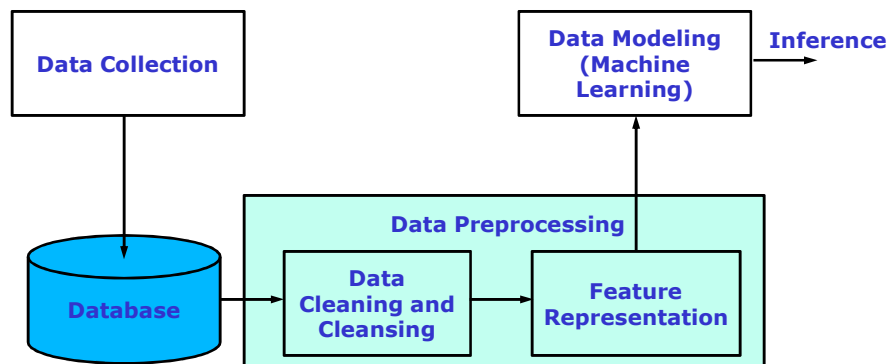# Data, Types of Data and Data Collection using Sensors

---

Data Science

- Multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insight from structured and unstructured data

- Central concept is gaining insight from data

- Machine learning uses data to extract knowledge
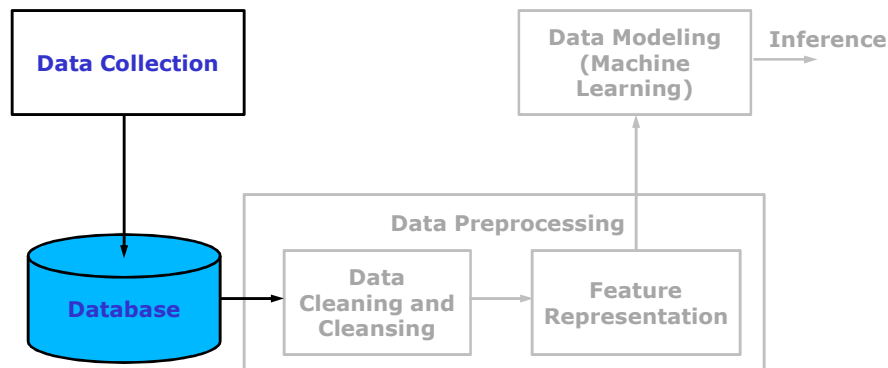

2

Data Science

- Multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insight from structured and unstructured data
- Central concept is gaining insight from data
- Machine learning uses data to extract knowledge

| Data Collection |

| Data Modeling (Machine Learning) | Inference |

| Database |

| Data Preprocessing |
| Data Cleaning and Cleansing | Feature Representation |

3

Data Collection

- Data manifests itself in many different forms
- Different forms of data require different ways to collect them and different storage solutions
- Collection of data may consists of sending out surveys, polls or doing other experiments
- Data based on the way it is collected:
  - Data that comes from surveys
    - Usually textual form of data or mixed
  - Data in the form of signals (comes from sensors)
    - Speech/Audio,
    - Images and videos
    - Temperature readings, Humidity
    - Seismic data
    - EEG (all bio-type signals) etc.
- According to the objective of the task, the way the data is collected will change

Types of Data: Based on Organization

- Unstructured data:
  - Rawest form of data
  - Example: Any type of files like texts, images, sounds or videos etc.
  - This type of data stored in a repository of files
    - Well organised directories on the computer hard drive
- Structured data:
  - It is a tabular data (rows and columns), which are very well defined
  - Stored in databases
    - Spreadsheets [Comma Separated Value (CSV) format]
    - Oracle
    - DB2
    - MySQL etc.

