



# Review on Applications of Machine Learning for Landslide Monitoring and Prediction

Wasit Limprasert, PhD

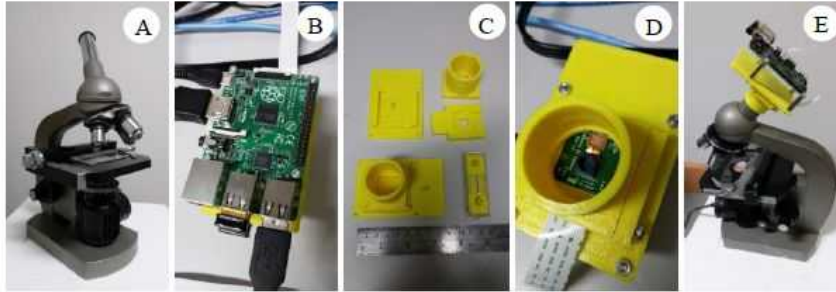
01/06/2017

# Contents

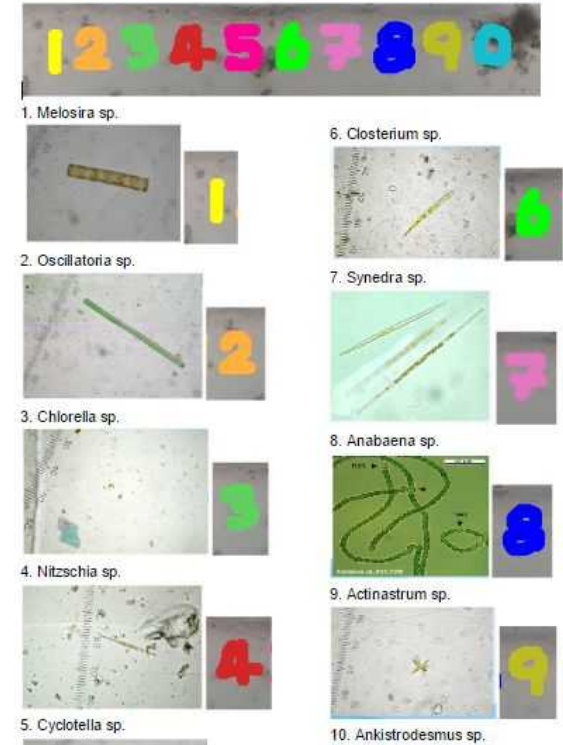
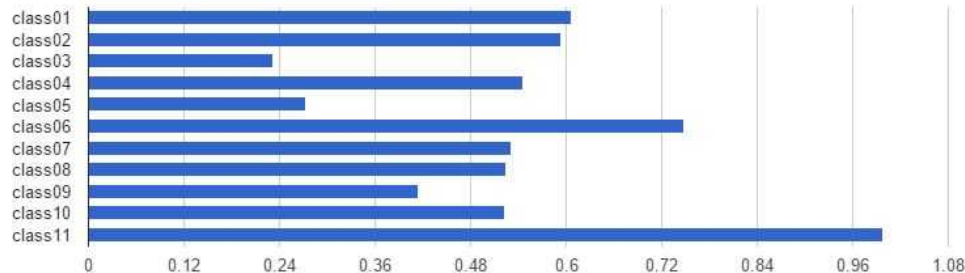
- Our research
- Landslide and Machine learning
- Intro Dee learning

# Algae Image Classification using Parallel Randomforest

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



Precision of classes



# อุปกรณ์เก็บภาพสาหร่าย

## 6.2.1 การติดตั้งอุปกรณ์เก็บภาพที่สถานีสูบน้ำสำแล



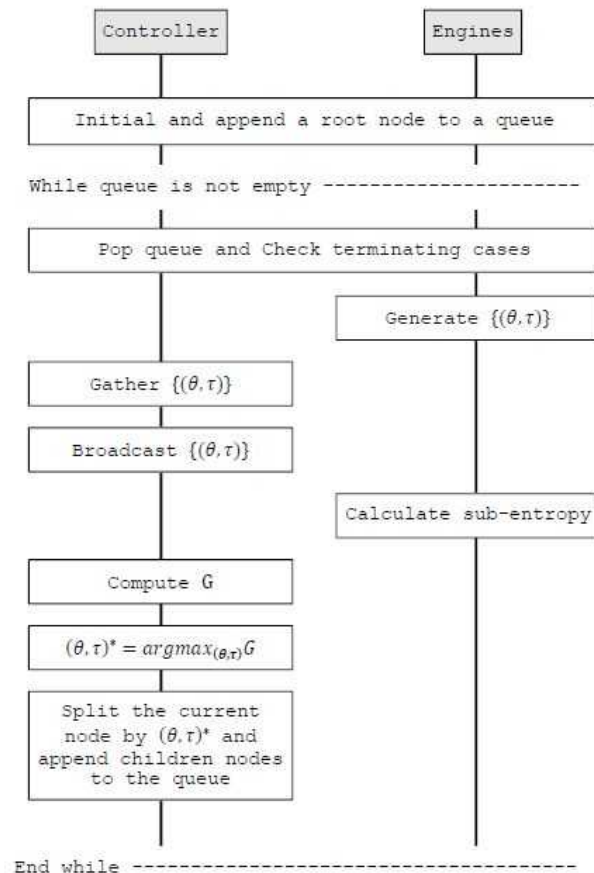
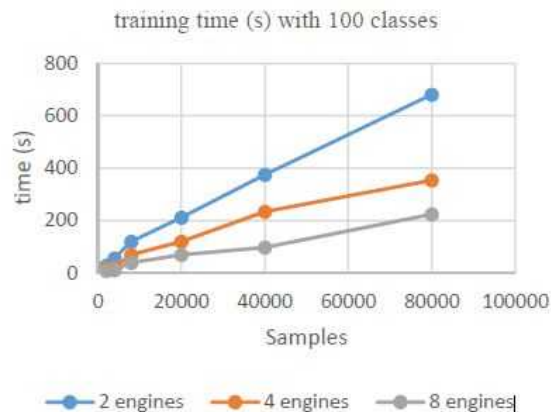
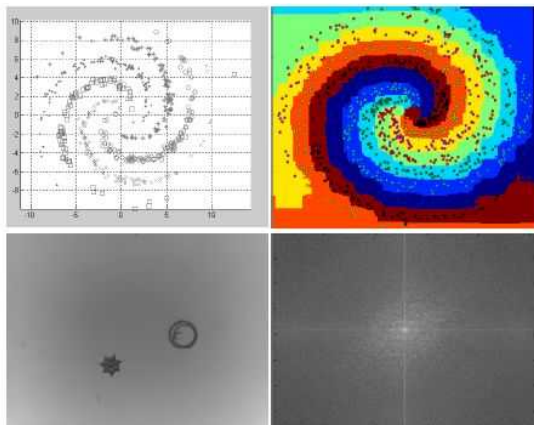
รูปที่ 5: ลักษณะการดำเนินการซ่อมบำรุง



รูปที่ 6: ส่วนประกอบของระบบอิเล็กทรอนิกส์

# Parallel Randomforest

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



# Algae Image Detection

Algae Monitoring and Prediction

Water History

Images

## Algae Detection



raw: img/160112\_162212\_08.jpg



detection: img/160112\_162212\_08\_recall.jpg

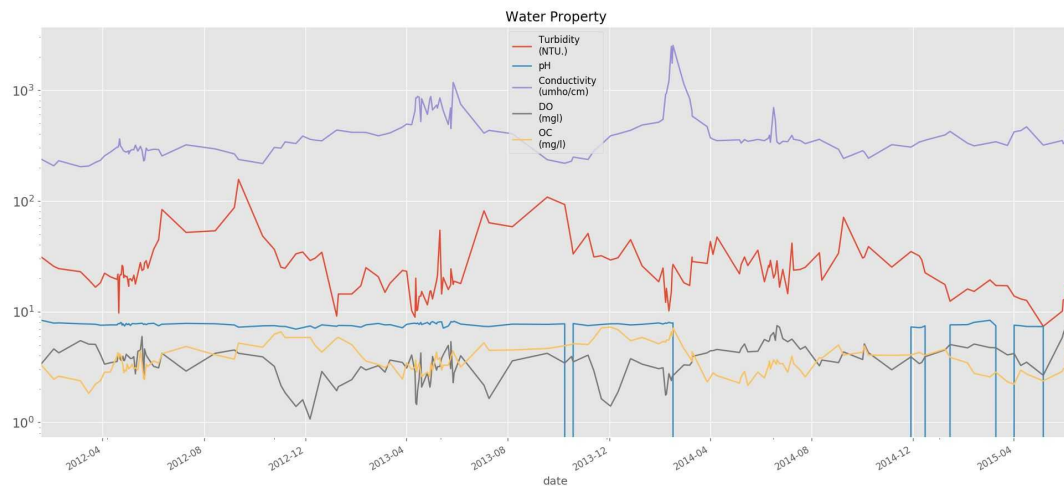
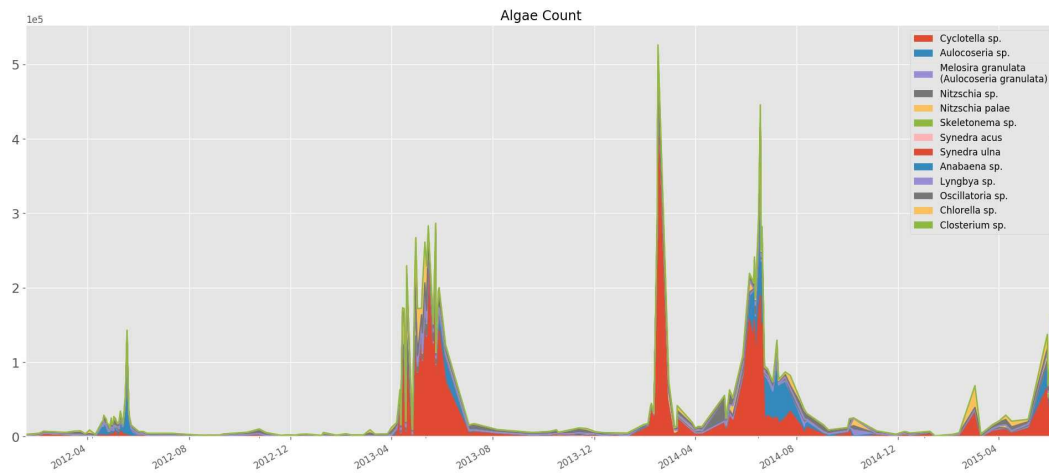
created: May 7, 2017, 8:33 a.m.  
updated: May 7, 2017, 8:33 a.m.  
detected pixels: 4035  
Aulocoseria sp.: 1



created: May 7, 2017, 8:33 a.m.  
updated: May 7, 2017, 8:33 a.m.  
detected pixels: 2035  
Aulocoseria sp.: 1

