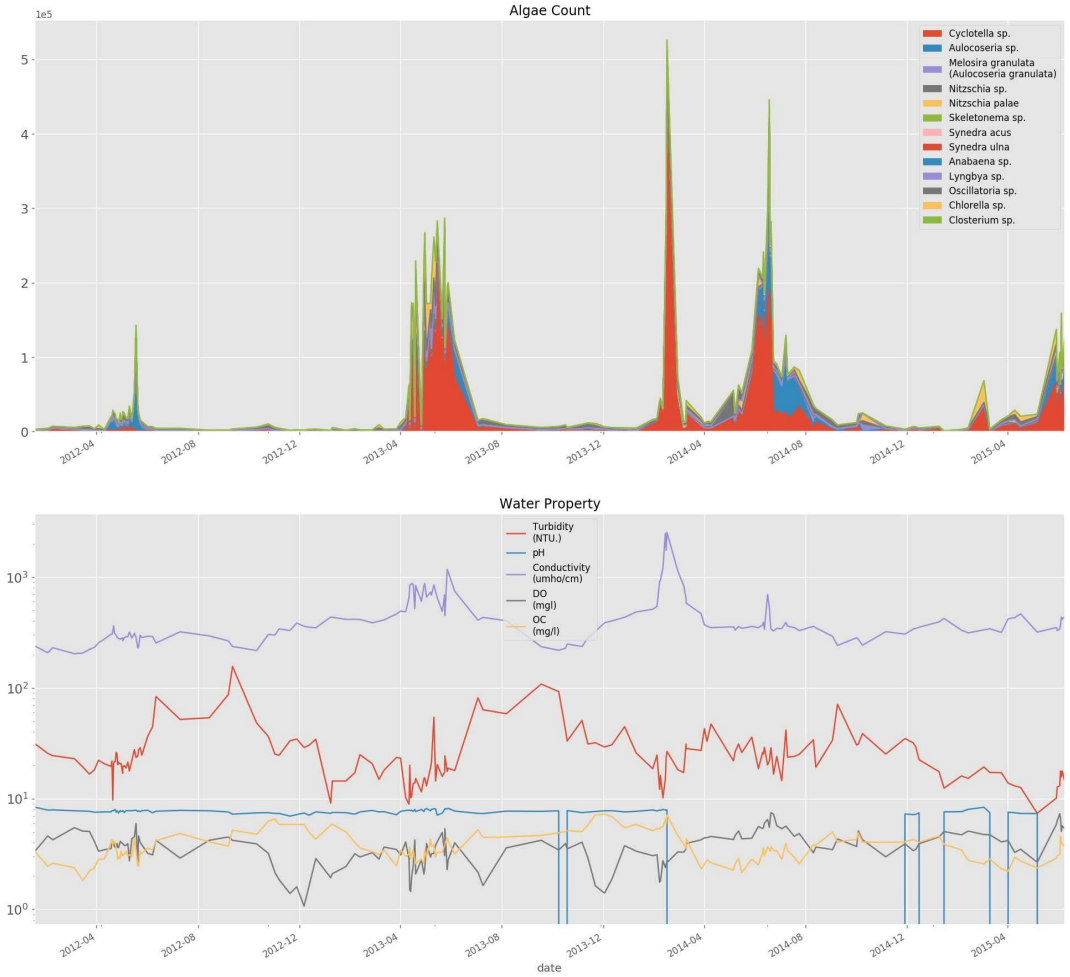


Algae Image Detection



Water Information

- Algae population
- Water Property
- To monitor and prediction population of algae



A Case Study of Data Analysis for Educational Management

- To help students choose courses for next enrolment
- To develop grade prediction system
- To analyze prerequisite and generate better curriculum

Before appending

STUDENTID	TERM	COURSEID	GRADE
0001	2009/1	X	A
0001	2010/1	Y	B
0001	2010/2	Z	C
0002	2009/1	X	D
0002	2010/2	Z	F

appending prior results

After appending

STUDENTID	TERM	COURSEID	GRADE	X	Y	Z
0001	2009/1	X	A	-	-	-
0001	2010/1	Y	B	A	-	-
0001	2010/2	Z	C	A	B	-
0002	2009/1	X	D	-	-	-
0002	2010/2	Z	F	D	-	-

Data mining

Cleaning → Transformation → Training → Evaluation

pd.dropna()

Before appending

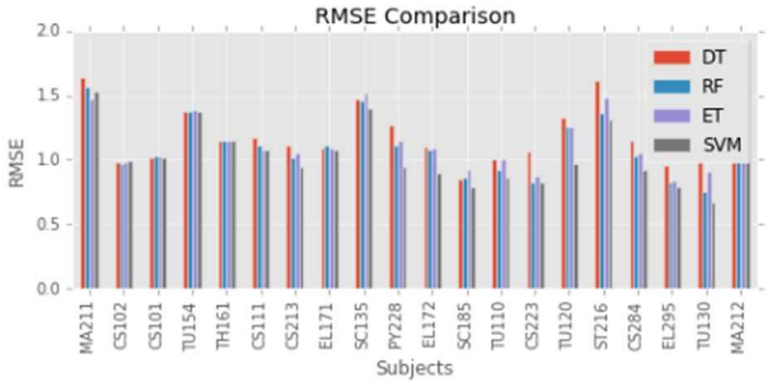
STUDENTID	TERM	COURSEID	GRADE
0001	2009/1	X	A
0001	2010/1	Y	B
0001	2010/2	Z	C
0002	2009/1	X	D
0002	2010/2	Z	F

After appending

appending prior results

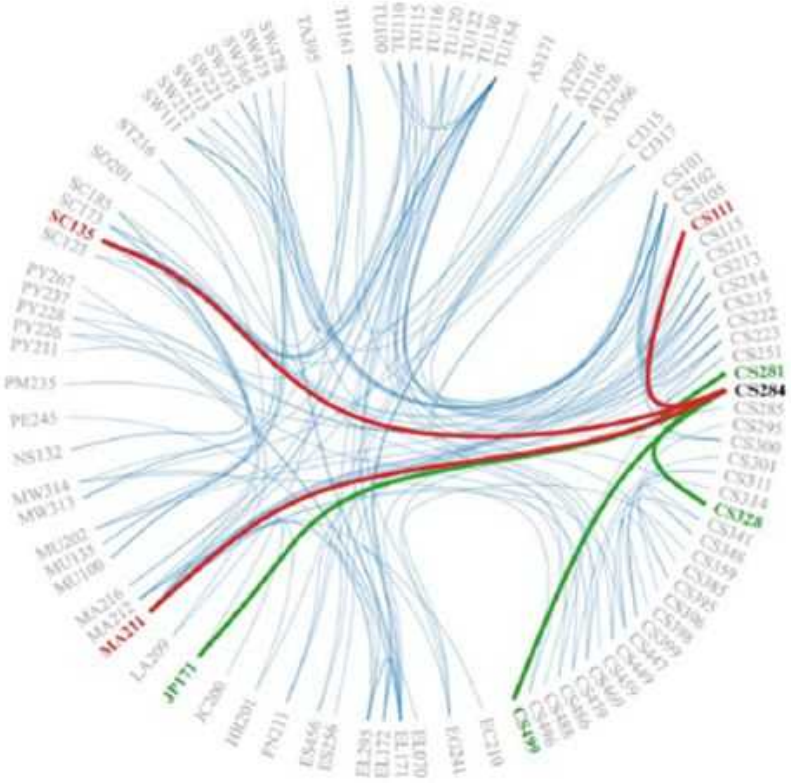
STUDENTID	TERM	COURSEID	GRADE	X	Y	Z
0001	2009/1	X	A	-	-	-
0001	2010/1	Y	B	A	-	-
0001	2010/2	Z	C	A	B	-
0002	2009/1	X	D	-	-	-
0002	2010/2	Z	F	D	-	-

et.fit()



Result

- Overall prediction result
- Grade prediction with RMSE 1.01
- Small number of data 28,272 records
- An new approach to extract new perquisites from existing curriculum



Arrival Time Prediction and Train Tracking Analysis

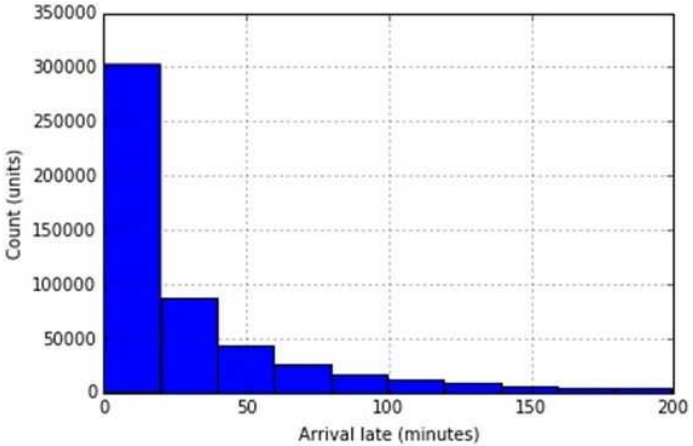
Problem

- Arrival late
- Unpredict time table schedule
- The mean difference between schedule time and actual arrival time is about ± 16 minutes

Result

- Data set 1 year of 2015 around 975,386 records
- ± 3.8 minutes
- Rescheduled time table
- Feature important score
- Visualization of relation among features

