Machine Learning

LG전자 고급 데이터 사이언티스트 양성 과정 3기

이영기

서울대학교 컴퓨터공학부

Success is not final. Failure is not fatal
It is the courage to continue that counts.   Winston
Churchill

Quick Introduction

- Name: Youngki Lee

- Office: 301-413

- Phone: 02-880-1726

- E-mail: youngkilee@snu.ac.kr

- Research Homepage:

  http://youngkilee.blogspot.com

Quick Introduction

Experiences

- 2020-Now: Associate Professor, Department of CSE, Seoul National University
- 2018-2020: Assistant Professor, Department of CSE, Seoul National University
- 2013-2018: Assistant Professor, School of Information Systems, Singapore Management University

Research Lab

- Human-centered computer systems lab (https://hcs.snu.ac.kr/people/) in computer science and engineering dept.
- 6 PhD students, 2 Maters students, and 5+ undergraduate research interns

Research Interests

- Mobile and Embedded ML/DL Systems
- Behavior and Context sensing
- Future Media Systems
- IoT and Edge Systems

Research Activities

- 30↑ top conferences papers (ACM MobiSys, ACM MobiCom, ACM UbiComp, ACM SenSys, ACM CHI)
- Program Co-Chair and General Co-Chair, ACM UbiComp 2018
- Steering committee, ACM UbiComp, 2019 - Now
- Technical Program Committee (ACM MobiSys, ACM MobiCom, ACM SenSys, ACM UbiComp, IEEE INFOCOM)

Agenda for Today

- Introduction to the class

- Introduction to the machine learning

- Introduction to the statistical learning

  ✔ Chapter 2 of our textbook

- Lab for the chapter 2

My Areas of Research
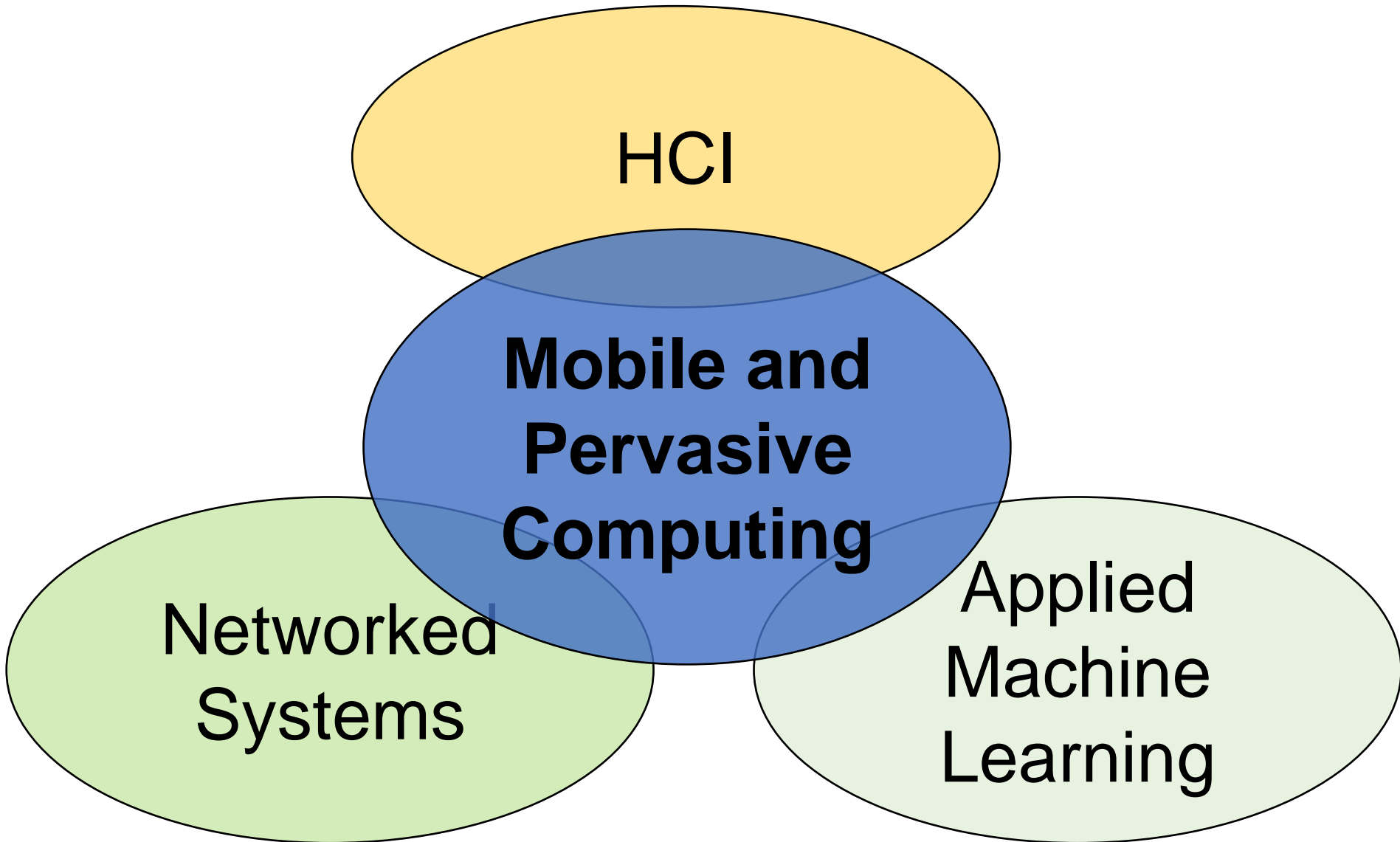
Research Intro: Mobile Computing
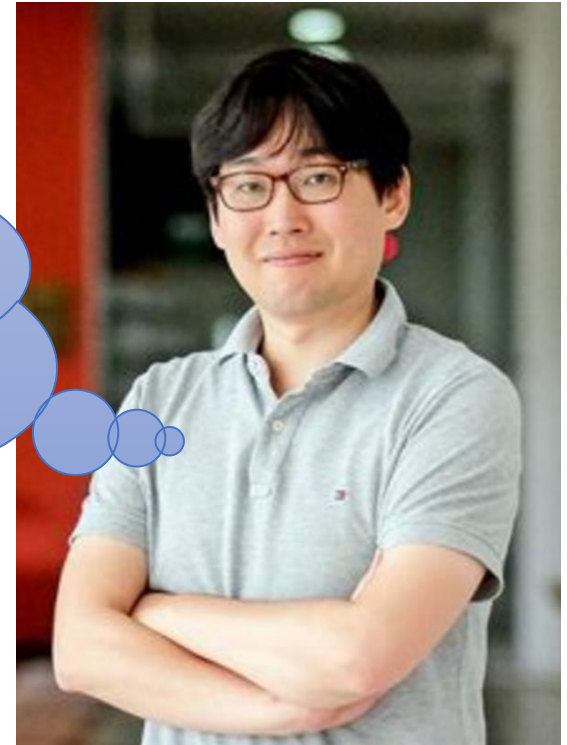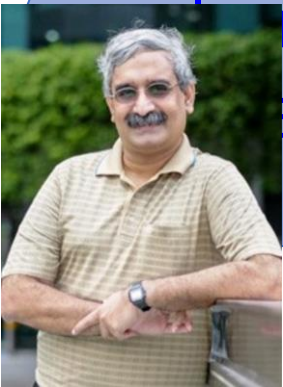
"Mobile" "Computer" or Beyond?