# 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*

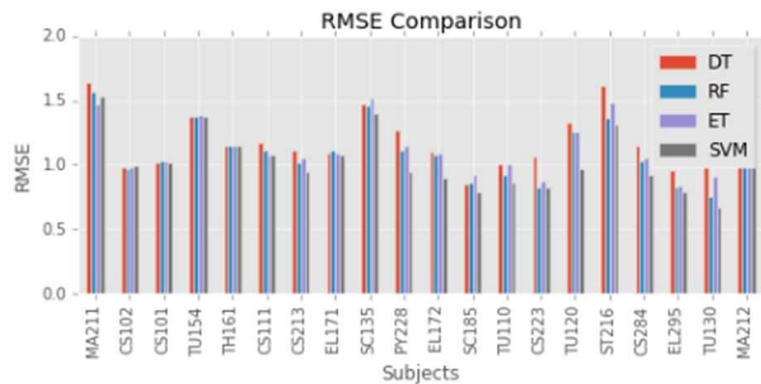
STUDENTID	TERM	COURSEID	GRADE
0001	2009/1	X	A
0001	2010/1	Y	B
0001	2010/2	Z	C
0002	2009/1	X	D
0002	2010/2	Z	F

*After appending*

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	-	-

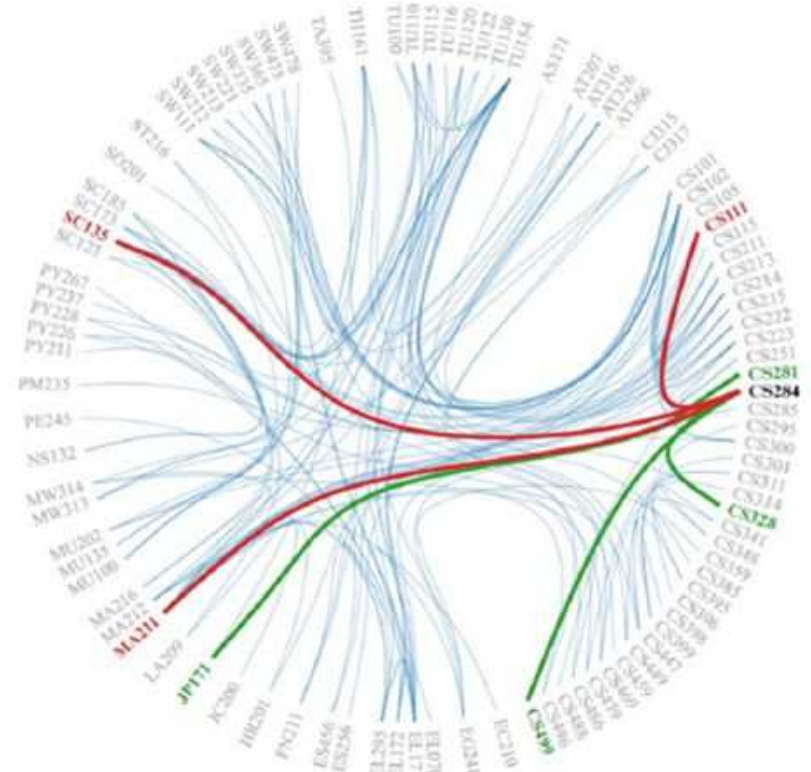
*appending prior results*

**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

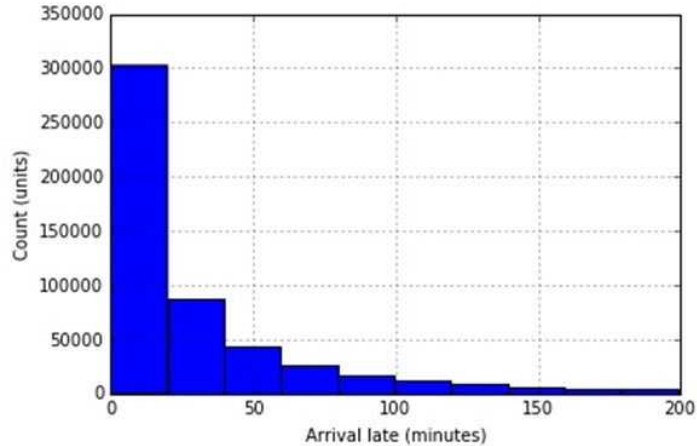
## Problem

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

## Result

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

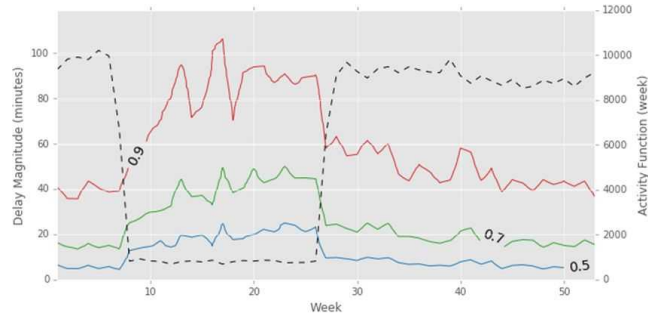
# Arrival Time Prediction and Train Tracking Analysis



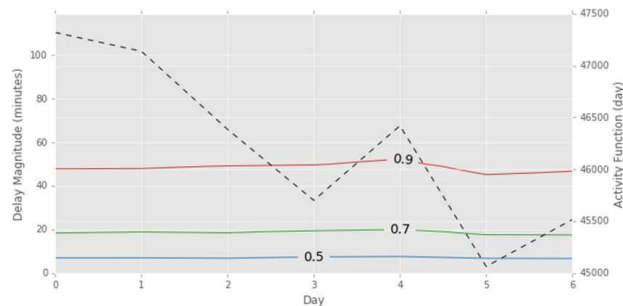
Ranking	Feature	FIS
1	<i>week</i>	0.273
2	<i>day</i>	0.117
3	<i>station</i>	0.113
4	<i>train_no</i>	0.107
5	<i>arrive_time</i>	0.072
6	<i>leave_time</i>	0.054
7	<i>arrive_cause</i>	0.049
8	<i>default_arrivetime</i>	0.041
9	<i>leave_cause</i>	0.041
10	<i>default_leavetime</i>	0.041

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

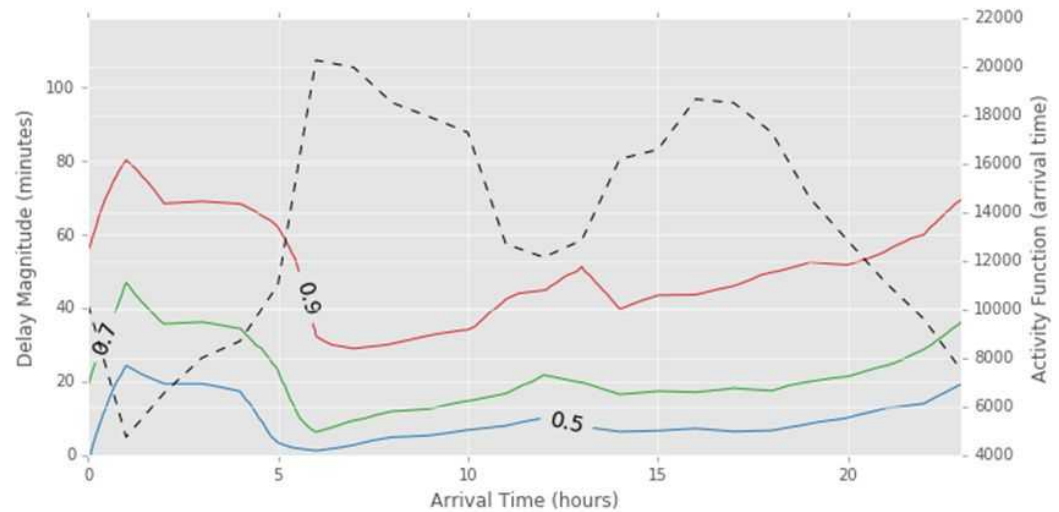
# Arrival Time Prediction and Train Tracking Analysis