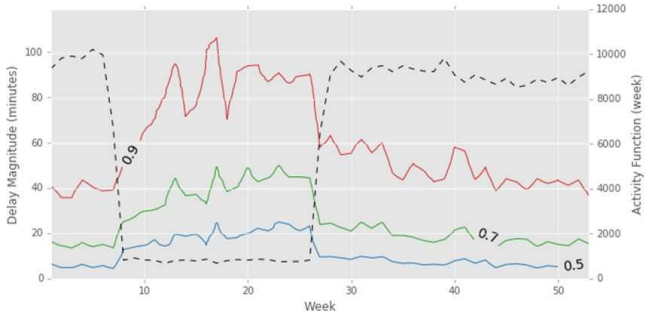
Arrival Time Prediction and Train Tracking Analysis



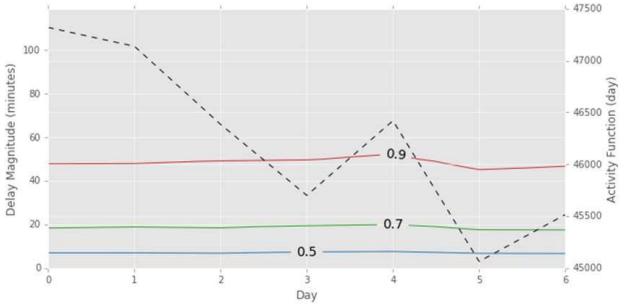
Ranking	Feature	FIS
1	*week*	0.273
2	*day*	0.117
3	*station*	0.113
4	*train_no*	0.107
5	*arrive_time*	0.072
6	*leave_time*	0.054
7	*arrive_cause*	0.049
8	*default_arrivetime*	0.041
9	*leave_cause*	0.041
10	*default_leavetime*	0.041

Regression	Random Forest	ANN	Linear
RMSE	3.863	124.907	25.380
MAE	2.001	60.582	14.976

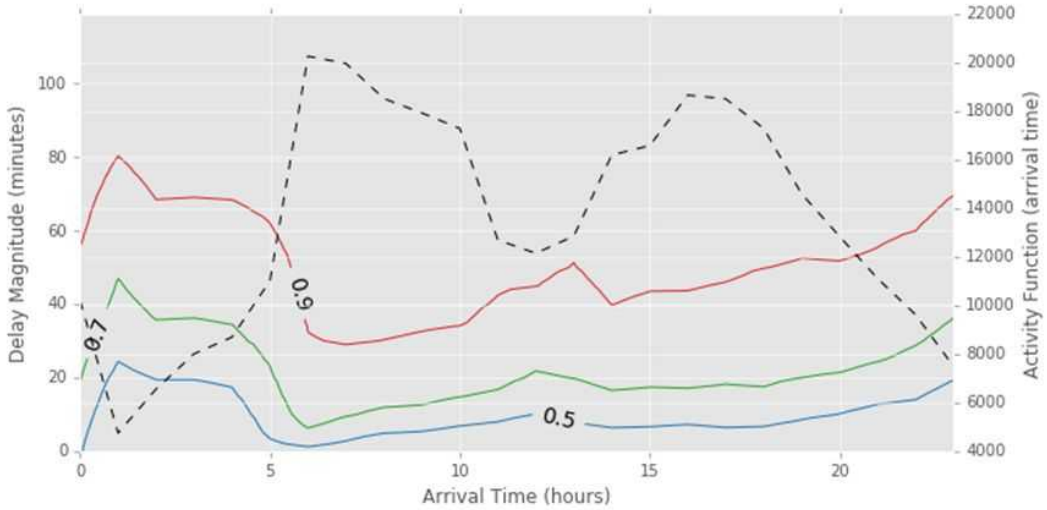
Arrival Time Prediction and Train Tracking Analysis



Week8-25 has significantly less number of services

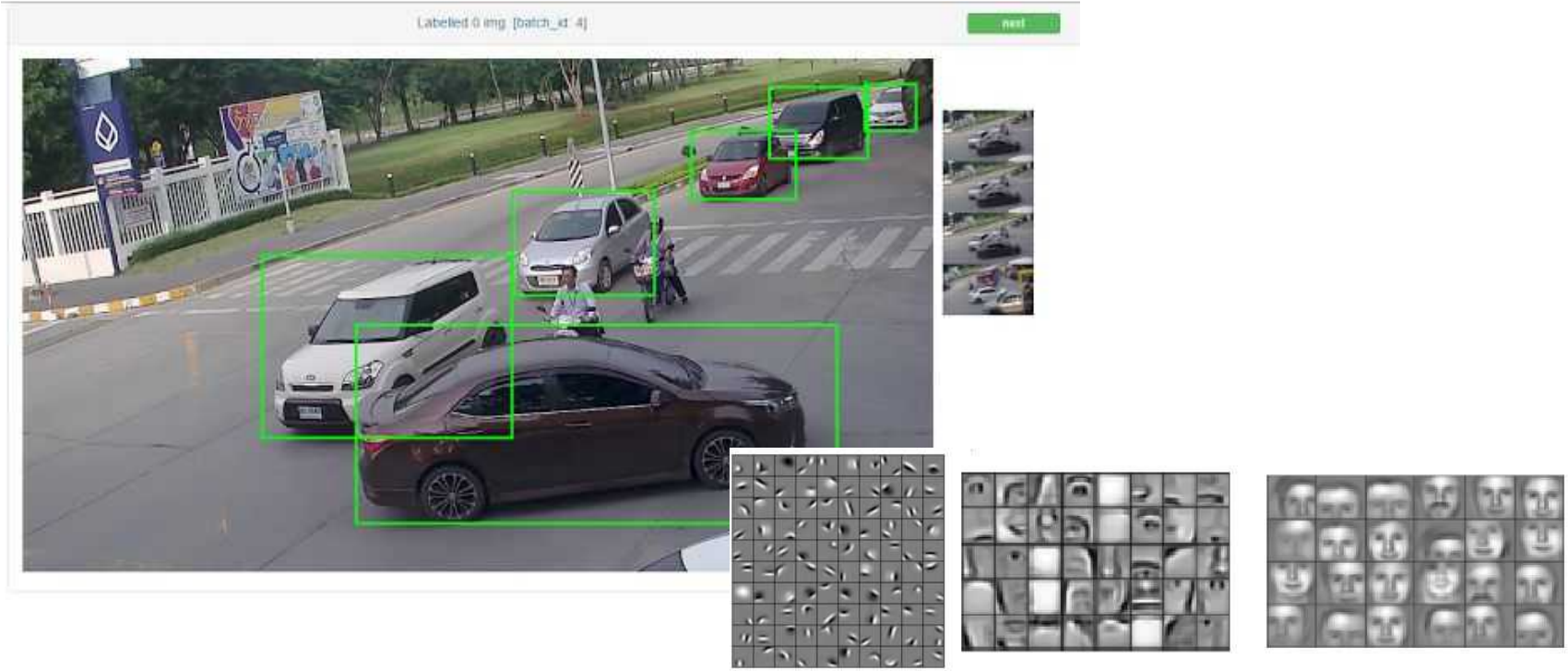


Thursday and Saturday have less number of passengers



- Delay and activity are inversely proportional
- Main Activity start at 6:00
- Delay at noon is interesting and need more investigation

