Machine Learning

LG전자 고급 데이터 사이언티스트 양성 과정

이영기

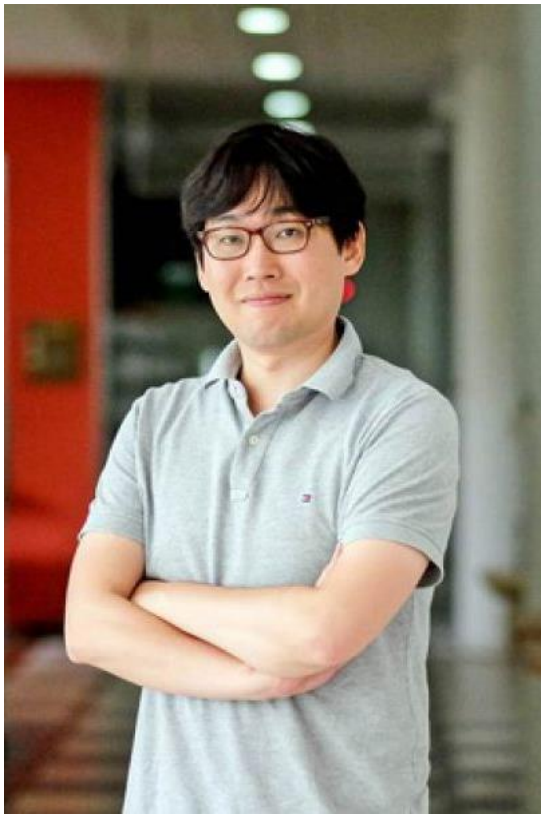
서울대학교 컴퓨터공학부

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



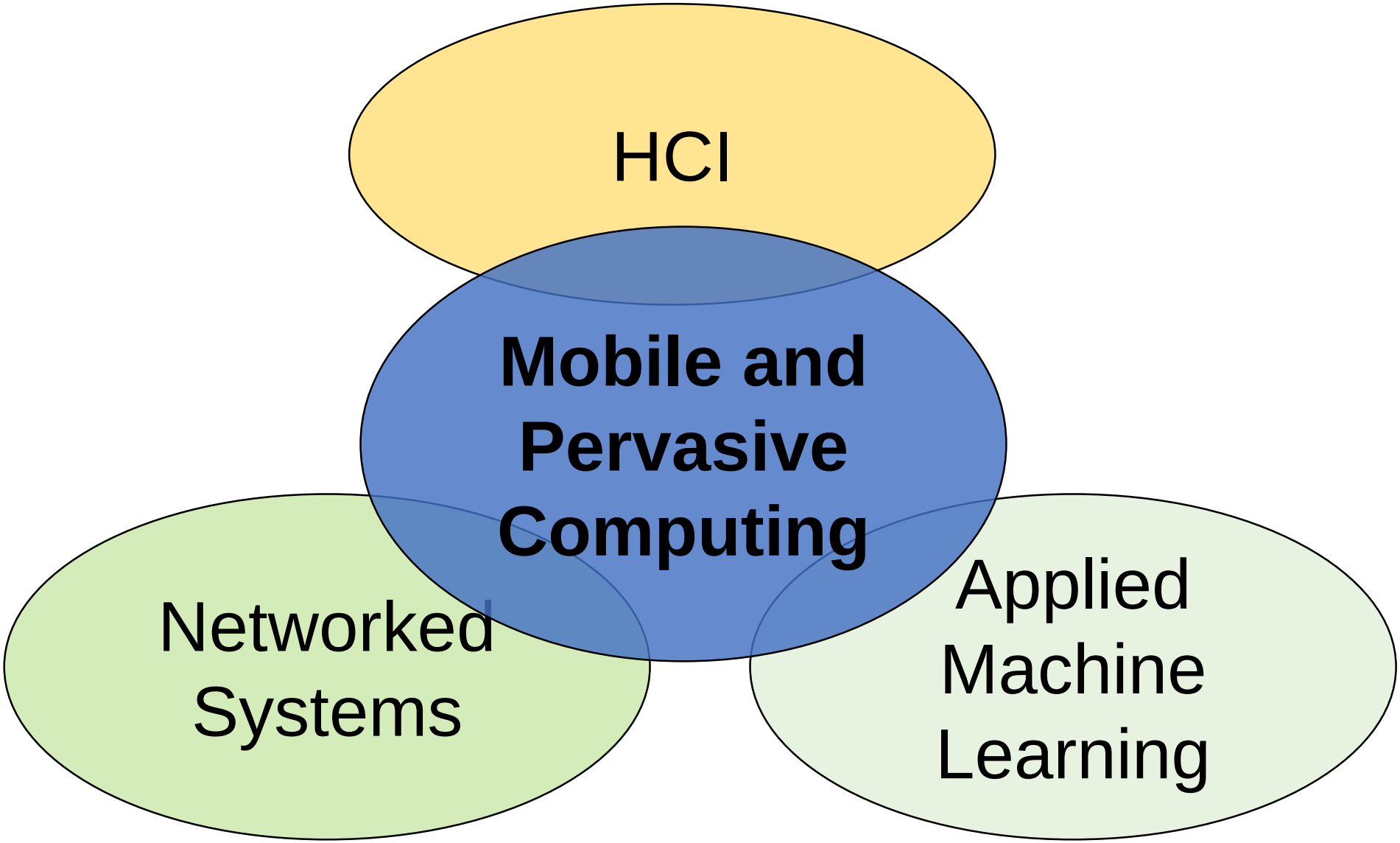
Quick Introduction

- **Education**
  - ✓ **Ph.D.**, Computer Science, KAIST, Korea, 2012
  - ✓ **B.S**., Computer Science, KAIST, Korea, 2003
- **Experience**
  - ✓ Assistant Professor, Computer Science and Engineering, Seoul National University, Korea, September 2018 - Present