5

Types of Data: Based on Organization

- Semi-Structured data:
  - Anywhere between unstructured and structured data
  - A consistent format is defined, however there is no strict structure and parts of data may be incomplete or different type
  - Example: Data in the form of XML and JSON
    - Stored in document oriented databases

6

Types of Data: Based on Organization

• Semi-Structured data:

– Anyw ... ata
– A co ... ot strict struc ... lete or differ
– Exam
  • St

```xml
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>

  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>

  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>

  <book category="web">
    <title lang="en">XQuery Kick Start</title>
    <author>James McGovern</author>
    <author>Per Bothner</author>
    <author>Kurt Cagle</author>
```

7

Type of Data: Based on Variables (Value) found in Data

• Mainly in Structured Data:

• Numerical data:

– Data represented as numbers
– Data in which information is measurable
– This type of data is called quantitative data as its describes a quantity
– Two types:
  • Continuous valued data:
    – There is no limit on the range of the values
    – Example: Cost of the books, temperature etc.
  • Discrete valued data:
    – There is a specific limit on the range of the values
    – Example: number of members of family, number of days in a month etc.

8

Type of Data: Based on Variables (Value) found in Data

- Categorical data:
  - Data that is not a number. It can be string of text or date
  - It describe an item or event to one of few different categories
  - Example: Ethnicity, gender, eye colour, etc.
  - This type of data is called qualitative data as its describes a quality
  - Three types values they hold:
    - Ordinal values: Values that have a set order to them
      - Example: Severity of a alarm as "Critical", "Medium" and ""Low", Ranking of a race as " First", Second", Third" …
    - Nominal values: Values that have no set order to them
      - Example: Values for the variables "Marital Status", "Country" etc.
    - Binary values: Special type of categorical data
   - Have only two values – "Yes" and "No" OR "True" and "False" OR "1" and "0" **9**

Type of Data: Based on Variables (Value) found in Data

- Time series data:
  - It involve time and some kind of value
  - Example: Temperature at every hour
  - It is clearly structured and numeric in nature
  - Special case of numerical data
  - This type of data is important because of IoT and sensors
  - Data from sensors are almost always time-series in nature

Date/ Time	Temperature (C)/ Humidity (%)	Pressure (Pa)	Rain (inches)	Light Intensity (lux)	Accelerations (g)	Force (N)	Moisture (%)
2017-09-06 18:44:32	23.00,56.00	617.64	0.01	3	0.52,0.31,-0.80,0.00,0.00,0.00,31.36,-159.01	0.02	81.00
2017-09-06 18:33:32	24.00,58.00	619.47	0.01	12	0.52,0.30,-0.79,0.00,0.00,0.00,31.45,-159.12	0.02	82.00
2017-09-06 18:22:39	24.00,58.00	623.37	0.00	71	0.52,0.31,-0.80,0.00,0.00,0.00,31.35,-158.88	0.02	83.00
2017-09-06 18:11:31	25.00,60.00	627.02	0.05	194	0.51,0.31,-0.80,0.00,0.00,0.00,30.80,-159.00	0.02	81.00

Data Collection from Sensors

- Sensors are the devices that respond to the environment around it and convert the physical parameters into a signal (e.g., optical, electrical, mechanical) suitable for processing

Surrounding environment → Sensor → **signal** electrical, optical or mechanical
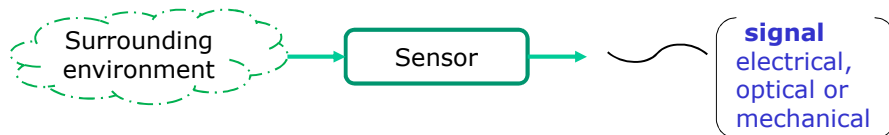
- ***Example***: a temperature sensor outputs an electrical signal whose voltage or current can be used to identify the temperature around it
- Sensors can be an electrical/mechanical component, a module or a subsystem