Group-Aware Mobile Ads

- Location-based mobile ads are often of no use.

- Group-aware promotion to satisfy the entire group. (used in Resort World Sentosa in Singapore).

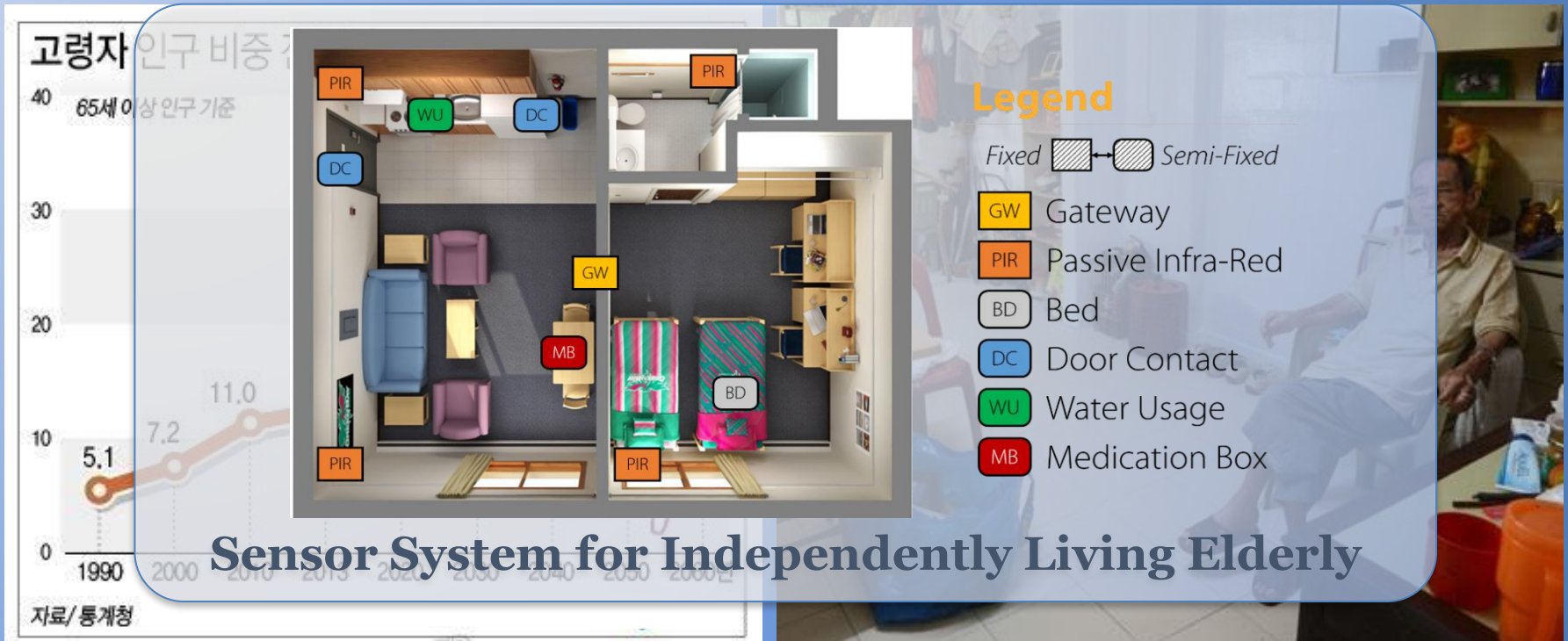Korean BBQ: 50% off for lunch today!!! Wow, this is a great deal. But...

- Based on our group detection system, GruMon [SenSys 14].