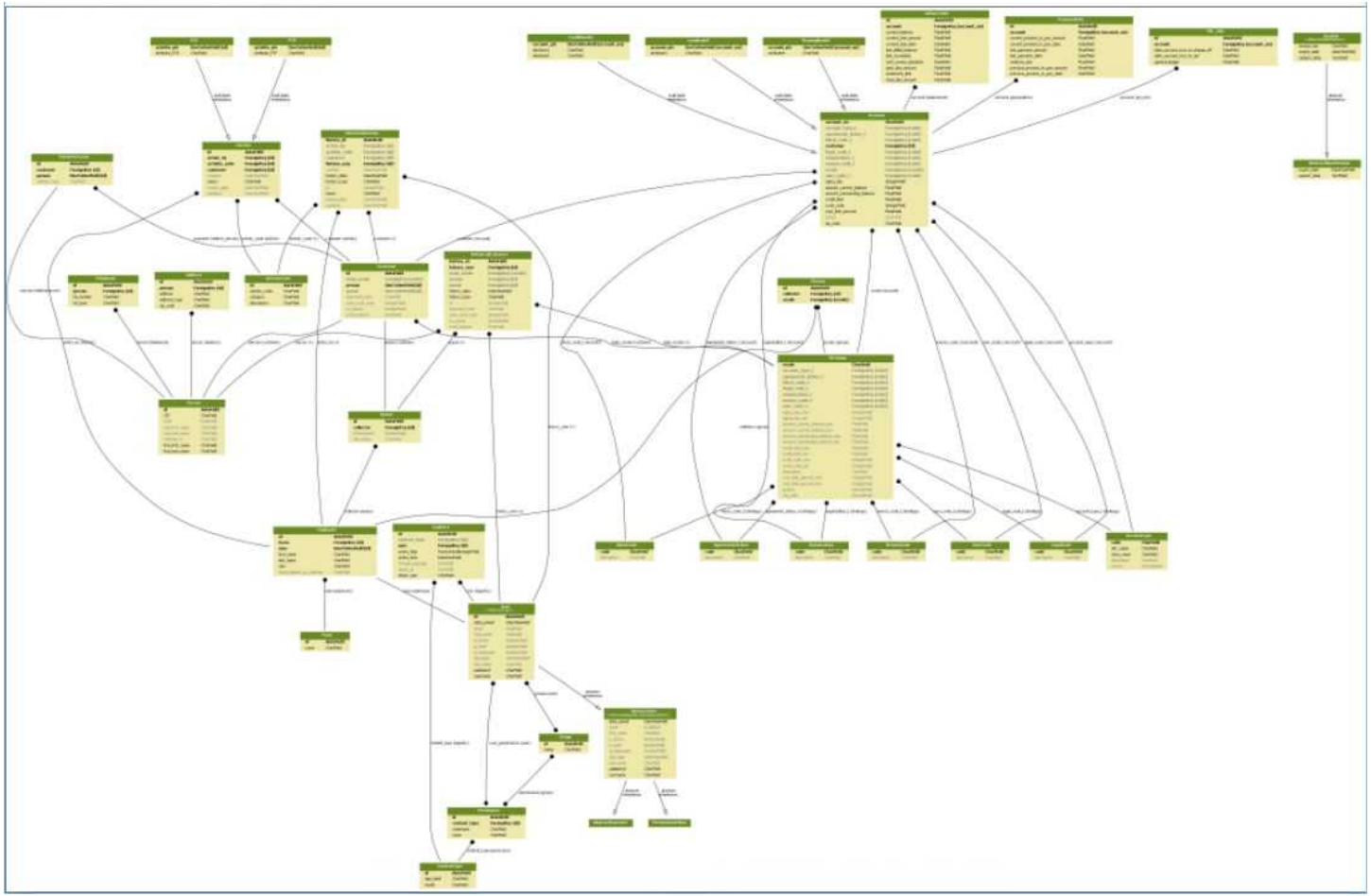
Deep Learning



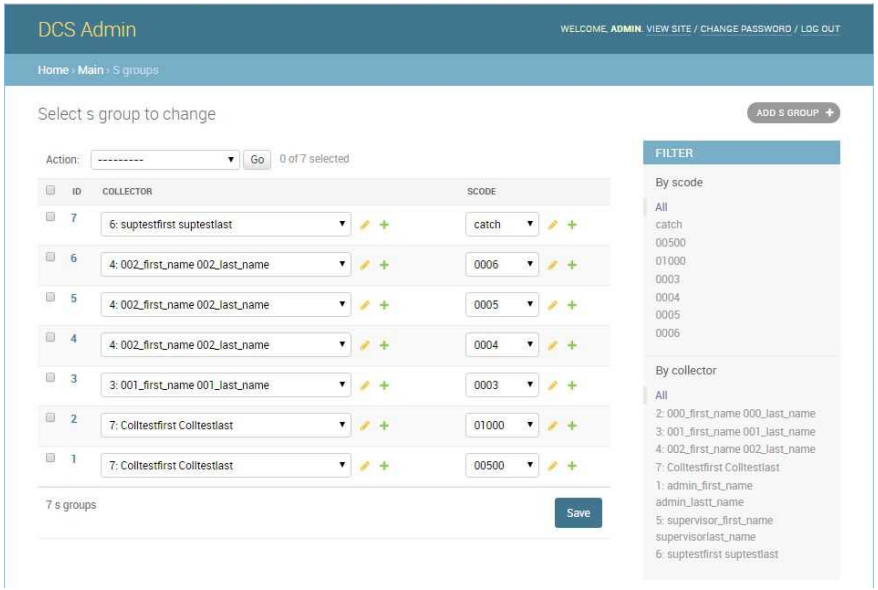
Convolution Neural Network

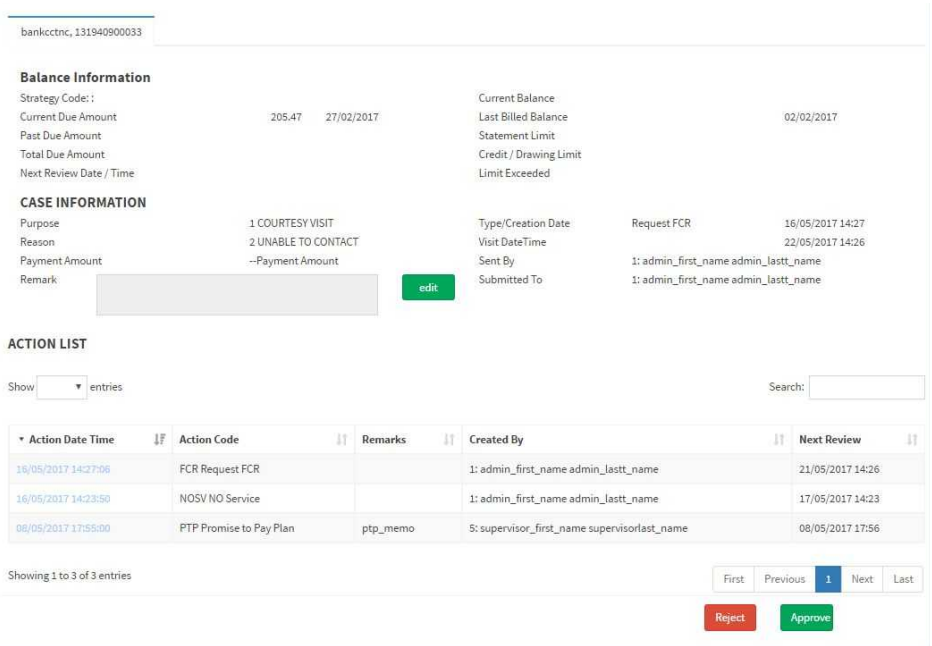
Titan

- GeForce Titan X
- GPU 3072 cores
- Peak computing power at 11 Tera flops
- Global memory 12GB
- RAM 128GB
- SSD



Software Development



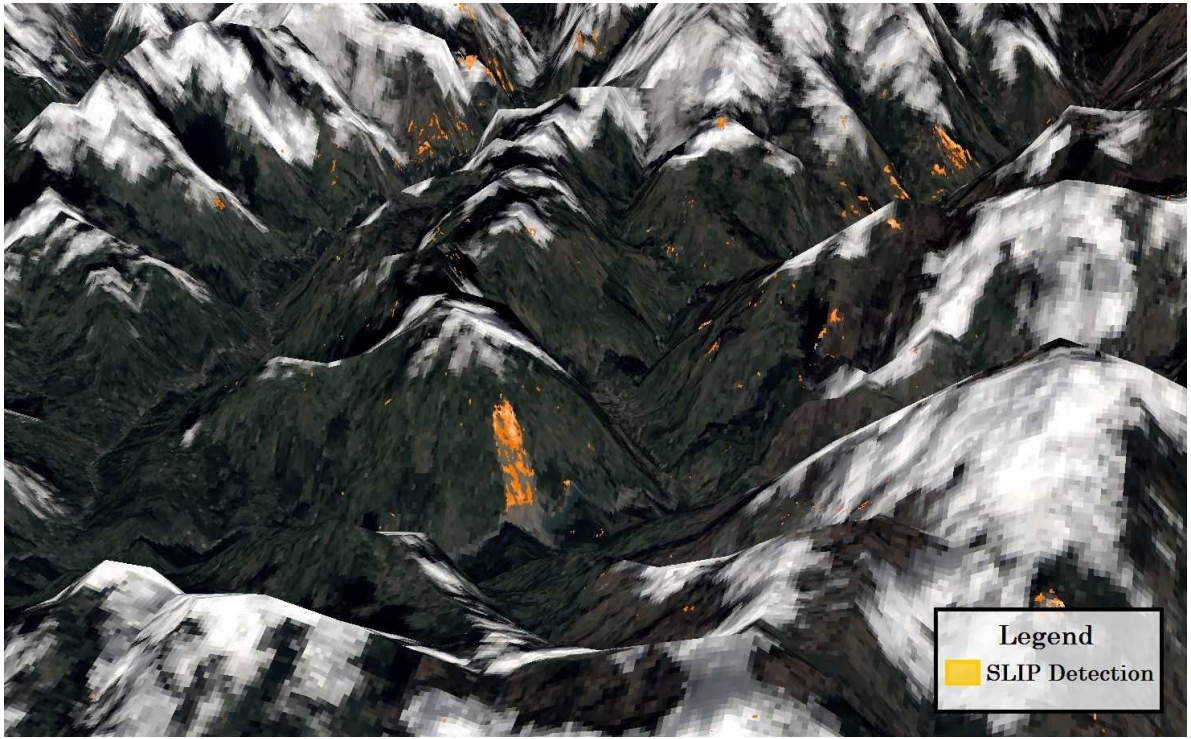


Review on Applications of Machine Learning for Landslide Monitoring and Prediction