11

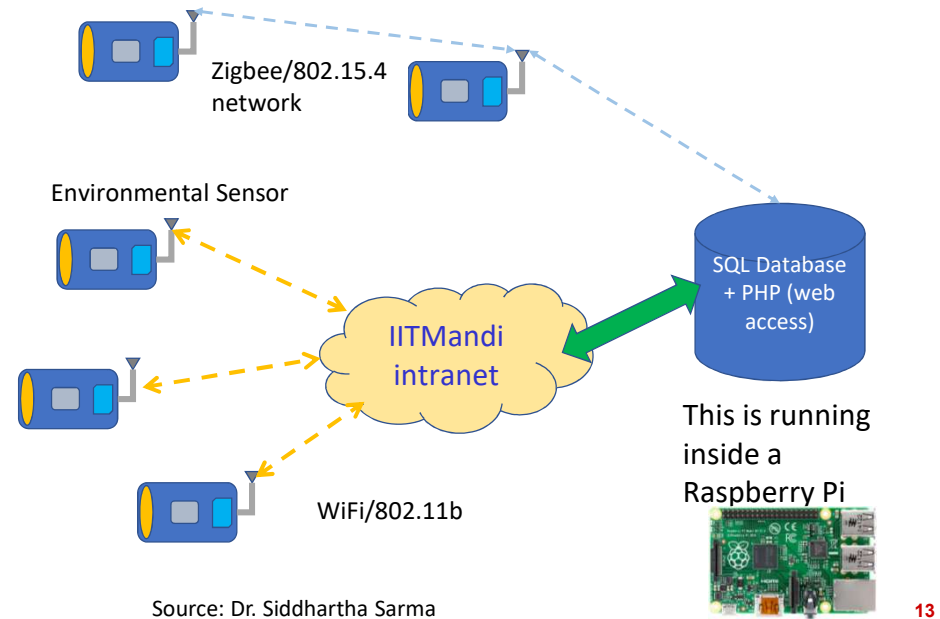## Different Types of Sensors

- Acoustic, sound sensors (e.g., microphone)

- Visual sensors (e.g. cameras)

- Environmental sensors (e.g., temperature, humidity, pressure etc.)

- Chemical sensors (e.g., NOx sensors)

- Flow sensors (e.g., water flow sensors)

- Motion sensors (e.g., gyroscope)

- Proximity or presence sensor (e.g., Passive Infrared (PIR))

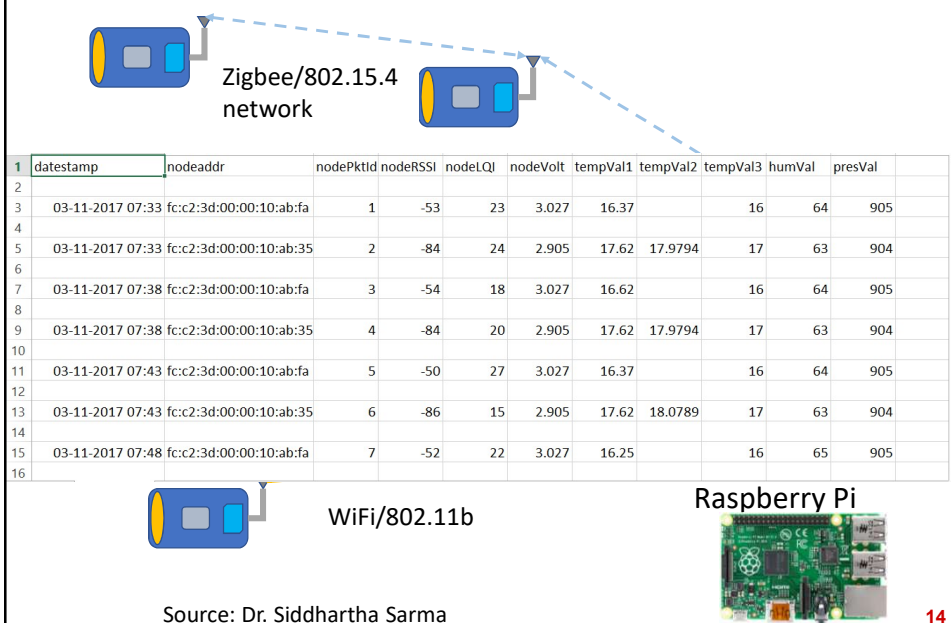- Biosensors (e.g., glucose monitor)