Independent Living Assistant

Not only to enhance physical wellness,
but also for socially and mentally healthy life

Sensor System for Independently Living Elderly

Legend

Fixed ⟷ Semi-Fixed

GW	Gateway
PIR	Passive Infra-Red
BD	Bed
DC	Door Contact
WU	Water Usage
MB	Medication Box

- On-going project at SMU (initial testbed being deployed to 100 elderly who live alone)

Life-Immersive Mobile Computing

Sense human behavior, emotion, and surrounding contexts

↓

Extract useful insights and knowledge

↓

Provide what people need right on time & place

Sleep Quality Monitoring

Pothole Monitoring

Location-aware Alarms

Physical Activity Diary

Bus Stop Queue Estimation

Proactive Advertisement

Key Building Block: Context Sensing

Comprehensive/ detailed behavior

- ✔ Centimeter-level indoor localization

- ✔ Eating
- ✔ Smoking
- ✔ Shopping
- ✔ Dancing
- ✔ Drumming

- ✔ Turn-takings
- ✔ Linguistic contents
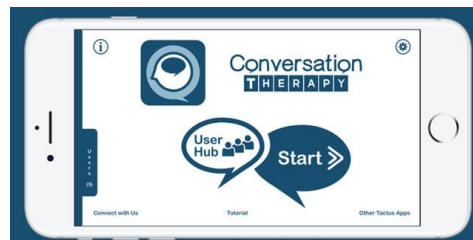- ✔ Emotional expressions