01/06/2017

Detecting Landslides in Nepal with Landsat

Input: RGB+infrared images

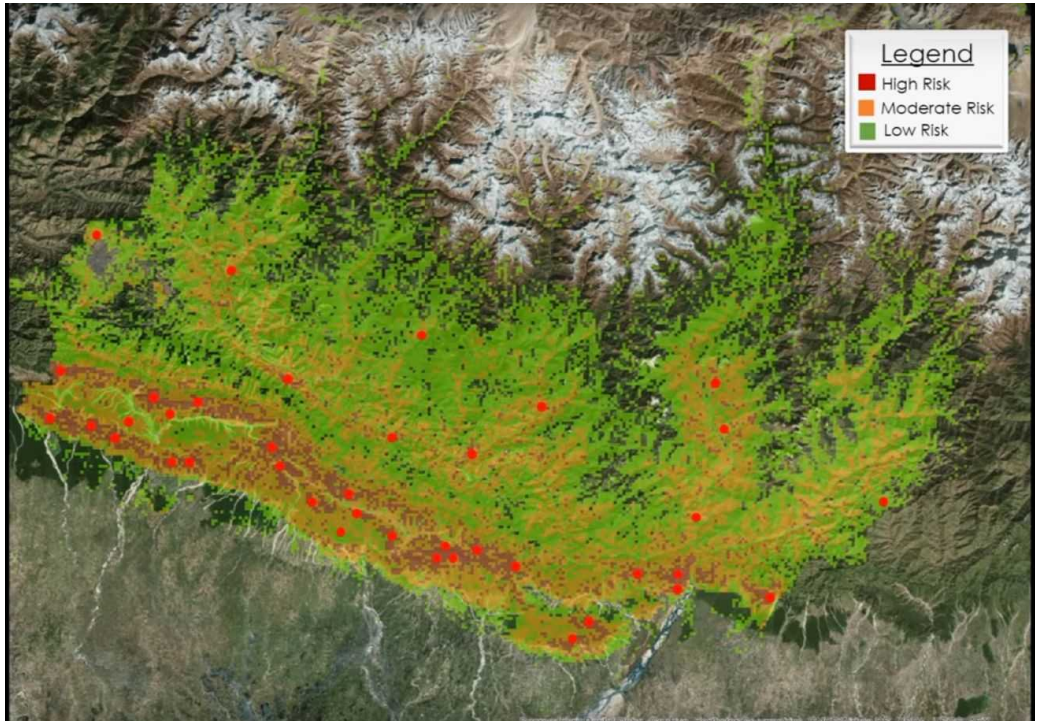


https://landsat.gsfc.nasa.gov/detecting-landslides-in-nepal-with-landsat/

Gorkha earthquake

April 2015, 7.8 Gorkha earthquake caused over 9,000 casualties a $1billion damages

Himalaya Disasters - NASA DEVELOP Fall 2015

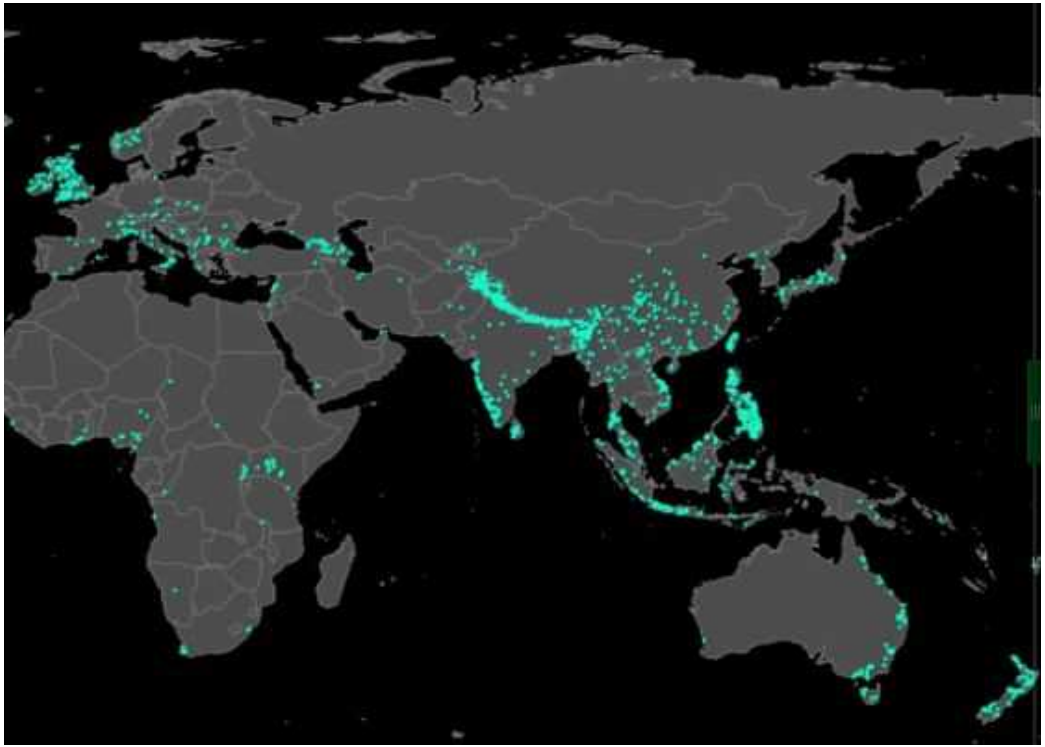


- Red band change means changing of vegetations
- m=NIR-SWIR
- Soil moisture change = m_{new}/m_{old}
- SRTM topography image

- Slip detection
  - Moisture change >40%
  - Slope > 15%
- Future
  - Web platform to calibrate the algorithm

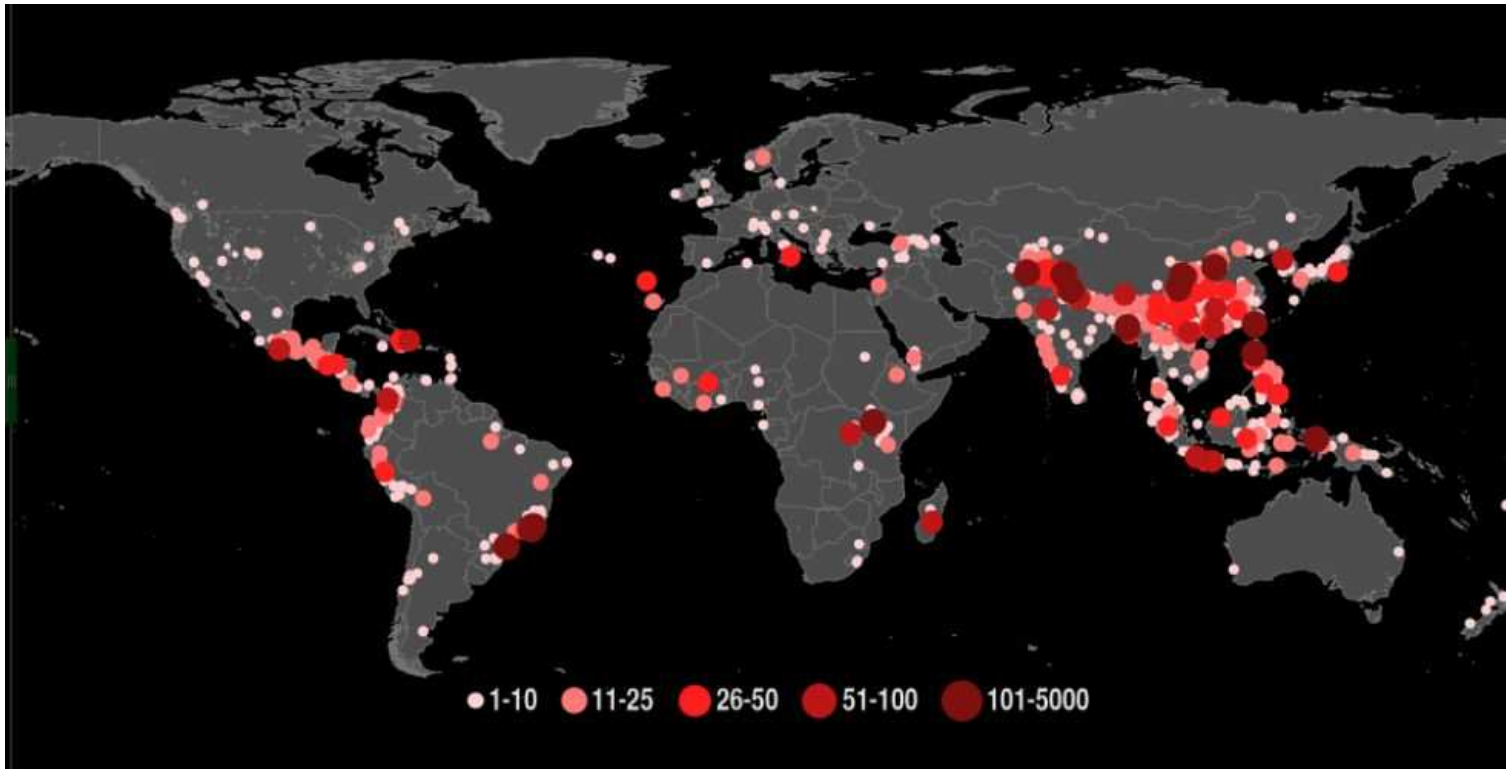
https://earthobservatory.nasa.gov/NaturalHazards/view.php?id=88319&src=nhrss