12

High-Level Overview: Environmental Data (Temperature, Humidity, Pressure etc) Collection
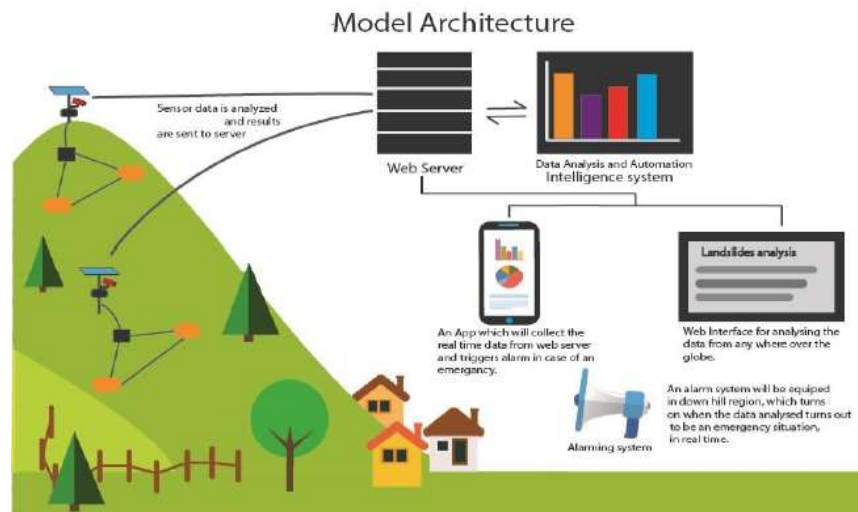
Zigbee/802.15.4 network

Environmental Sensor

IITMandi intranet

SQL Database + PHP (web access)

This is running inside a Raspberry Pi

WiFi/802.11b

Source: Dr. Siddhartha Sarma

13

High-Level Overview: Environmental Data (Temperature, Humidity, Pressure etc) Collection

Zigbee/802.15.4 network

	datestamp	nodeaddr	nodePktId	nodeRSSI	nodeLQI	nodeVolt	tempVal1	tempVal2	tempVal3	humVal	presVal	
1												
2												
3	03-11-2017 07:33	fc:c2:3d:00:00:10:ab:fa	1	-53	23	3.027	16.37		16	64	905	
4												
5	03-11-2017 07:33	fc:c2:3d:00:00:10:ab:35	2	-84	24	2.905	17.62	17.9794	17	63	904	
6												
7	03-11-2017 07:38	fc:c2:3d:00:00:10:ab:fa	3	-54	18	3.027	16.62		16	64	905	
8												
9	03-11-2017 07:38	fc:c2:3d:00:00:10:ab:35	4	-84	20	2.905	17.62	17.9794	17	63	904	
10												
11	03-11-2017 07:43	fc:c2:3d:00:00:10:ab:fa	5	-50	27	3.027	16.37		16	64	905	
12												
13	03-11-2017 07:43	fc:c2:3d:00:00:10:ab:35	6	-86	15	2.905	17.62	18.0789	17	63	904	
14												
15	03-11-2017 07:48	fc:c2:3d:00:00:10:ab:fa	7	-52	22	3.027	16.25		16	65	905	
16												

WiFi/802.11b

Raspberry Pi

Source: Dr. Siddhartha Sarma

14

Land Slide Monitoring System (LMS)

- LMSs that rely on Internet of Things (IoT) and low-cost Micro-Electro-Mechanical Systems (MEMS) sensors

Source: Dr. Varun Dutt

15

Components of LMS

GY 61 Accelerometer Sensor — Pin Diagram of GY-61 — YL 69 Soil Moisture Sensor — SIM 900A GSM Module — Force Sensor