  - ✓ Assistant Professor, School of Information Systems, Singapore Management University, Singapore, 2013 - 2018
- **Research Interest**
  - ✓ Human Behavior and Context Sensing
  - ✓ Mobile and Embedded Machine Learning
  - ✓ Systems and Tools for Emerging Devices
  - ✓ Large-Scale Data Analytics Platform

Agenda for Today

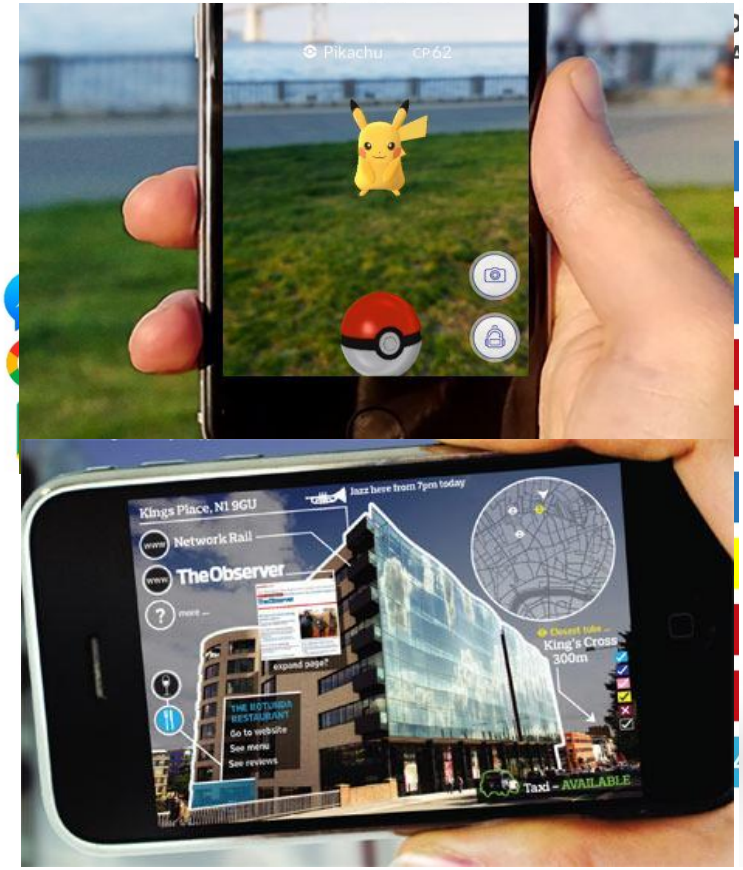
- Introduction to the class
- Introduction to the machine learning
- Introduction to the statistical learning
  - ✓Chapter 2 of our textbook
- Lab for the chapter 2

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).