NASA Global Landslide Catalog

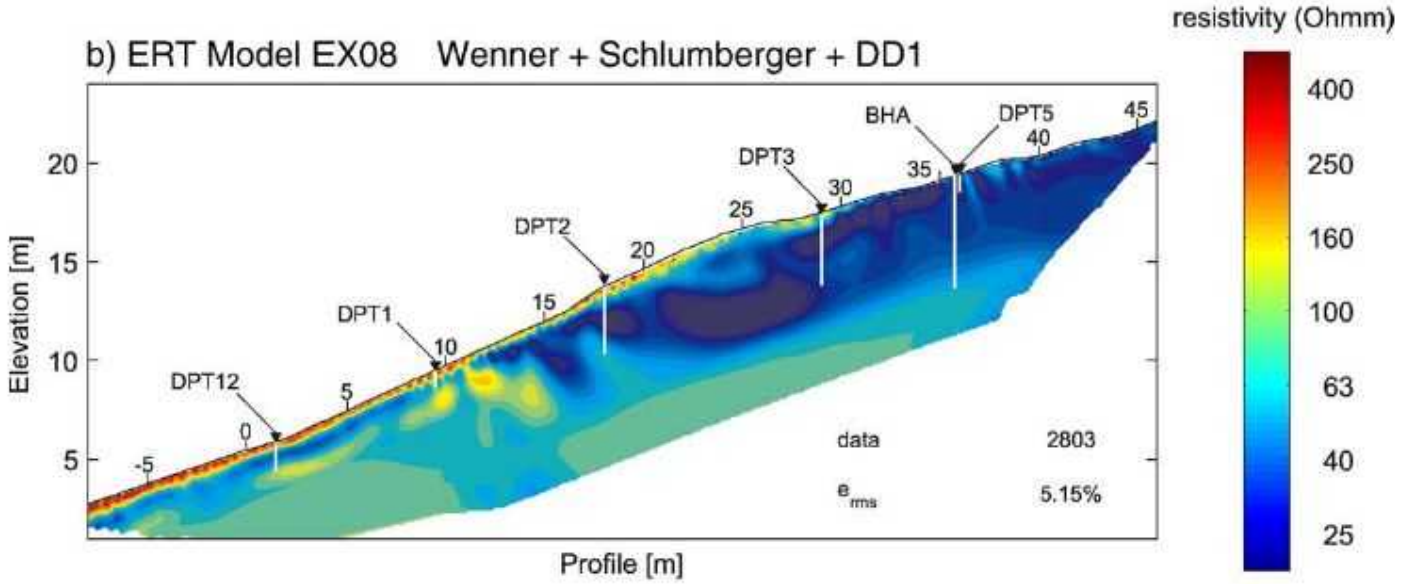


- Location and detail of previous landslides
- Mostly caused by Asian Monsoon

Landslide casualties



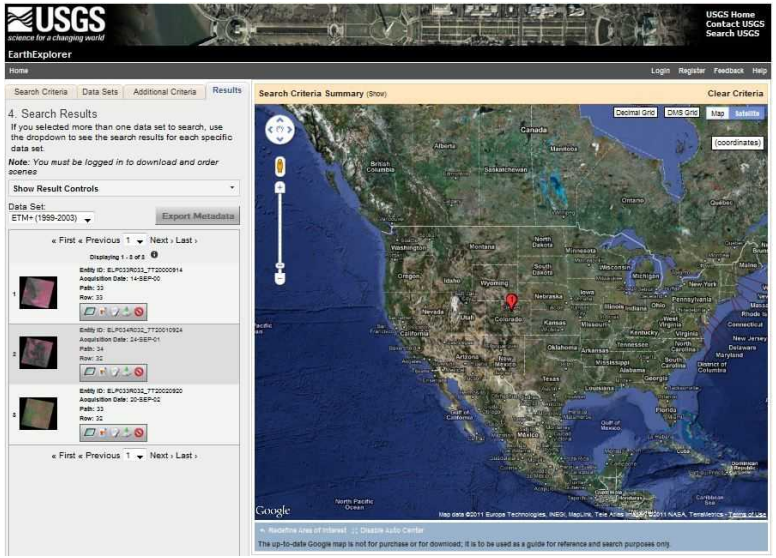
Electrical resistivity tomography (ERT)



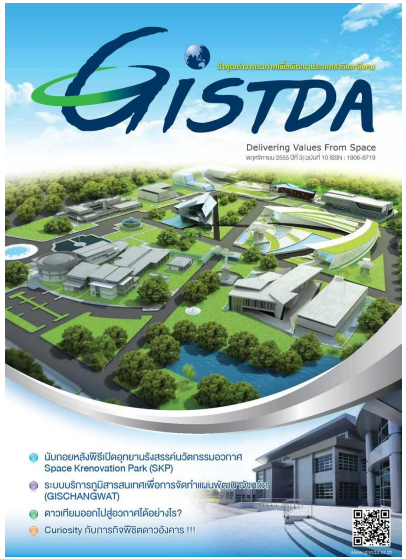
To convert from resistivity map to moisture map, expert need to identify structure geological layer

Friedel, Thielen and Springman 2006

Satellite Image repository



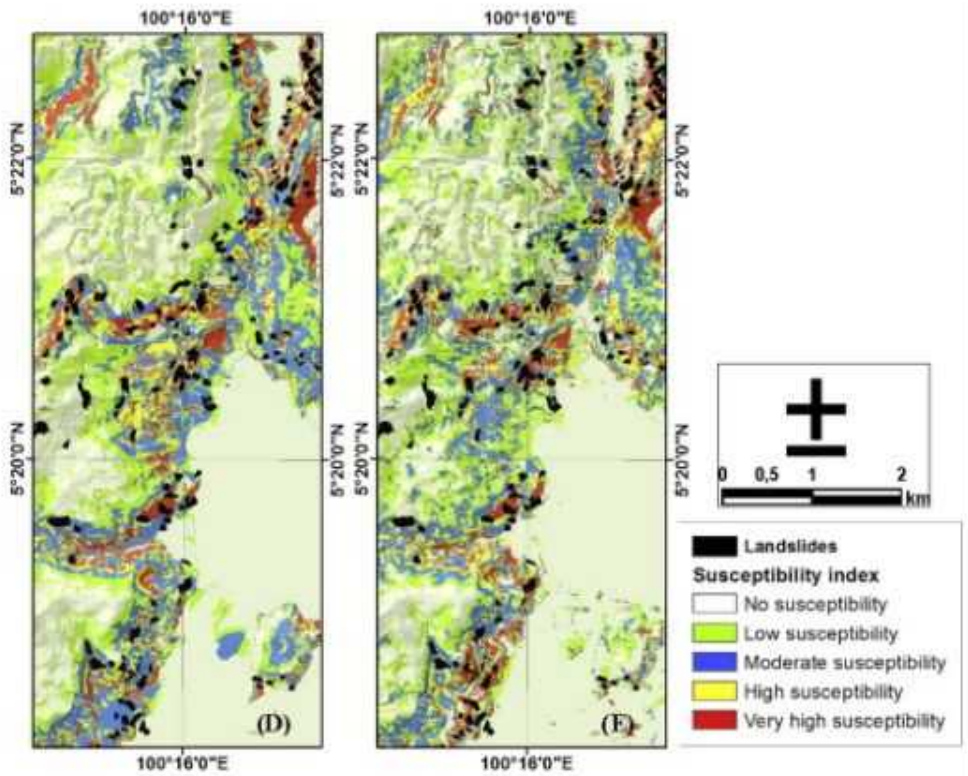
Earth Explorer



GISTDA

- Rapid survey unpopulated area
- Covering larger area
- Still need a ground team for investigation

Landslide susceptibility mapping



Input Data

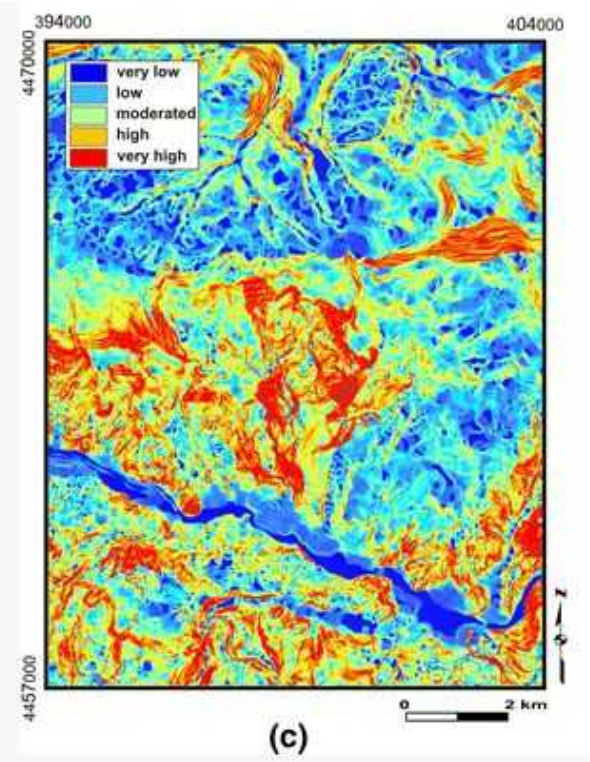
- Altitude
- Slope
- soil type
- plan curvature
- proximity to drainage
- TWI
- proximity to road
- Normalized difference vegetation index

Result using neuro-fuzzy inference

- Area under the ROC curve 94.21%

Biswajeet Pradhan, 2012

landslide susceptibility mapping



Işık Yilmaz 2009

Input Data

- Slope angle
- Slope aspect
- Elevation
- Distance from drainage
- Distance from road
- Distance from sattle ment
- TWI
- SPI