Week 8-25 has significantly less number of services

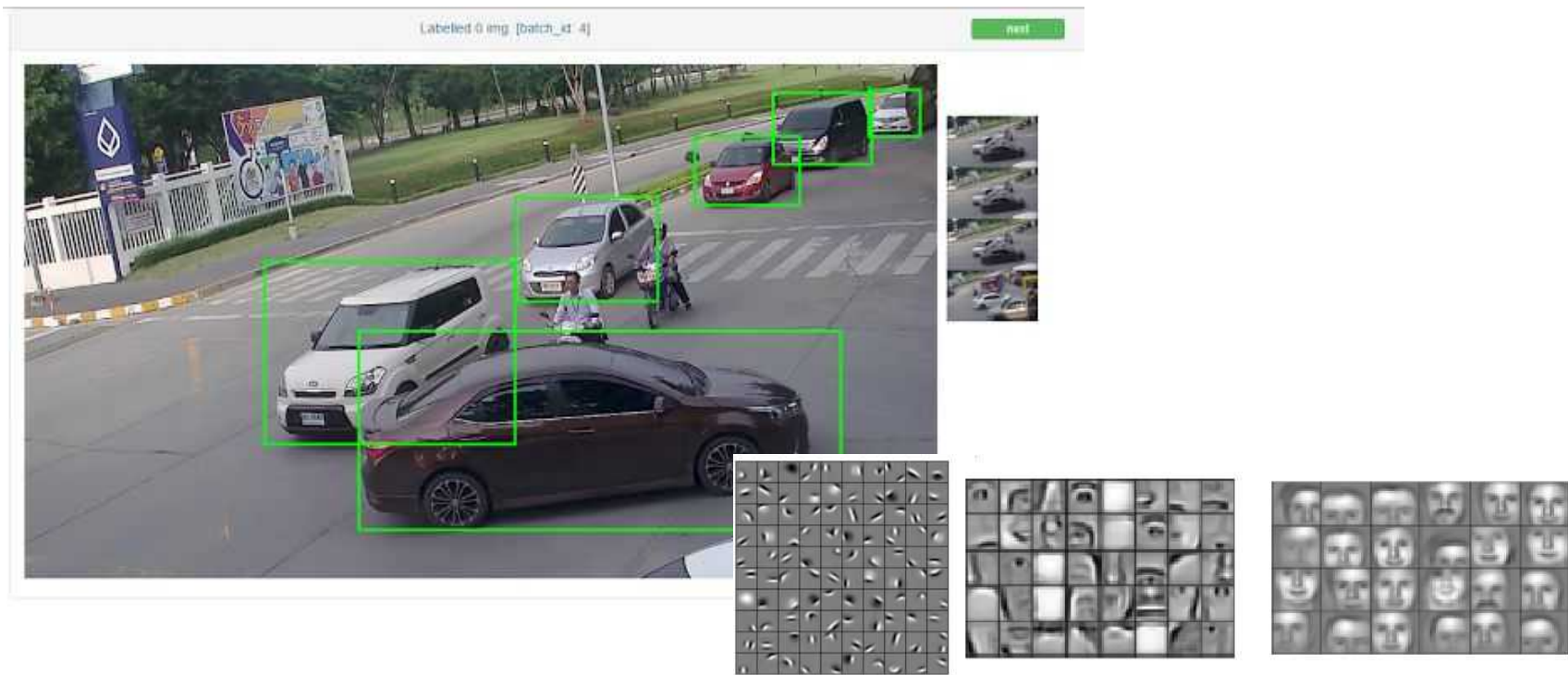


Thursday and Saturday have less number of passengers



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

# Deep Learning

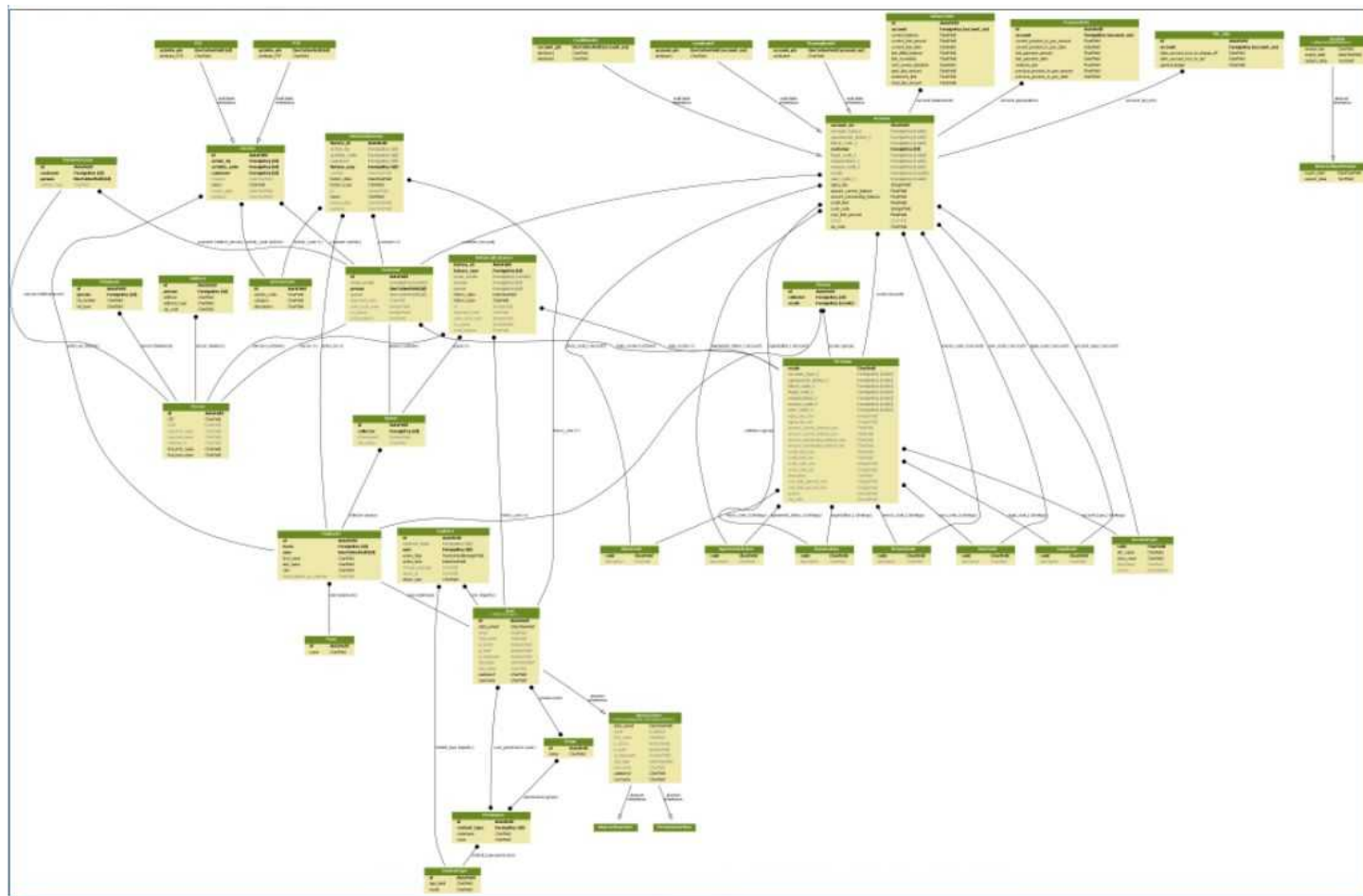


Convolution Neural Network

# Titan

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







# Software Development

DCS Admin

WELCOME ADMIN VIEW SITE / CHANGE PASSWORD / LOG OUT

Home » Main » S groups

Select s group to change

Action: 

-----

 Go 0 of 7 selected

ID	COLLECTOR	SCOPE
7	6: suptestfirst suptestlast	catch
6	4: 002_first_name 002_last_name	0006
5	4: 002_first_name 002_last_name	0005
4	4: 002_first_name 002_last_name	0004
3	3: 001_first_name 001_last_name	0003
2	7: Coltestfirst Coltestlast	01000
1	7: Coltestfirst Coltestlast	00500

7 s groups

Save

FILTER

By scope

All

catch

00500

01000

0003

0004

0005

0006

By collector

All

000\_first\_name 000\_last\_name

3: 001\_first\_name 001\_last\_name

4: 002\_first\_name 002\_last\_name

7: Coltestfirst Coltestlast

1: admin\_first\_name admin\_last\_name

5: supervisor\_first\_name supervisorlast\_name

6: suptestfirst suptestlast

bankctcnc, 131940900033

Balance Information

Strategy Code: :  
Current Due Amount 205.47 27/02/2017  
Past Due Amount  
Total Due Amount  
Next Review Date / Time

Current Balance  
Last Billed Balance  
Statement Limit  
Credit / Drawing Limit  
Limit Exceeded

CASE INFORMATION

Purpose  
Reason  
Payment Amount  
Remark

1 COURTESY VISIT  
2 UNABLE TO CONTACT  
--Payment Amount

Type/Creation Date  
Visit DateTime  
Sent By  
Submitted To

Request FCR  
16/05/2017 14:27  
22/05/2017 14:26  
1: admin\_first\_name admin\_last\_name  
1: admin\_first\_name admin\_last\_name

edit

ACTION LIST

Show 

▼

 entries

Search:

Action Date Time	Action Code	Remarks	Created By	Next Review
16/05/2017 14:27:06	FCR Request FCR		1: admin_first_name admin_last_name	21/05/2017 14:26
16/05/2017 14:23:50	NOSV NO Service		1: admin_first_name admin_last_name	17/05/2017 14:23
08/05/2017 17:55:00	PTP Promise to Pay Plan	ptp_memo	5: supervisor_first_name supervisorlast_name	08/05/2017 17:56