Generic external context

Location

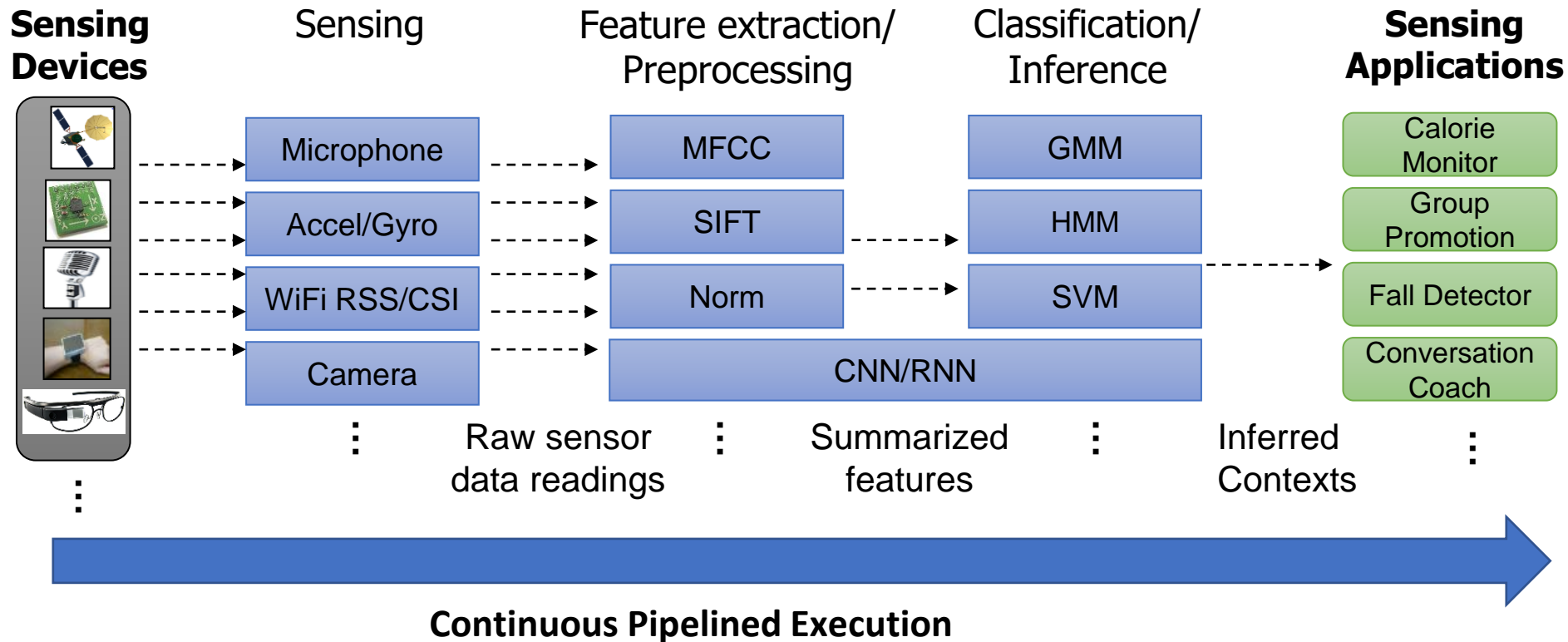Physical Activity

Conversation

Internal States

- ✔ Heartrate
- ✔ Stress
- ✔ Mood
- ✔ Sleep quality
- ✔ Distractibility
- ✔ Intention
- ✔ Engagement
- ✔ Attention
- ✔ Mindfulness
- ✔ Emotion
- ✔ Anxiety
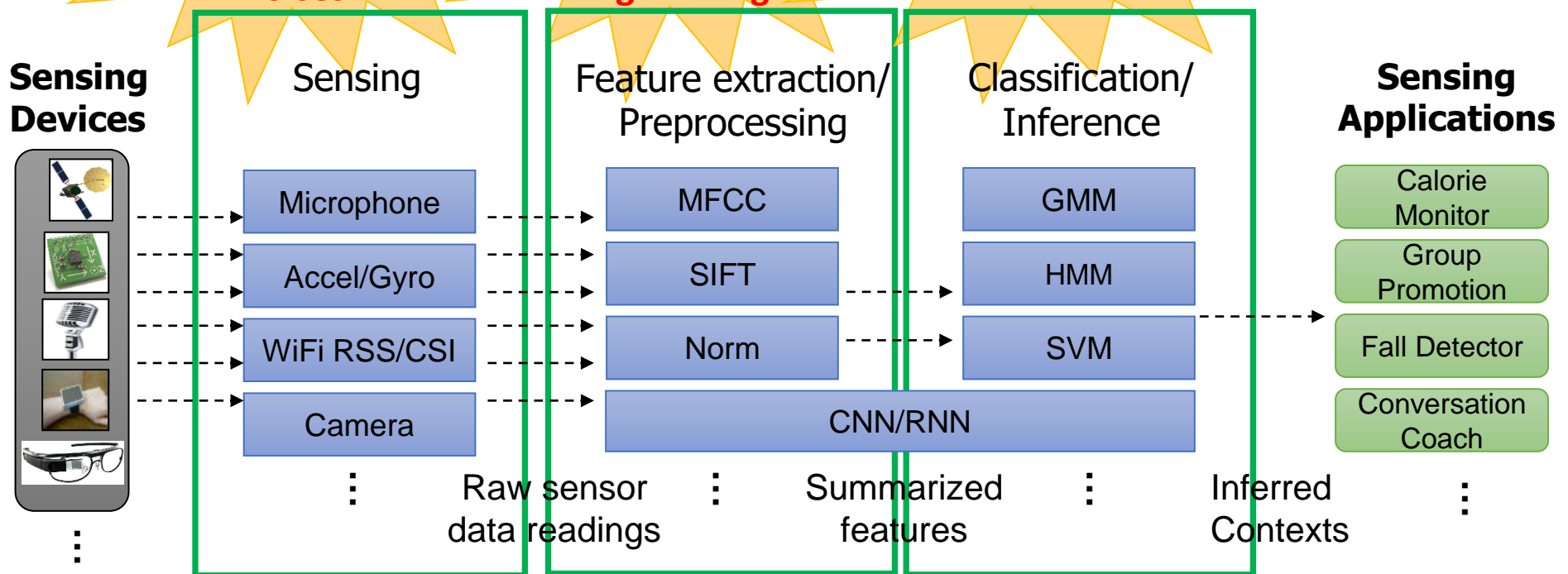- ✔ Depression
- ✔ Boredom
- ✔ Fatigue
- ✔ …

Common Computational Flow

Continuous sensing and analytics of user activities, location, emotions, and surroundings with mobile/IoT/wearable devices

Sensing Devices | Sensing | Feature extraction/ Preprocessing | Classification/ Inference | **Sensing Applications**

Sensing:
- Microphone
- Accel/Gyro
- WiFi RSS/CSI
- Camera

Feature extraction/Preprocessing:
- MFCC
- SIFT
- Norm
- CNN/RNN

Classification/Inference:
- GMM
- HMM
- SVM

Sensing Applications:
- Calorie Monitor
- Group Promotion
- Fall Detector
- Conversation Coach

Raw sensor data readings

Summarized features

Inferred Contexts

Continuous Pipelined Execution

Challenge 1: Inference Accuracy

- > 90% accuracy is extremely challenging.

- Errors are caused in each layers

Uncalibrated readings, and noises

Challenges in Feature Engineering

Non-Representative Model

Sensing Devices

Sensing

Feature extraction/ Preprocessing

Classification/ Inference

Sensing Applications

Sensing	Feature extraction/ Preprocessing	Classification/ Inference	Sensing Applications
Microphone	MFCC	GMM	Calorie Monitor
Accel/Gyro	SIFT	HMM	Group Promotion
WiFi RSS/CSI	Norm	SVM	Fall Detector
Camera	CNN/RNN		Conversation Coach