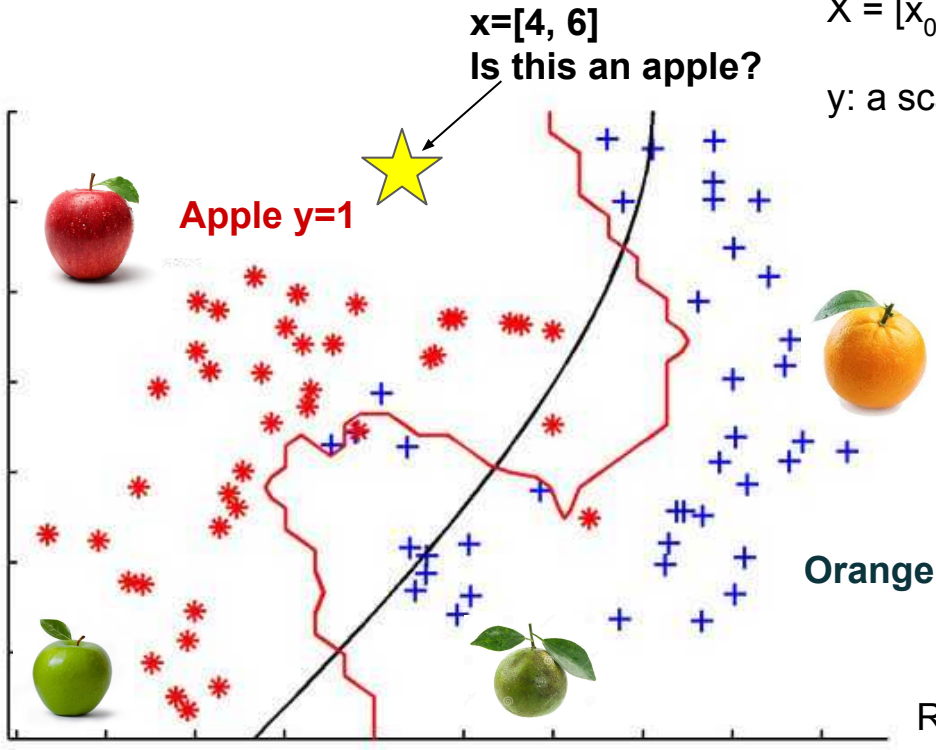
Result using ANN

- Area under the ROC curve 84.6%

Machine Learning

- Classification
- CNN
- RBM and Autoencoder

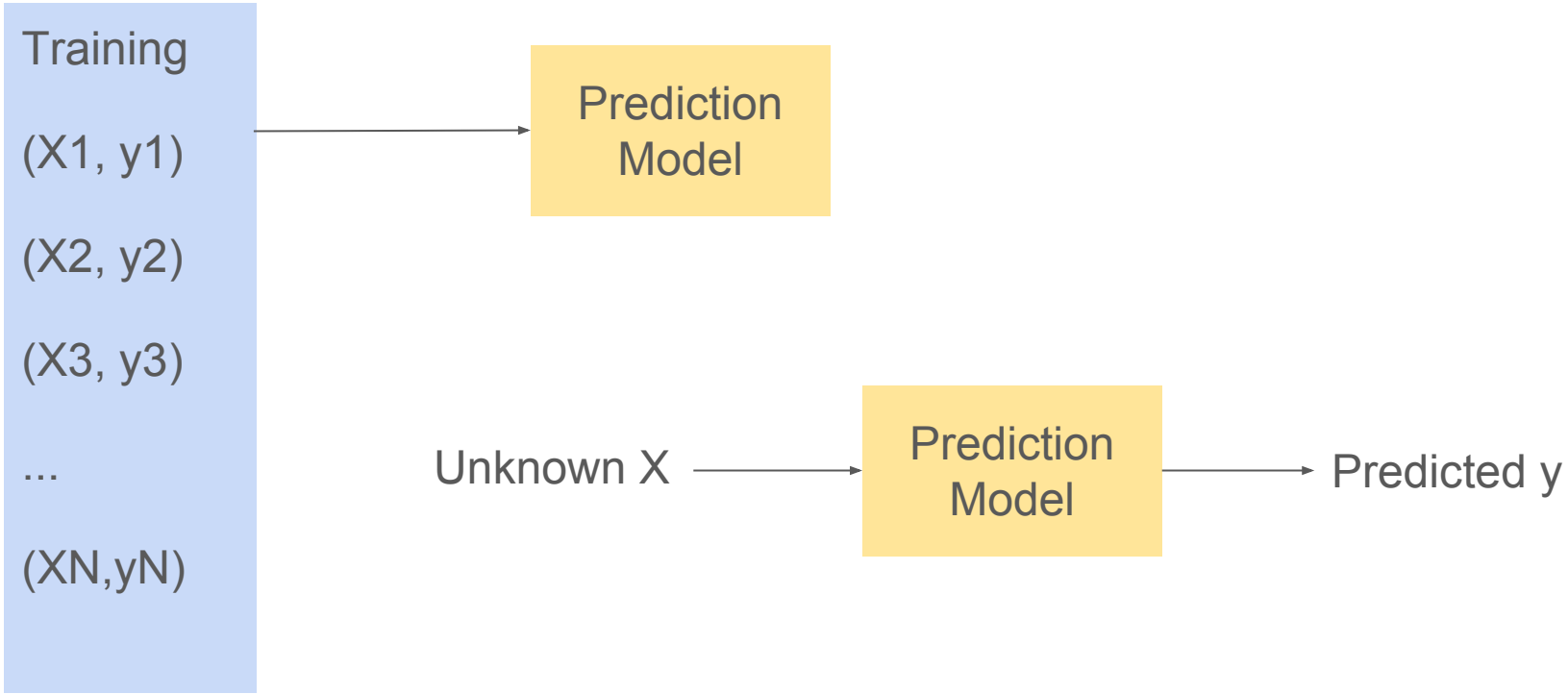
Classification



$X = [x_0, x_1]$: feature vector

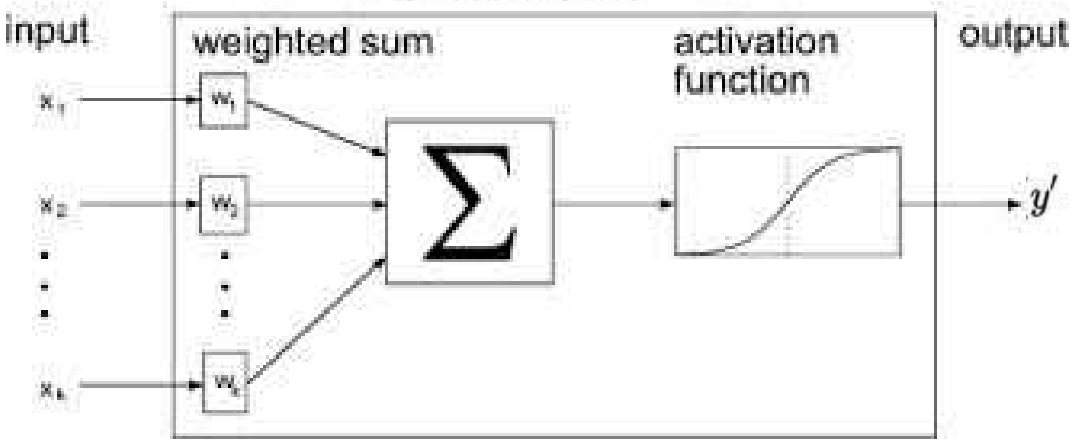
y: a scalar to indicate class

Supervise Learning



Neural network

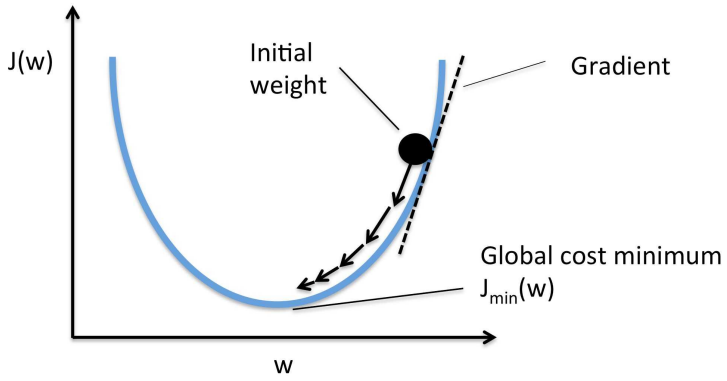
Architecture of a single perceptron



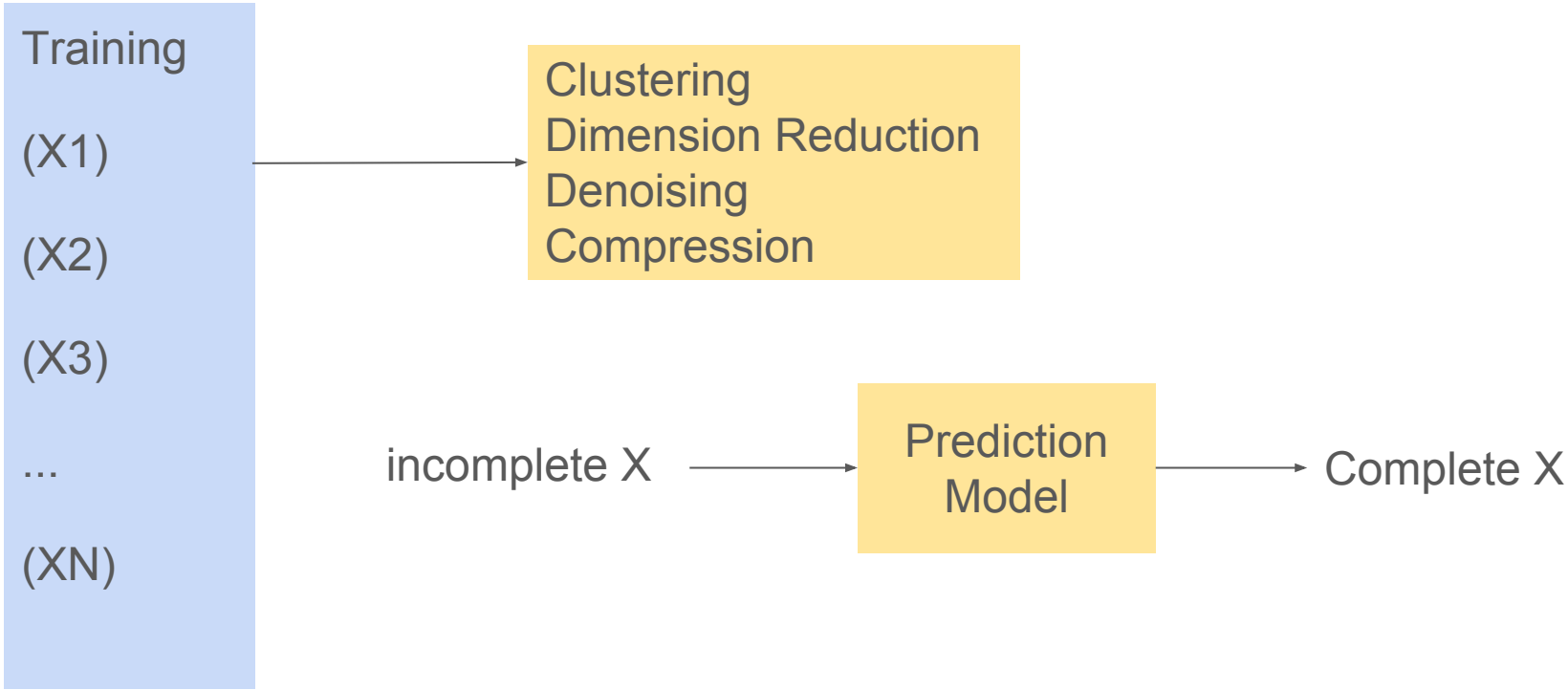
Backpropagation

$$E = \frac{1}{2n}\sum_x \|(y(x) - y'(x))\|^2$$