Showing 1 to 3 of 3 entries

First

Previous

1

Next

Last

Reject

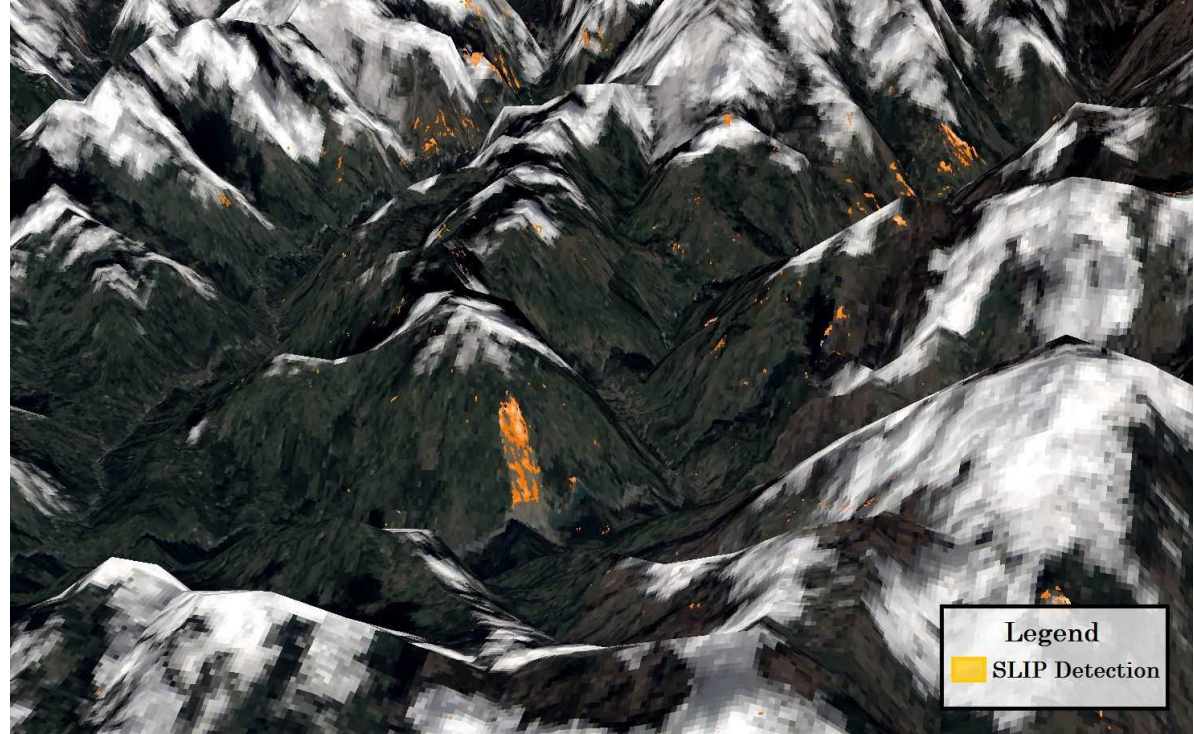
Approve

# 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/>

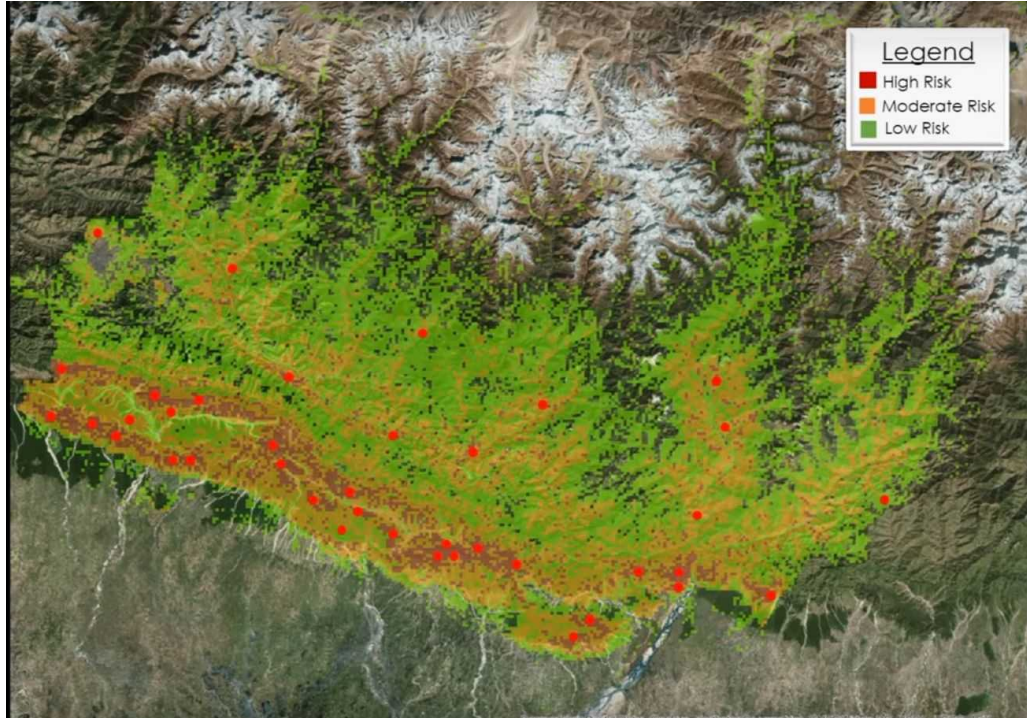
# Gorkha earthquake

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



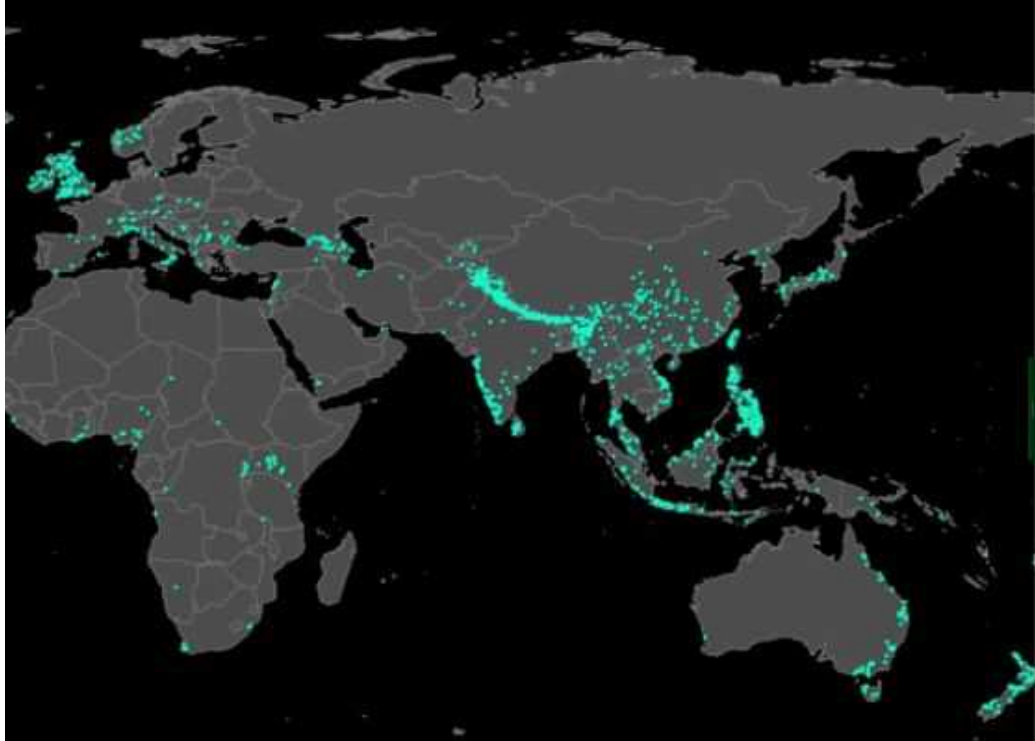


# Himalaya Disasters - NASA DEVELOP Fall 2015



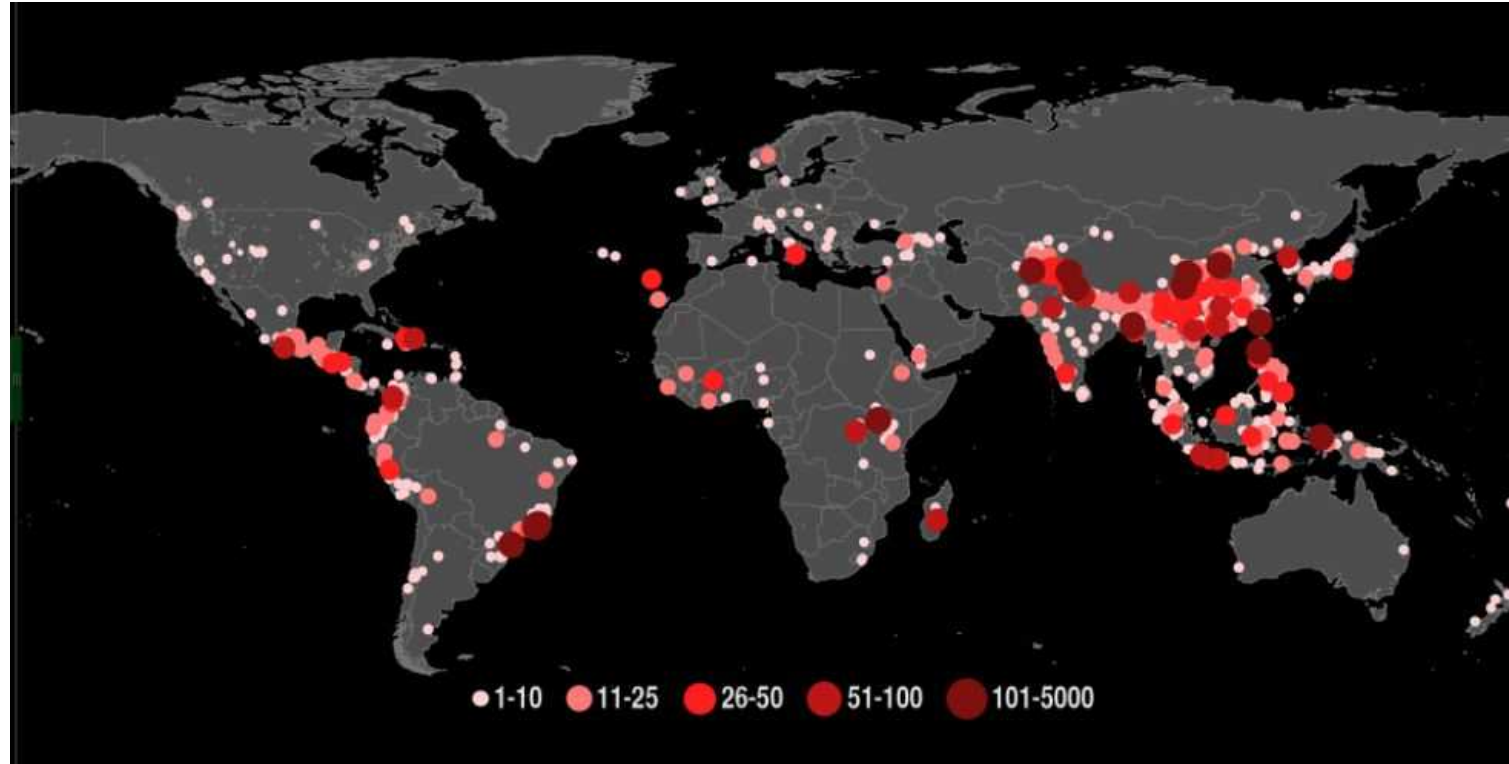
- Red band change means changing of vegetations
- $m = \text{NIR-SWIR}$
- Soil moisture change =  $m_{\text{new}}/m_{\text{old}}$
- SRTM topography image
- Slip detection
  - Moisture change >40%
  - Slope > 15%
- Future
  - Web platform to calibrate the algorithm