Raw sensor data readings

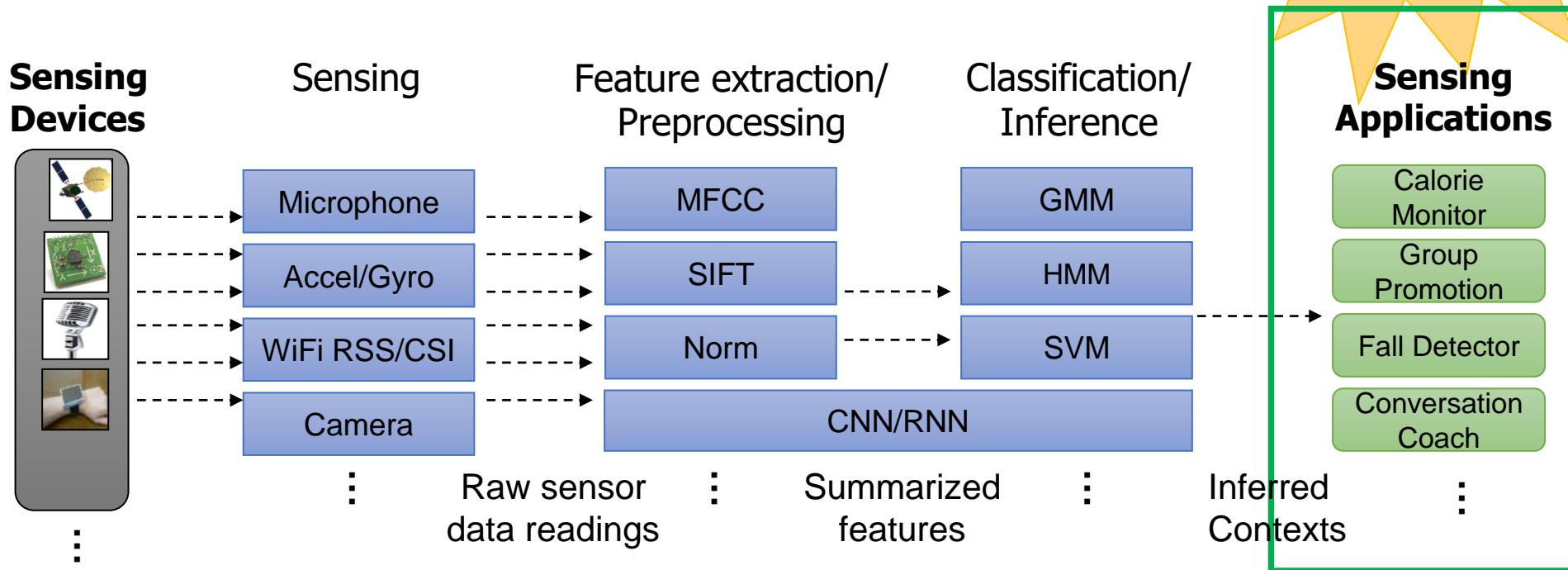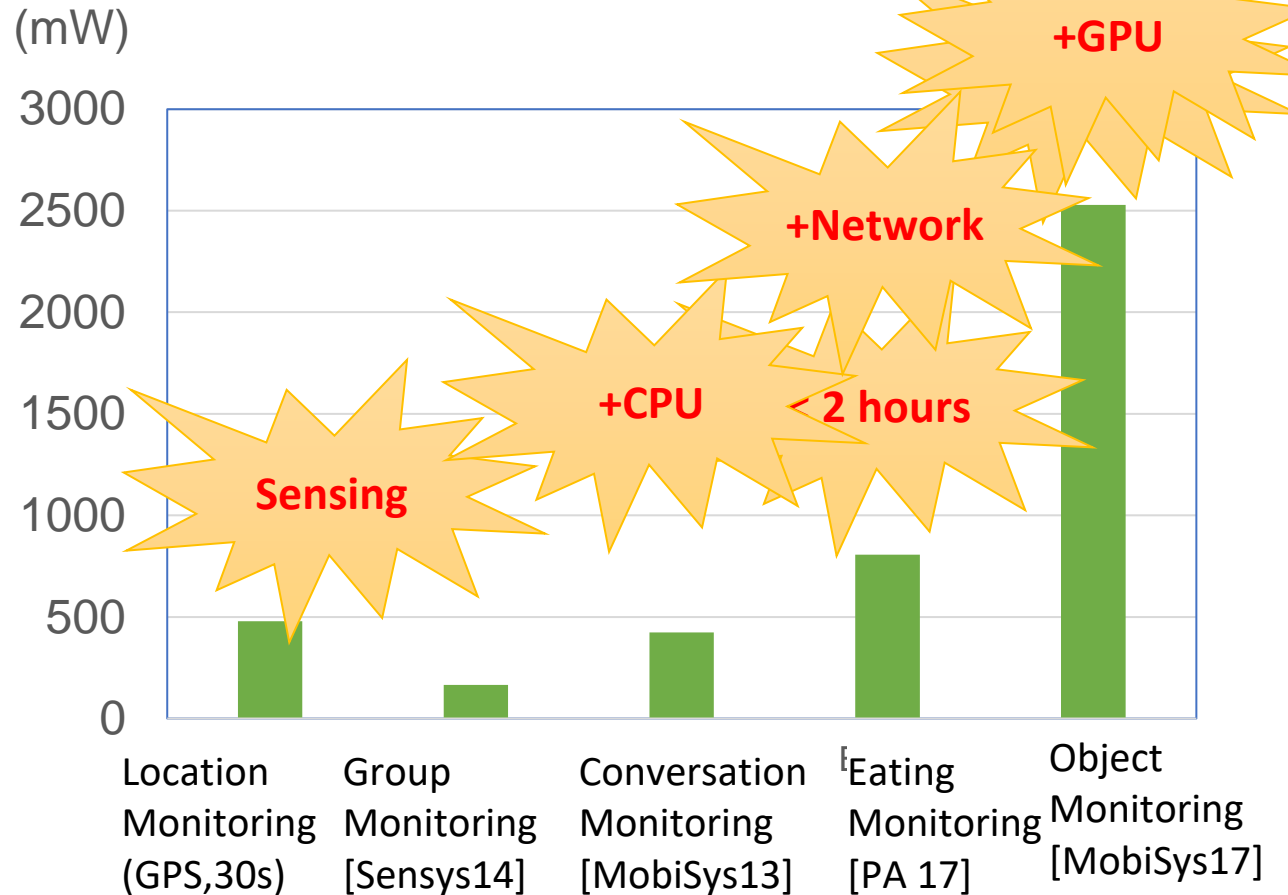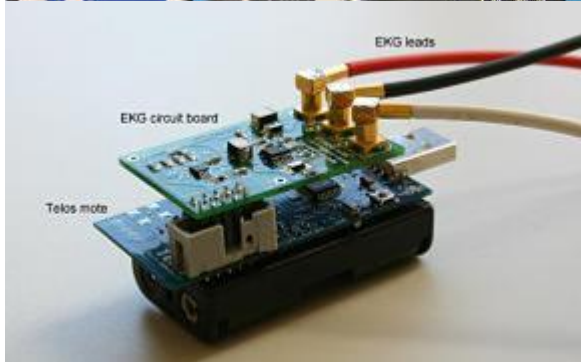Summarized features

Inferred Contexts

Challenge 2: Application Usability

- The inference results are not 100% correct.