$$\frac{\partial E}{\partial y'} = (y' - y)$$



Unsupervised Learning

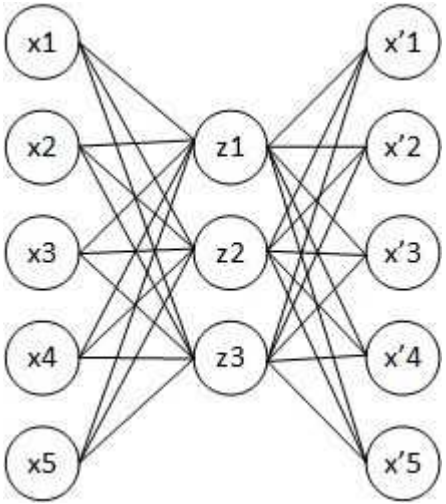


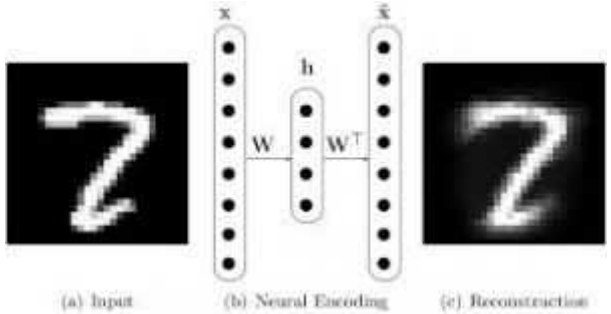
Clustering

MNIST dataset

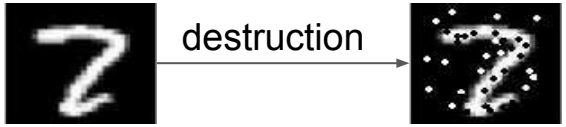
Automatically
Organize

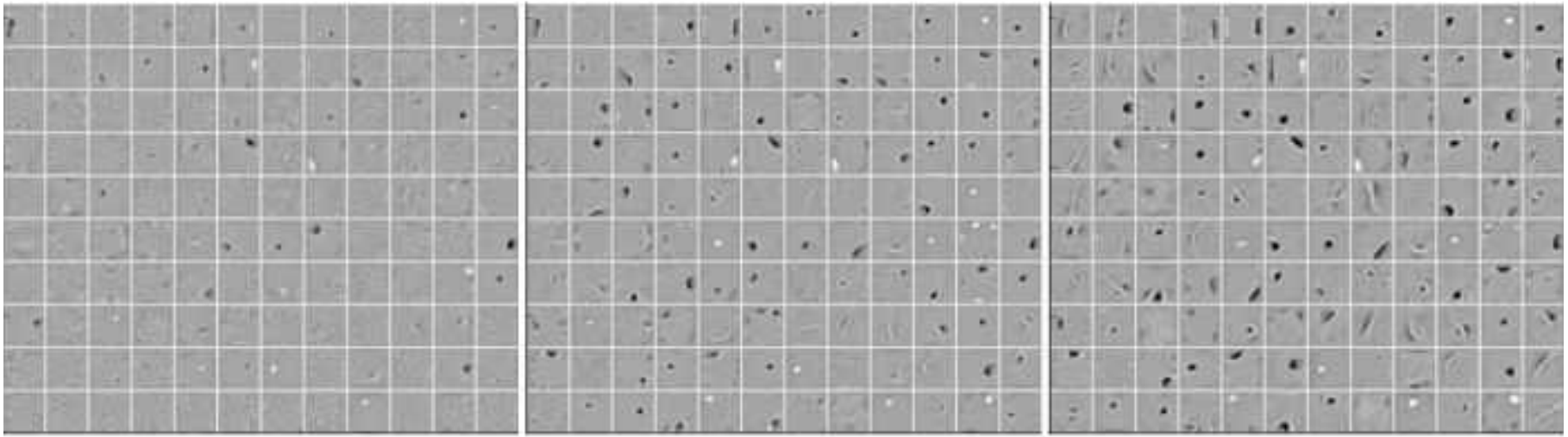
Autoencoder





Autoencoder with noisy training

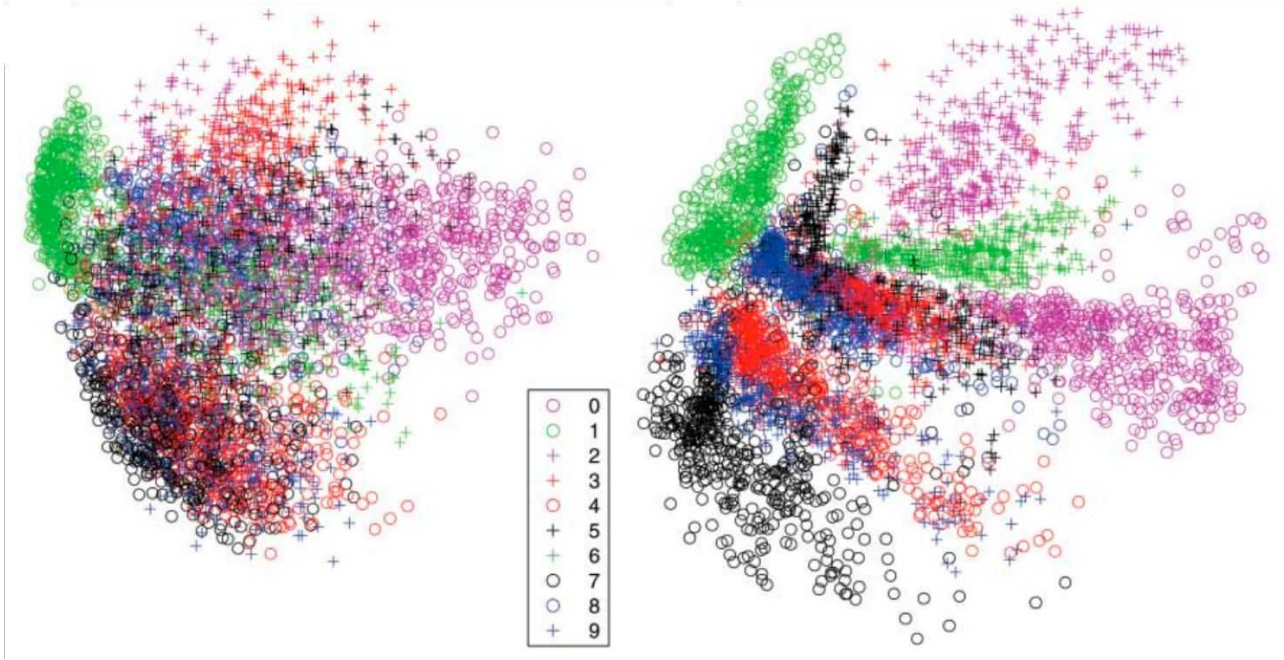




(a) No destroyed inputs (b) 25% destruction (c) 50% destruction

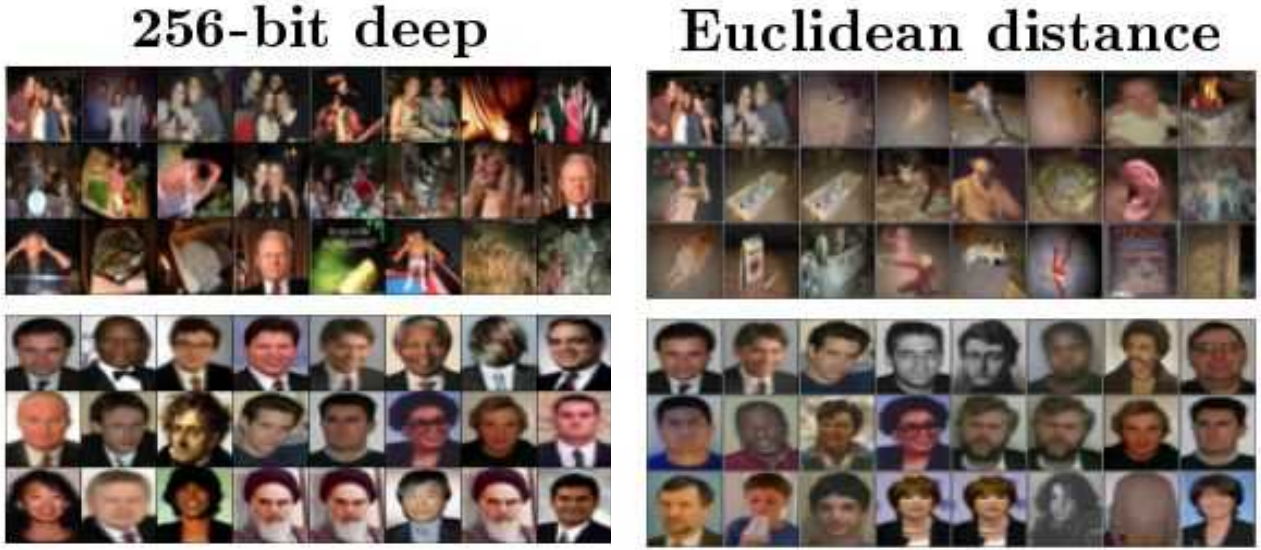
Vincent, Larochelle, Bengio and Manzagol, ICML2007