# NASA Global Landslide Catalog

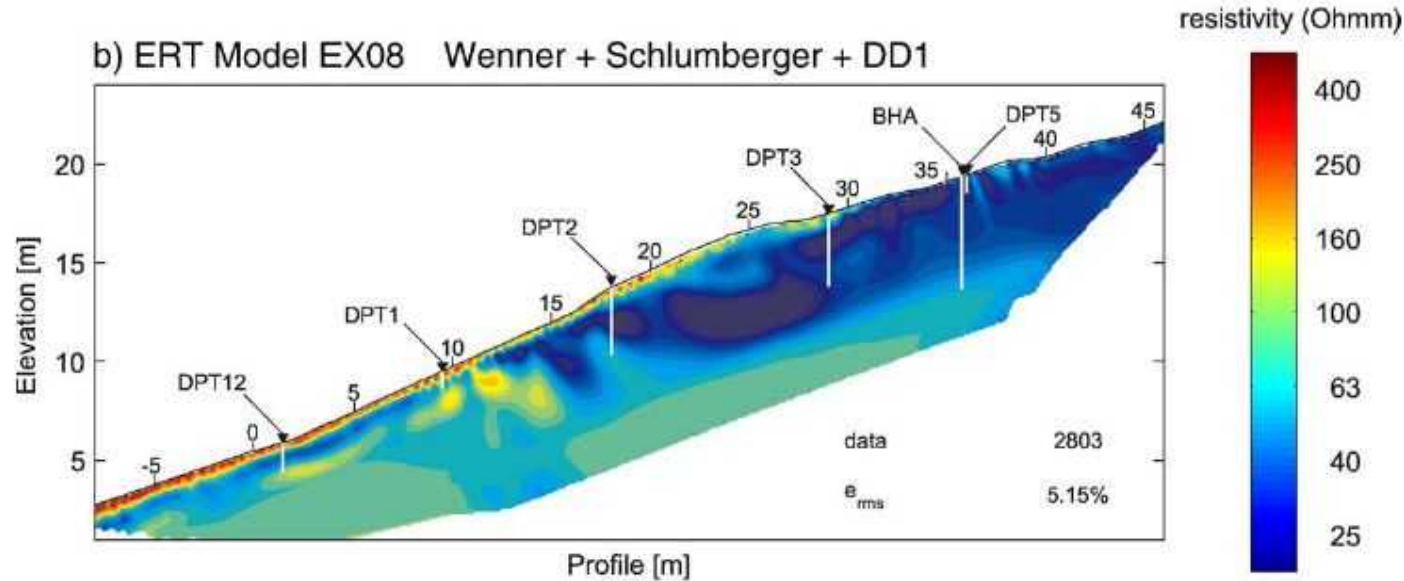


- Location and detail of previous landslides
- Mostly caused by Asian Monsoon

# Landslide casualties



# Electrical resistivity tomography (ERT)

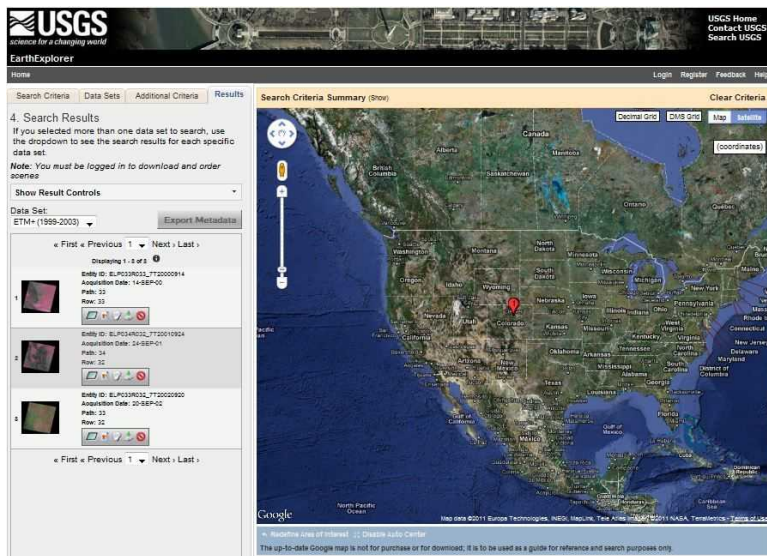


To convert from resistivity map to moisture map, expert need to identify structure geological layer