- 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

- 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

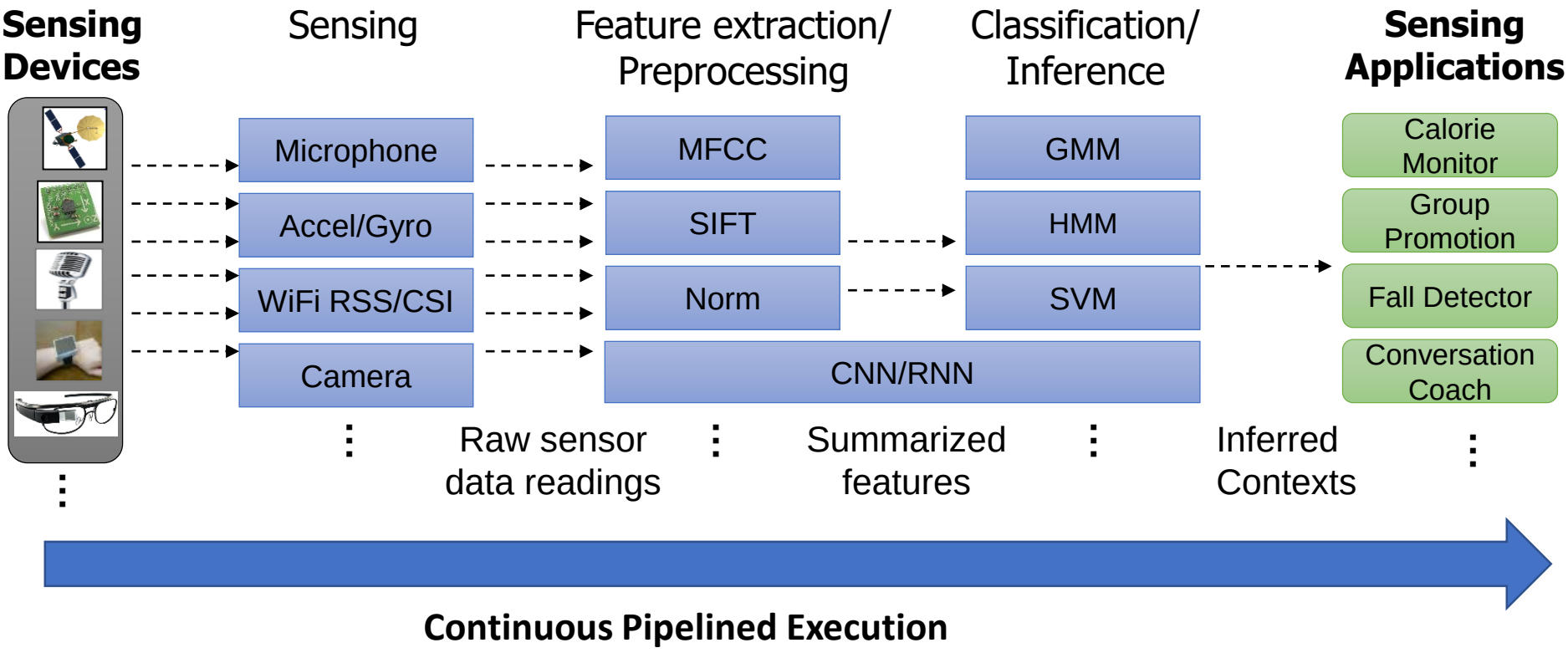
Conversation