Autoencoder



(left) PCA and (right) an autoencoder on the MNIST dataset

Content-Based Image Retrieval

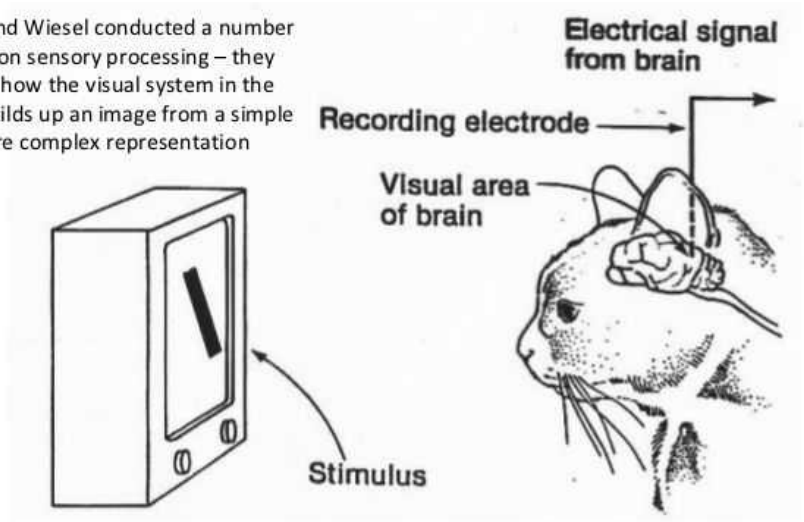


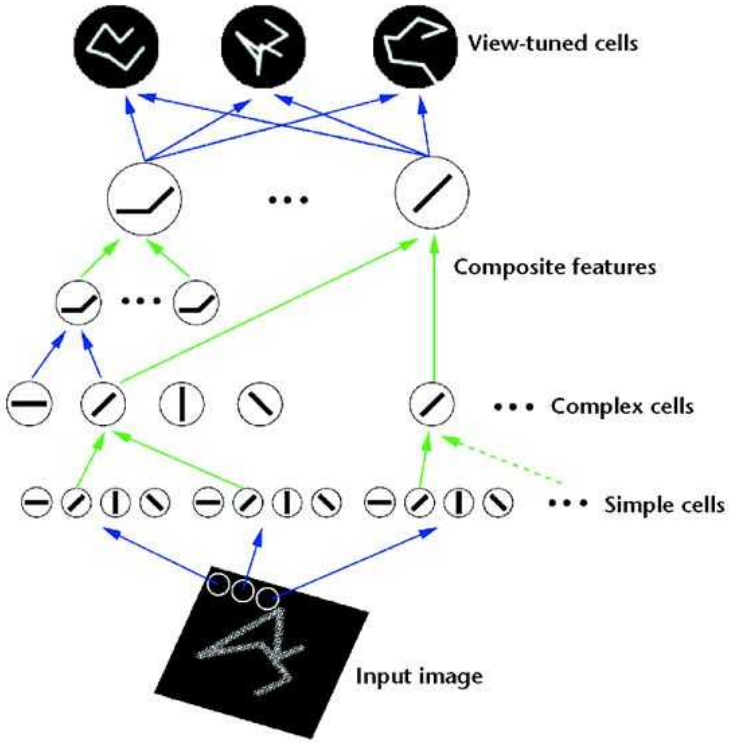
- 1.6 million of 32x32 color images
- Each image is 3072Byte
- Encoded to 256bit (3Byte)
- Searching quicker 1,000 times

Krizhevsky and Hinton 2011

Cat Brain (Hubel and Wiesel)

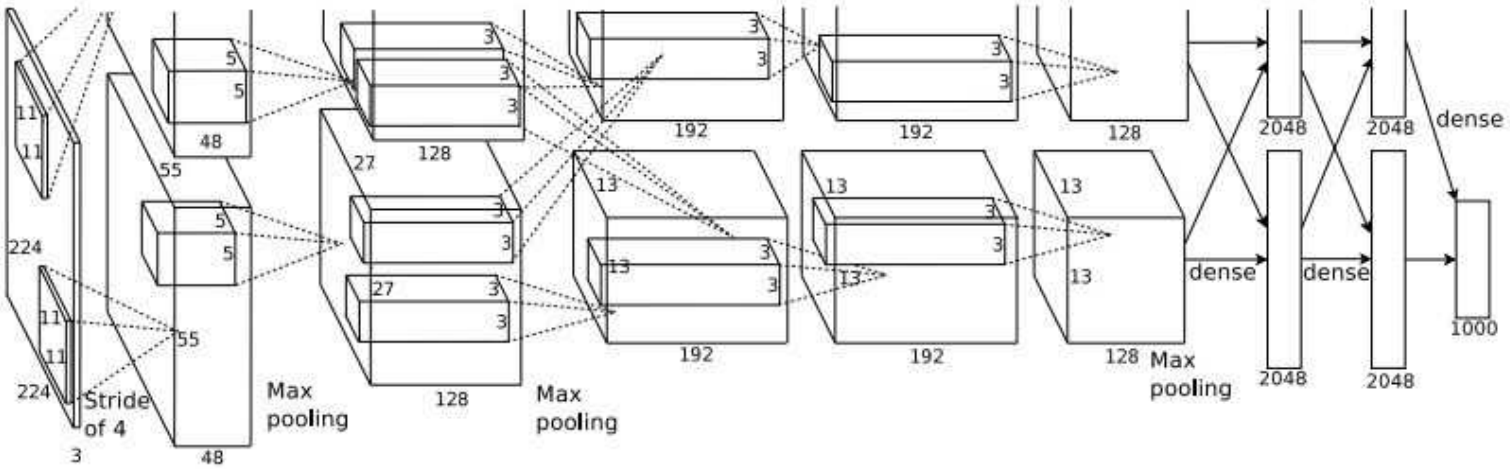
Hubel and Wiesel conducted a number of tests on sensory processing – they showed how the visual system in the brain builds up an image from a simple to a more complex representation





Koch & Poggio,
Predicting the visual world: silence is golden,
Nature Neuroscience 2, 9 - 10 (1999)

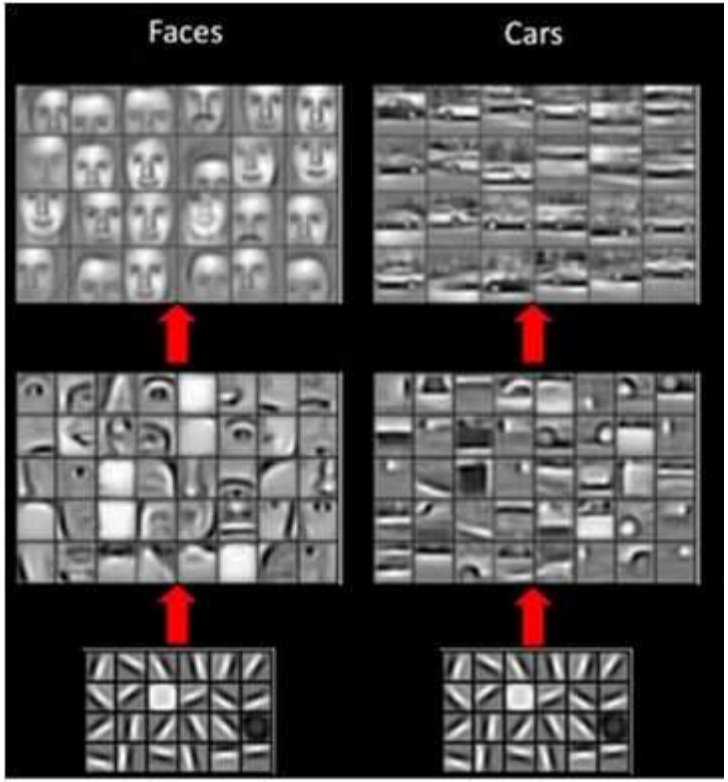
CNN architecture



Krizhevsky, Sutskeve and Hinton 2012

Convolution Neural Network

Krizhevsky, Sutskeve and Hinton 2012



Summary

Existing method to apply machine learning for detecting landslide susceptibility

Basic Machine learning

Sensor and image input

Possibility to use satellite image for suggest ground survey tem