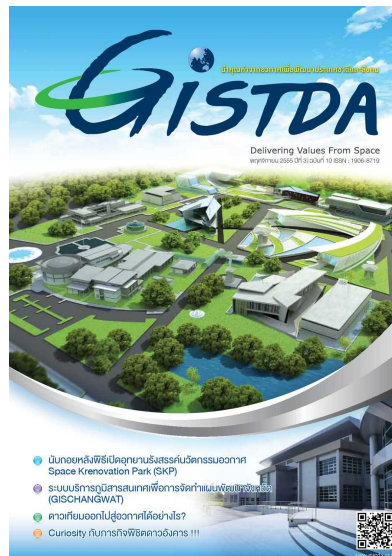
Friedel, Thielen and Springman 2006



# Satellite Image repository



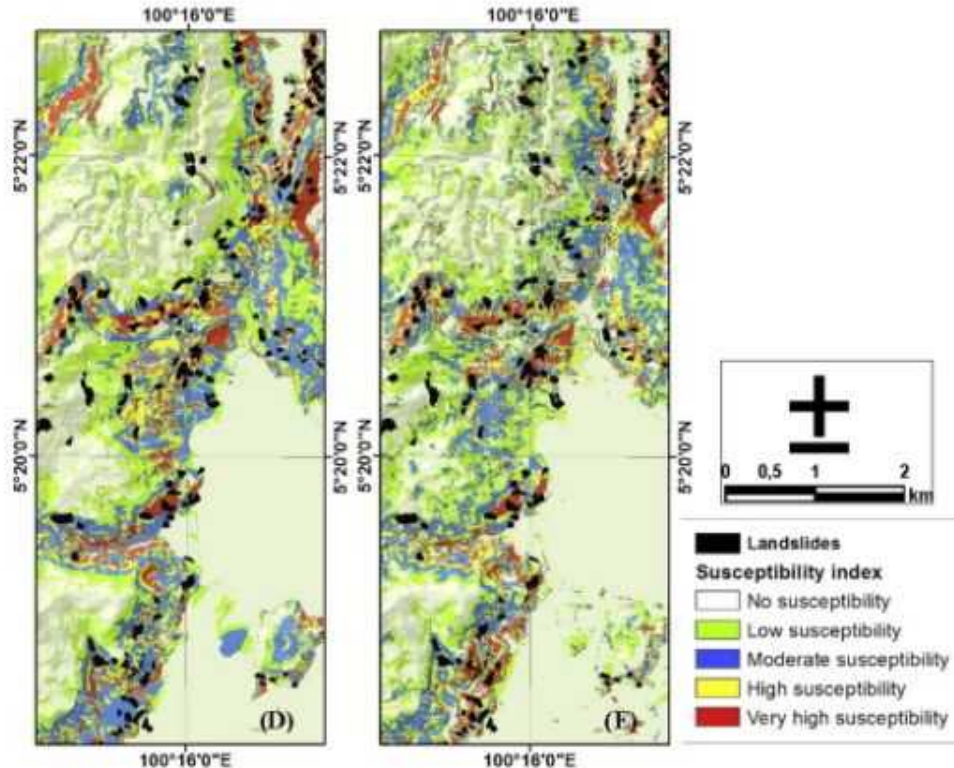
Earth Explorer



GISTDA

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

# Landslide susceptibility mapping



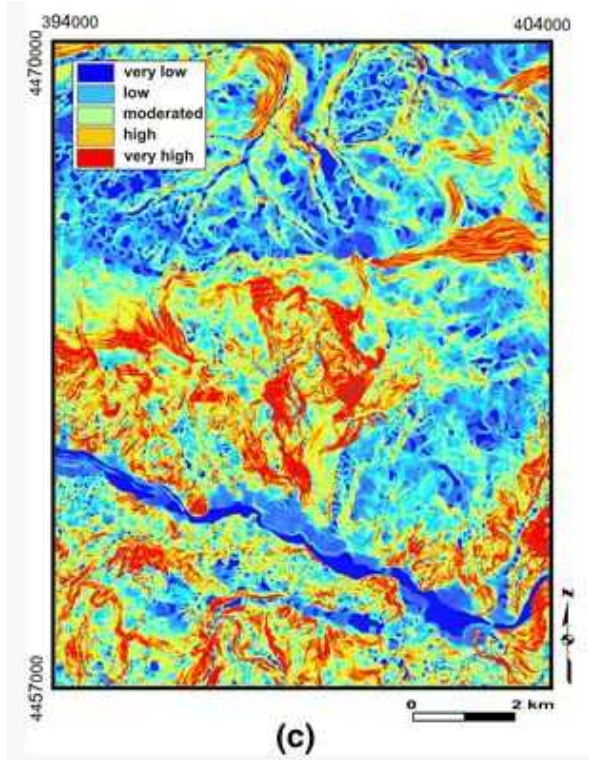
## Input Data

- 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%

# landslide susceptibility mapping



## Input Data

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

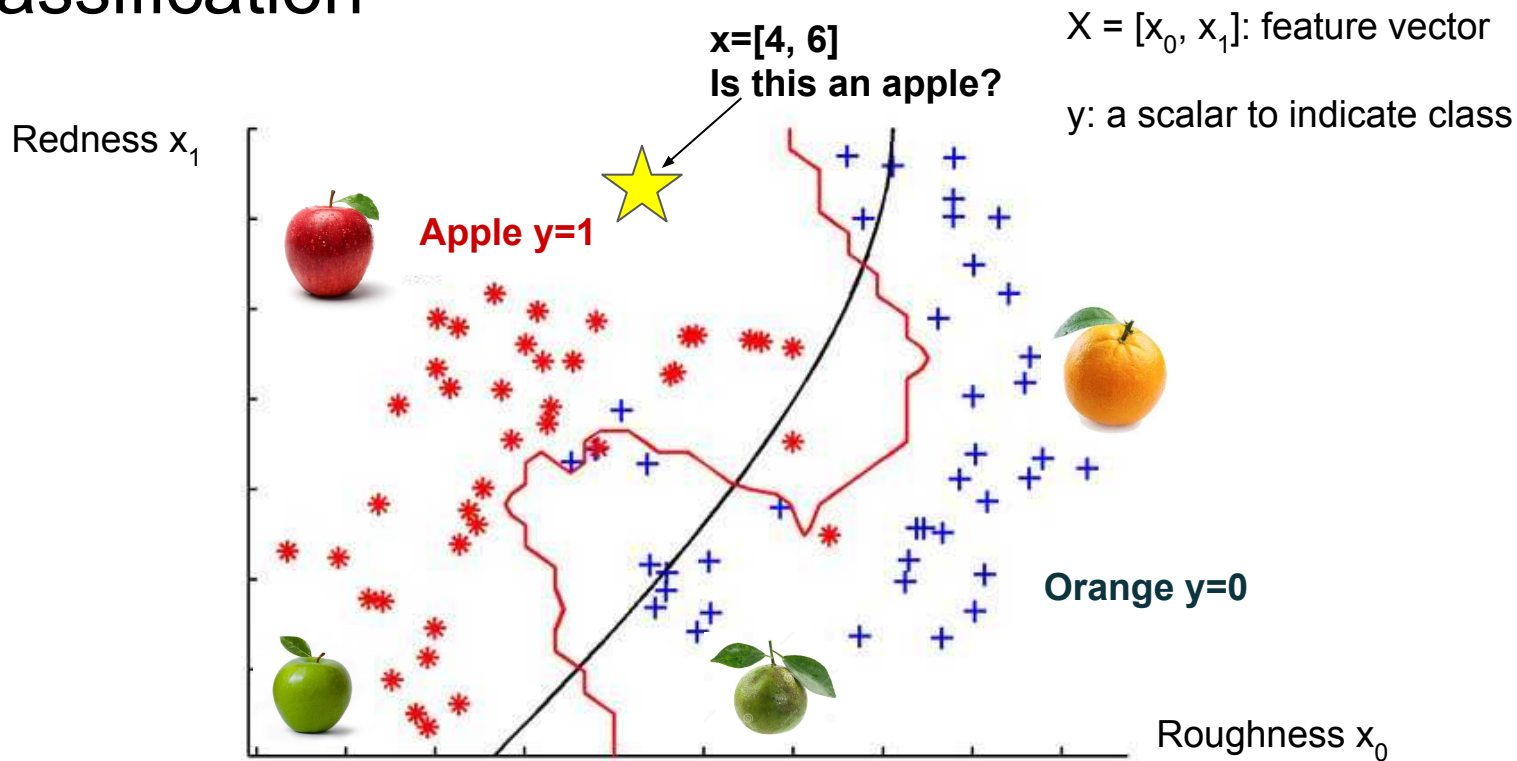
## Result using ANN

- Area under the ROC curve 84.6%

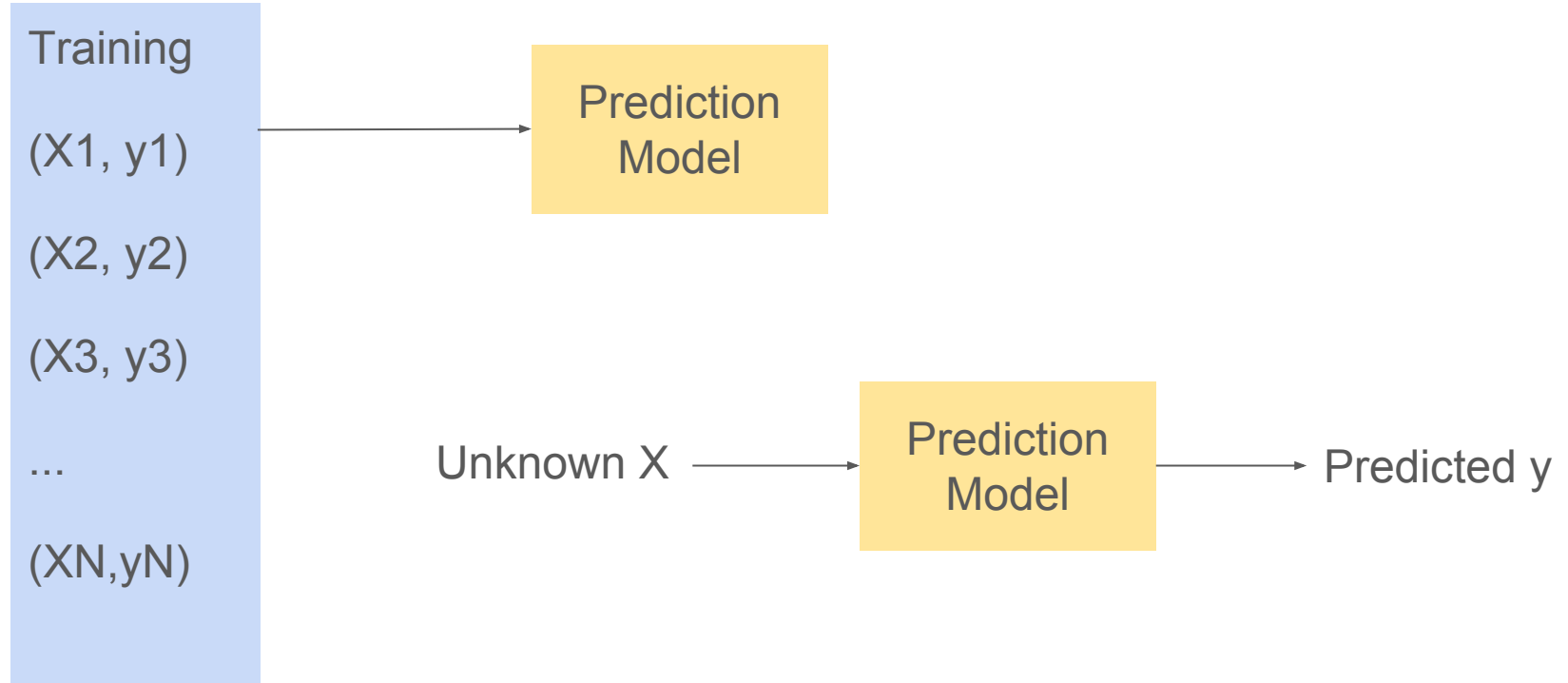
# Machine Learning

- Classification