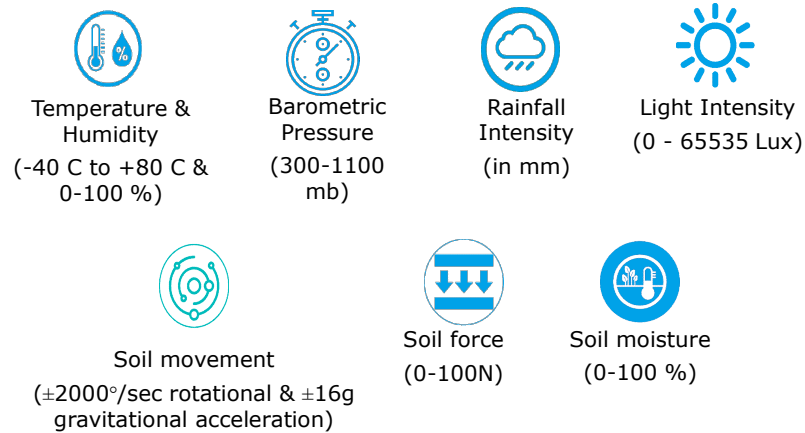Humidity Sensor DHT 22 — Light Sensor BH-1750 — Temperature and Pressure Sensor BMP-180 — Tipping Rain Gauge

Source: Dr. Varun Dutt

16

Architecture and Features of LMS

- The LMS monitors a number of weather and soil parameters via sensors on deployment location

Temperature & Humidity

(-40 C to +80 C & 0-100 %)

Barometric Pressure

(300-1100 mb)

Rainfall Intensity

(in mm)

Light Intensity

(0 - 65535 Lux)

Soil movement

(±2000°/sec rotational & ±16g gravitational acceleration)
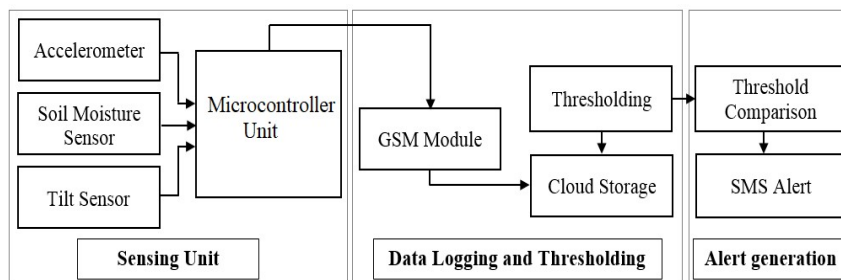
Soil force

(0-100N)

Soil moisture

(0-100 %)

Source: Dr. Varun Dutt

17

Architecture and Features of LMS

- The LMS monitors a number of weather and soil parameters via sensors on deployment location

Architecture diagram of LMS

Accelerometer		
Soil Moisture Sensor	Microcontroller Unit	GSM Module
Tilt Sensor		

Thresholding → Threshold Comparison

Cloud Storage

SMS Alert

Sensing Unit | **Data Logging and Thresholding** | **Alert generation**

The LMS will alert people via traffic lights, SMSs, or smart-apps on mobile phones about the danger of impending landslides

Source: Dr. Varun Dutt

18

Architecture and Features of LMS

- The LMS monitors a number of weather and soil parameters via sensors on deployment location

Architecture diagram of LMS

Accelerometer

Date/ Time	Temperature (C)/ Humidity (%)	Pressure (Pa)	Rain (inches)	Light Intensity (lux)	Accelerations (g)	Force (N)	Moisture (%)
2017-09-06 18:44:32	23.00,56.00	617.64	0.01	3	0.52,0.31,-0.80,0.00,0.00,0.00,31.36,-159.01	0.02	81.00
2017-09-06 18:33:32	24.00,58.00	619.47	0.01	12	0.52,0.30,-0.79,0.00,0.00,0.00,31.45,-159.12	0.02	82.00
2017-09-06 18:22:39	24.00,58.00	623.37	0.00	71	0.52,0.31,-0.80,0.00,0.00,0.00,31.35,-158.88	0.02	83.00
2017-09-06 18:11:31	25.00,60.00	627.02	0.05	194	0.51,0.31,-0.80,0.00,0.00,0.00,30.80,-159.00	0.02	81.00

Source: Dr. Varun Dutt

19