- App design should overcome the inaccuracy.

App Design with Inaccurate Results

Sensing Devices	Sensing	Feature extraction/ Preprocessing	Classification/ Inference	**Sensing Applications**
	Microphone	MFCC	GMM	Calorie Monitor
	Accel/Gyro	SIFT	HMM	Group Promotion
	WiFi RSS/CSI	Norm	SVM	Fall Detector
	Camera	CNN/RNN		Conversation Coach
		Raw sensor data readings	Summarized features	Inferred Contexts

Challenge 3: Power Scarcity

Chart axis: (mW), values 0, 500, 1000, 1500, 2000, 2500, 3000

Categories:
- Location Monitoring (GPS, 30s)
- Group Monitoring [Sensys14]
- Conversation Monitoring [MobiSys13]
- Eating Monitoring [PA 17]
- Object Monitoring [MobiSys17]

Annotations: **Sensing**, **+CPU**, **+Network**, **+GPU**, **< 2 hours**

- Measured with Samsung Note 4 (3220mAh battery)
- Used Samsung Gear (315 mAh battery) for Anapruna (eating detection)

Challenge 4: New Operational Mode

Vs.
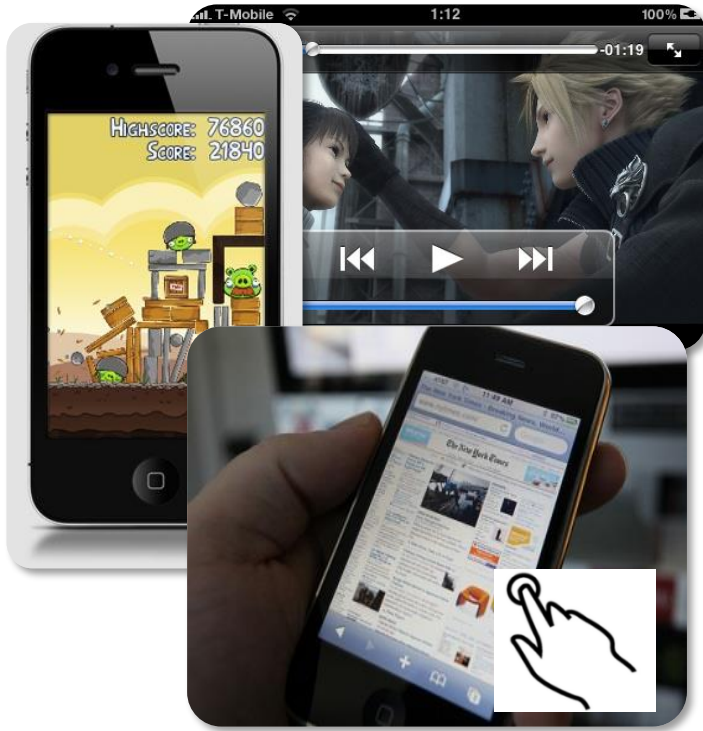
Small display, user mobility
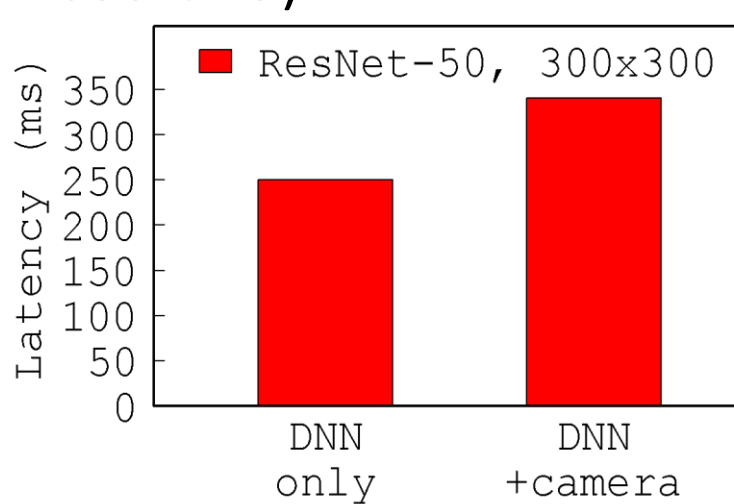☐ **A single user-interactive application**

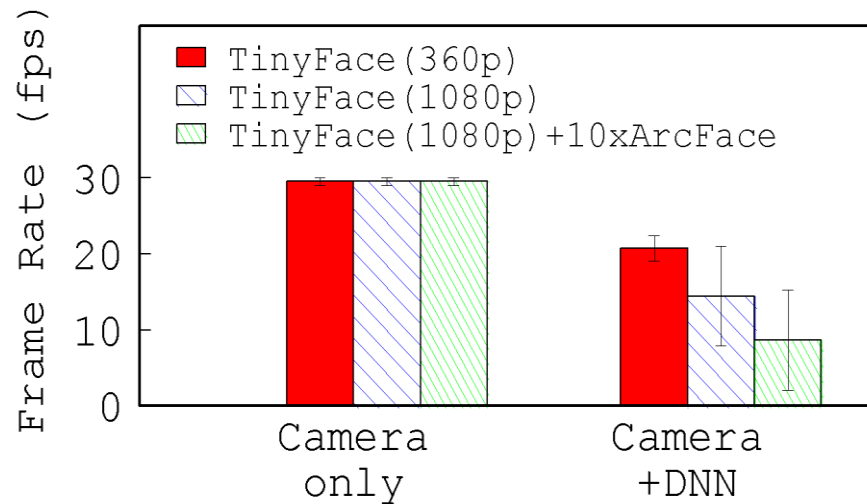**Mobile sensing** : *autonomous, situation-aware* services
☐ **Concurrent background sensing applications**

Challenge 5: Resource Contention

- Ran face detector continuously (with a TinyFace CNN-based Model) and a foreground camera rendering concurrently.