- CNN
- RBM and Autoencoder

# Classification

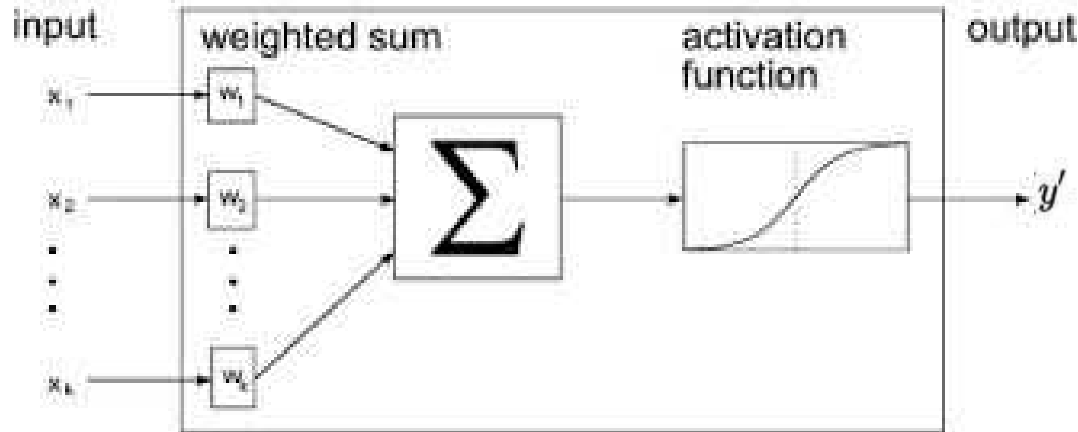


# Supervise Learning



# Neural network

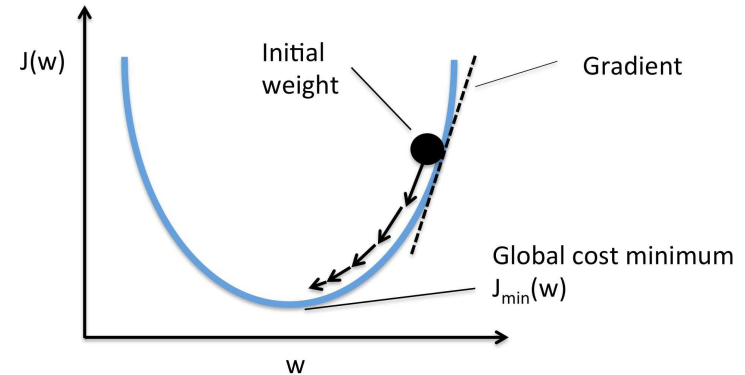
## Architecture of a single perceptron



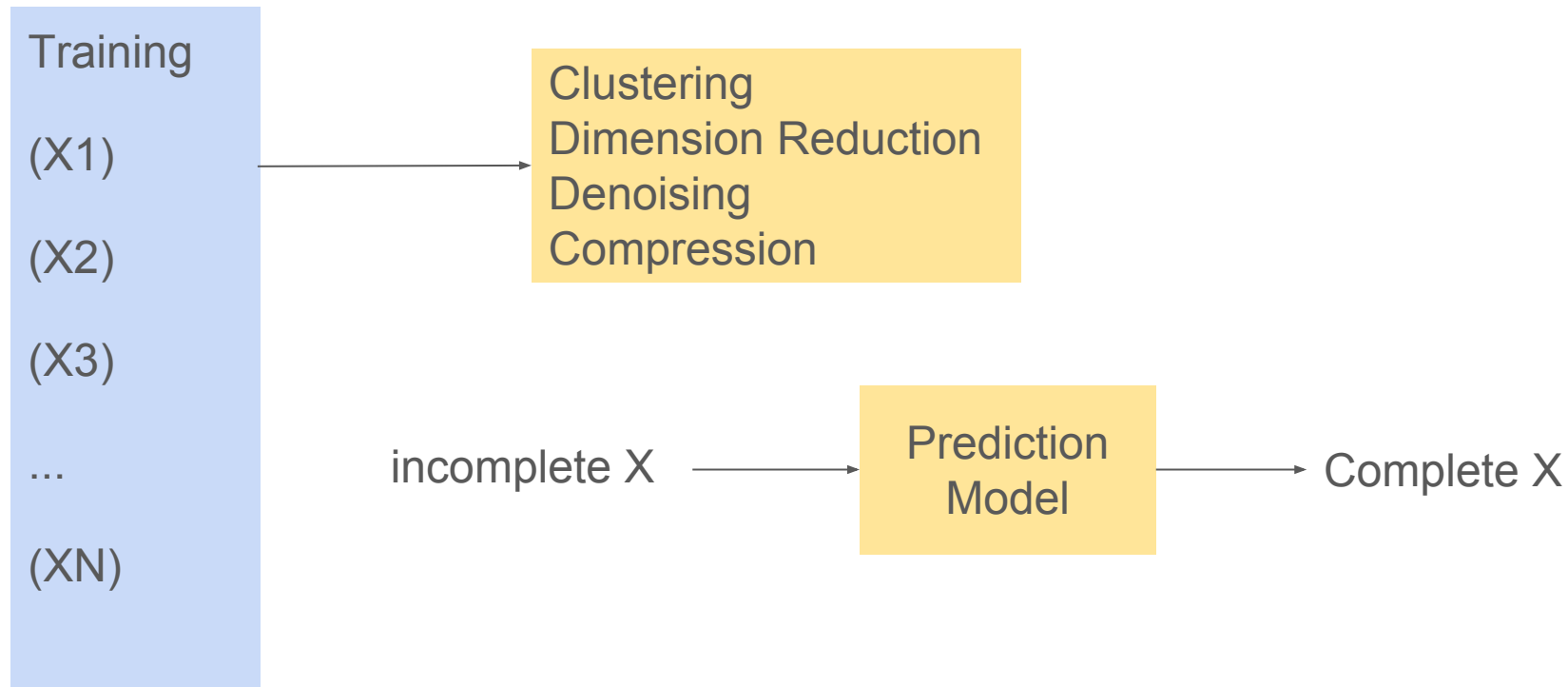
## Backpropagation

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

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



# Unsupervised Learning





# Clustering

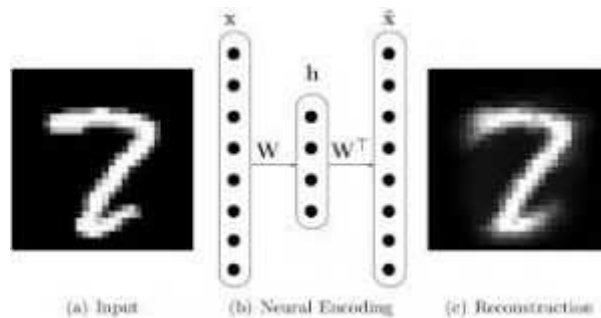
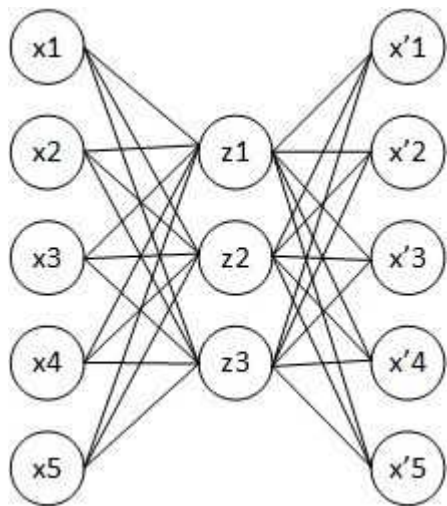


MNIST dataset

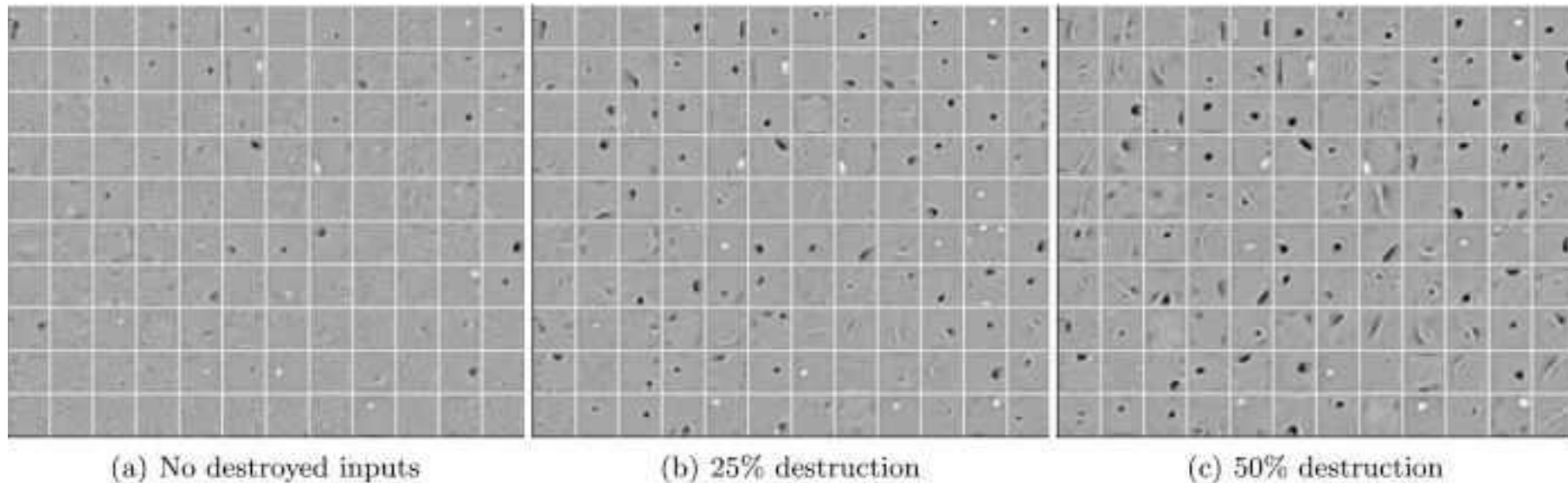
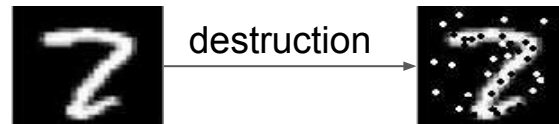
Automatically  
Organize