Internal States

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

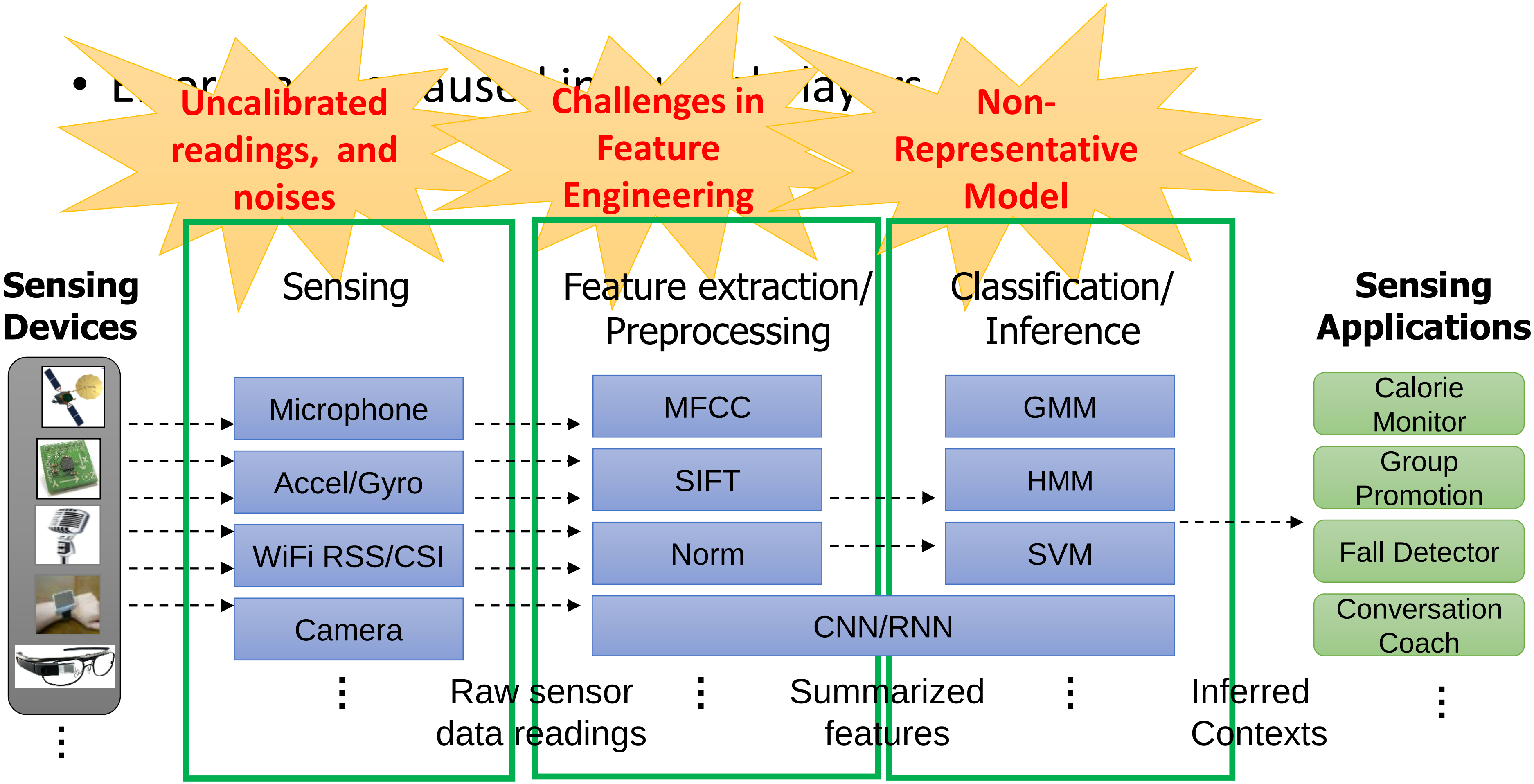
Common Computational Flow

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