- Measured frame rates on Google Pixel 3 XL (Qualcomm Adreno 630 GPU).

DNN Execution and Rendering Performance

Challenge 6: Poor Scalability

Amazing mobile s
How to test with

Lets test it with lab users and a small number of real users and consider it "real-world".

Wow! It does not work!
Need access to real venues
With real users on real devices
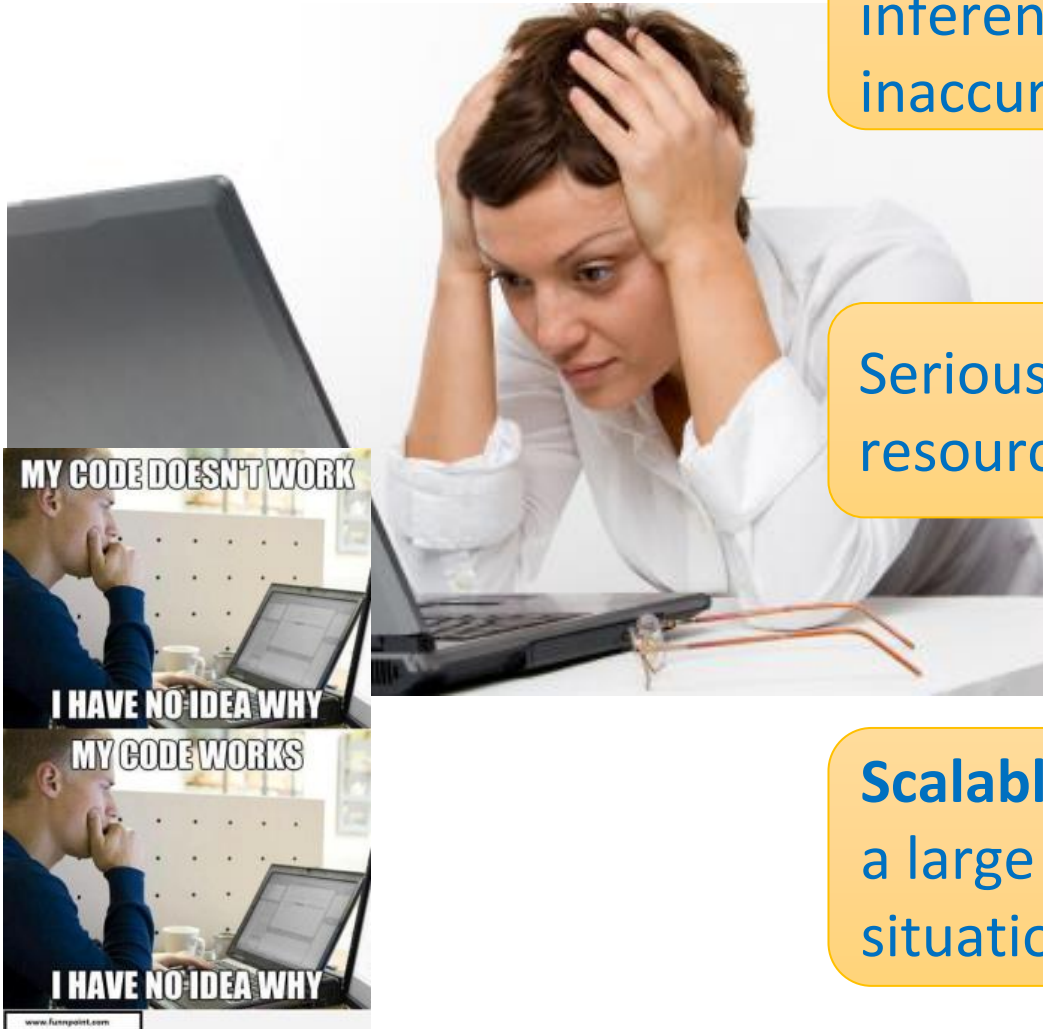HOW???

Individual Apps Solve All These?

Complexity to implement **accurate** inference logics and handling inaccuracy in app design

+

Serious **optimization** in battery and resource usage

+

Scalable deployment and testing with a large pool of real users in real-life situations

Full-Fledged Mobile Sensing Platform

Life-Immersive Sensing Applications

Group-aware ~~~~ **Fall monitoring**

"Notify me when the user is in a group of 3"

Simple and Intuitive Context Specification

Context Sensing and Analytics Platform
(on mobile/IoT/wearable devices and clouds)

Abstraction of Inference Logic and Runtime Resources

A rich set of mobile/IoT/wearable devices

Course Objectives

- Upon completion of the course, you should be able to:

    ✔ Understand key concepts and technical underpinnings of various machine learning techniques.

    ✔ Apply machine learning models to various real-world problems.