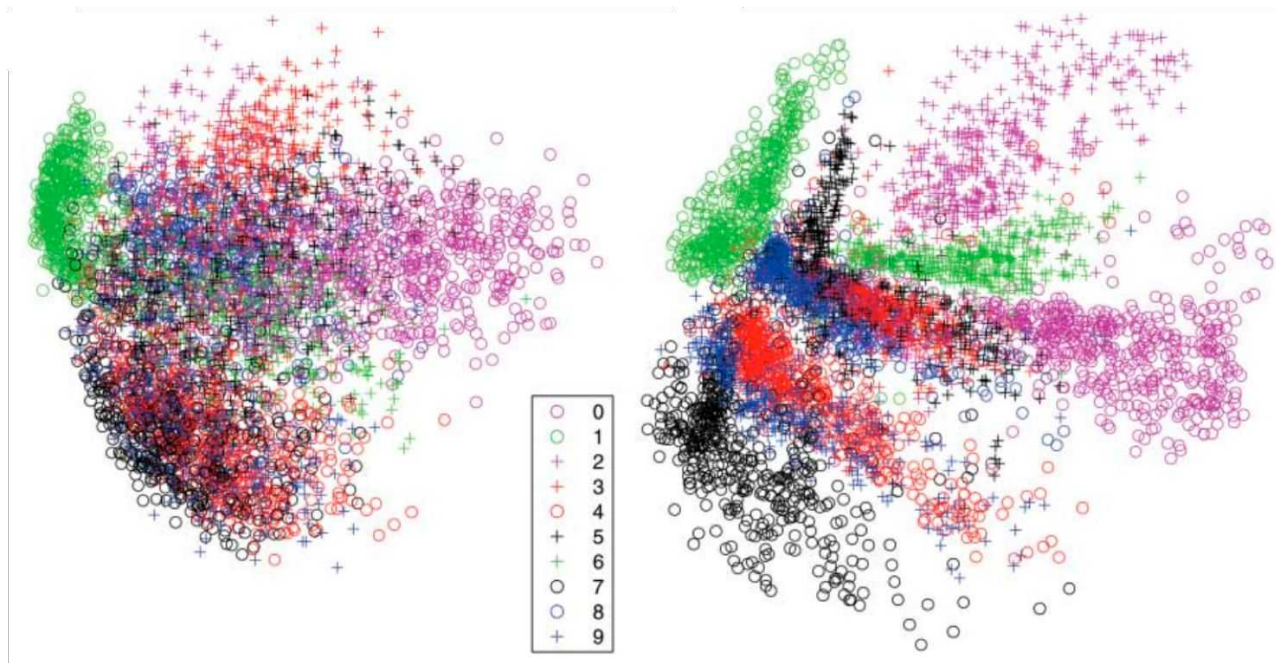
# Autoencoder



# Autoencoder with noisy training



# Autoencoder



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

# Content-Based Image Retrieval

**256-bit deep**



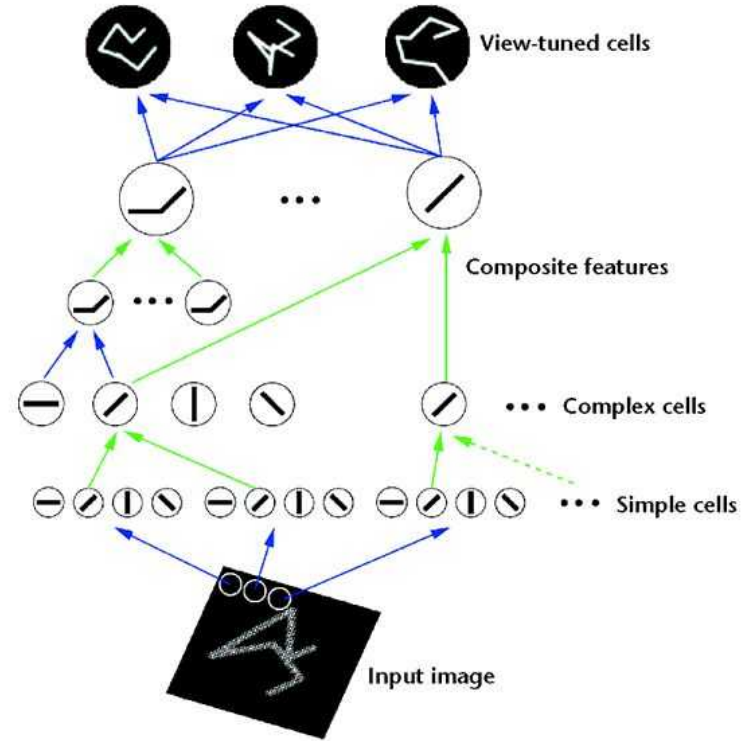
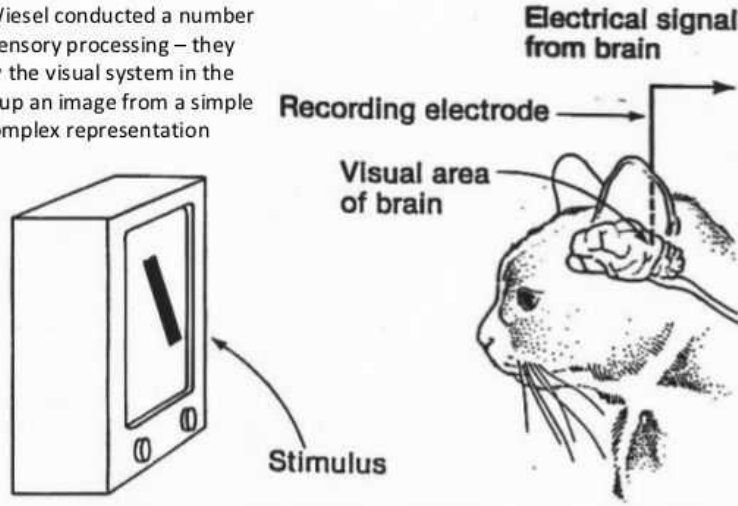
**Euclidean distance**



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

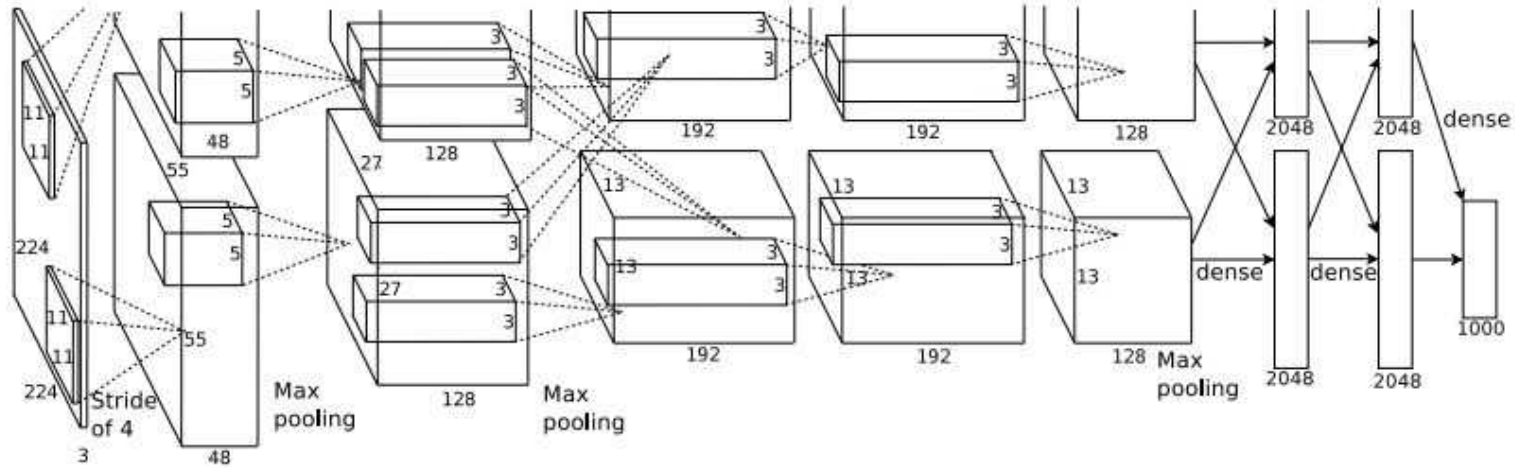
# Cat Brain (Hubel and Wiesel)

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



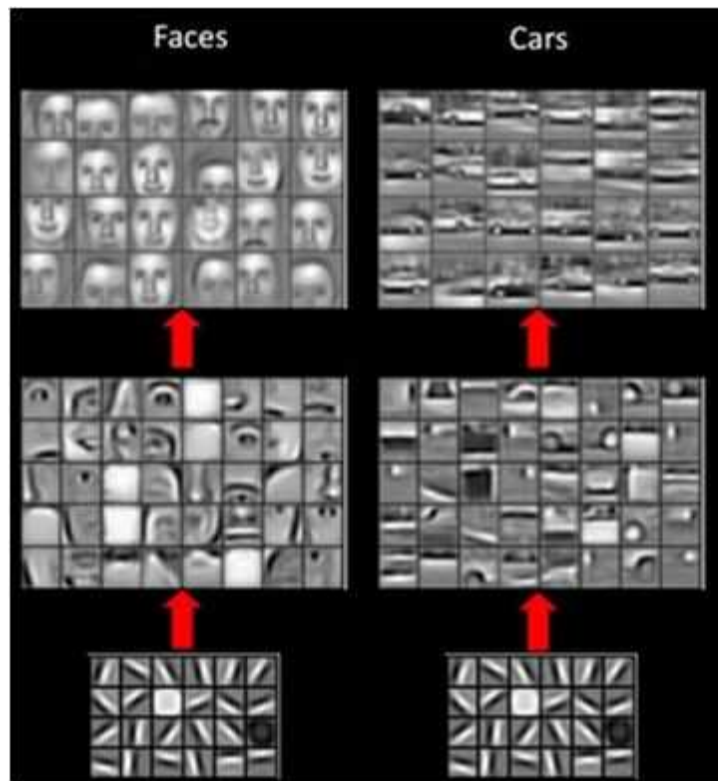
Krizhevsky, Sutskeve and Hinton 2012



# 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