Class Timings

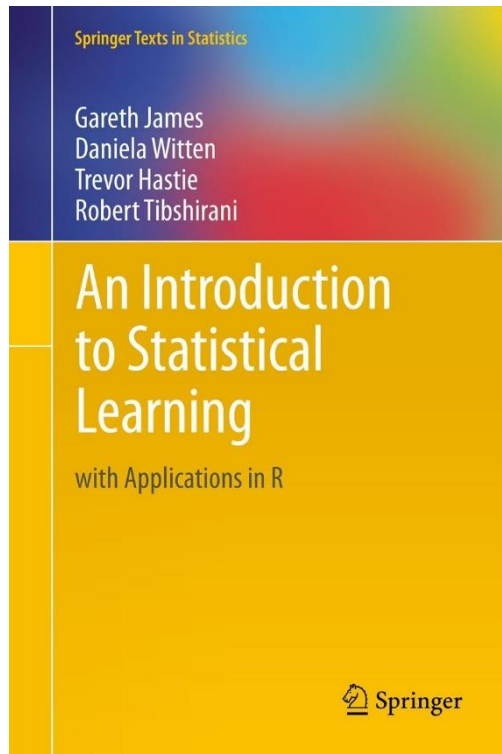| 3주차	1/17	1/18	1/19	1/20	1/21
오전	기계학습 강의 (이영기): 4시간				
오후	기계학습 실습 (조교): 4시간				
4주차	1/24	1/25	1/26	1/27	1/28
오전	기계학습 강의 (이영기): 4시간		강의 / 기말고사		
오후	기계학습 실습 (조교): 4시간				

Pre-Requisites

- Took an undergraduate "introduction to statistics" and "introduction to linear algebra" courses.
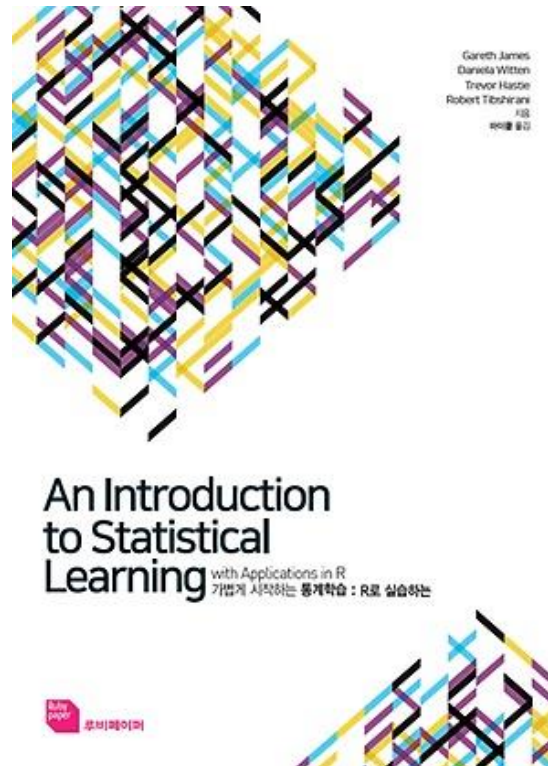
- Familiar with python.

Textbook

- Week 1

  ✔ 가볍게 시작하는 통계학습

- Week 2

  ✔ Lecture Slides

Philosophy of the Textbook

- It is important to understand the ideas behind the various techniques, in order to know how and when to use them.

- One has to understand the simpler methods first, in order to grasp the more sophisticated ones.

- It is important to accurately assess the performance of a method, to know how well or how badly it is working [simpler methods often perform as well as fancier ones!]

- This is an exciting research area, having important applications in science, industry and finance.

- Statistical learning is a fundamental ingredient in the training of a modern data scientist.

Labs

- GitLab

 ✔ https://gitlab.com/machine-learning-course1/ml-practice-lg-2022-spring

- Version

  ✔ Python >= 3.5.2

  ✔ requirements.txt

Tentative Lesson Plan

| Day	Lecture Topic	Evaluation Milestones
1	Class Intro & Intro to Machine Learning	
2	Linear Regression and Classification I	
3	Classification II and Resampling Methods	
4	Variable Selection	
5	Non-Linearity and Tree-based Methods I	
6	Tree-based Methods II and Support Vector Machines	Mini Project Due (40%)
7	Unsupervised Learning and Hidden Markov Models	
8	Convolution Neural Network Introduction	Final Exam (60%)

Teaching Assistant

- Jingyu Lee (이진규)

- Changmin Jeon (전창민)

- Hyunwoo Jeong (정현우)

- Juheon Yi (이주헌)

- How to contact TAs?

  ✔ Email to ta_ml@hcs.snu.ac.kr.
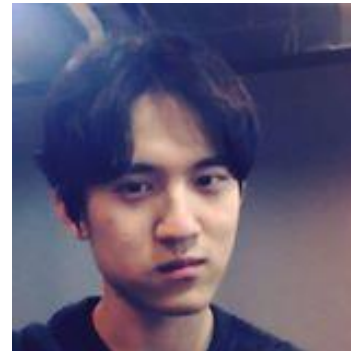
Assessment
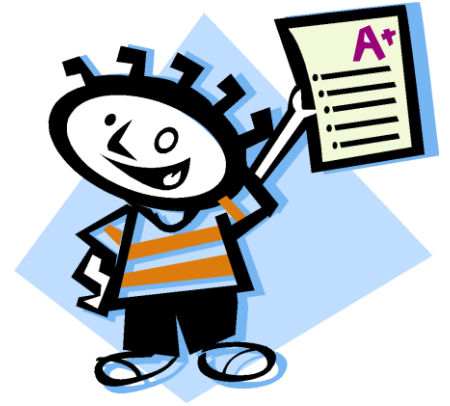
- Mini-Project (40%)

- Final Exam (60%)

  ✔ Week 2 Wednesday, Closed Book, All Materials

- Useful link

  ✔ Answers to textbook exercises:

  http://blog.princehonest.com/stat-learning/

Quick Intro of Everybody

- Brief Intro.

- What ML problems you are interested / solving.

Most Importantly …

- Let's work hard but have fun!