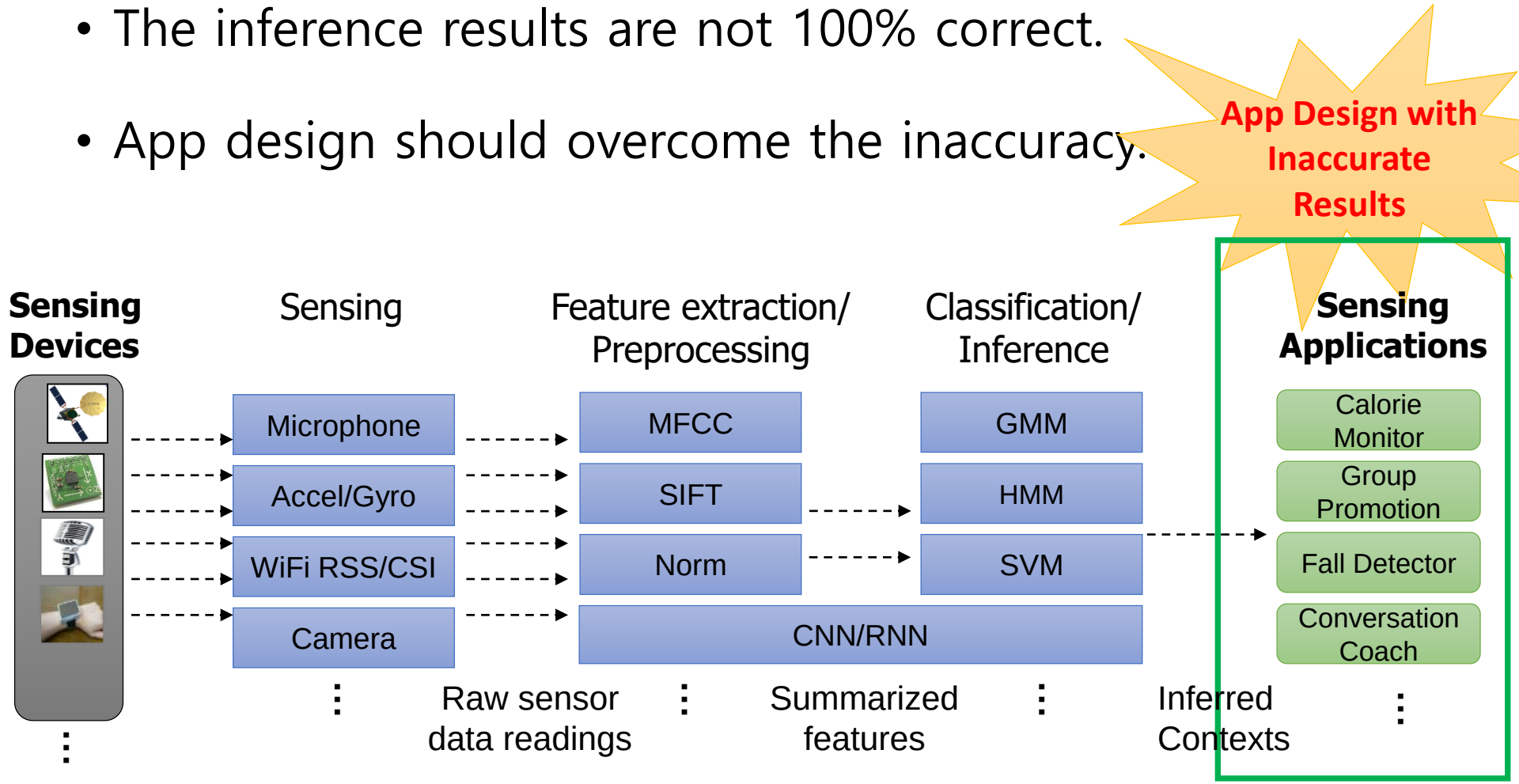
Challenge 1: Inference Accuracy

- > 90% accuracy is extremely challenging.

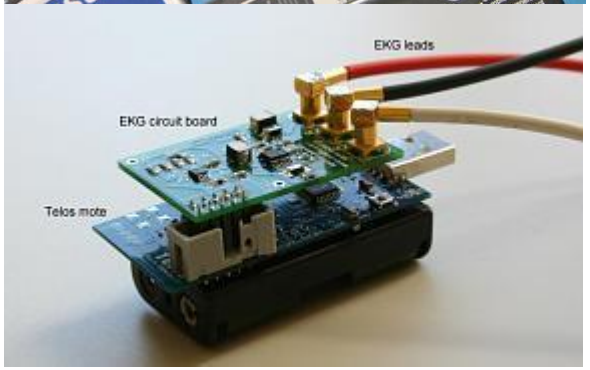


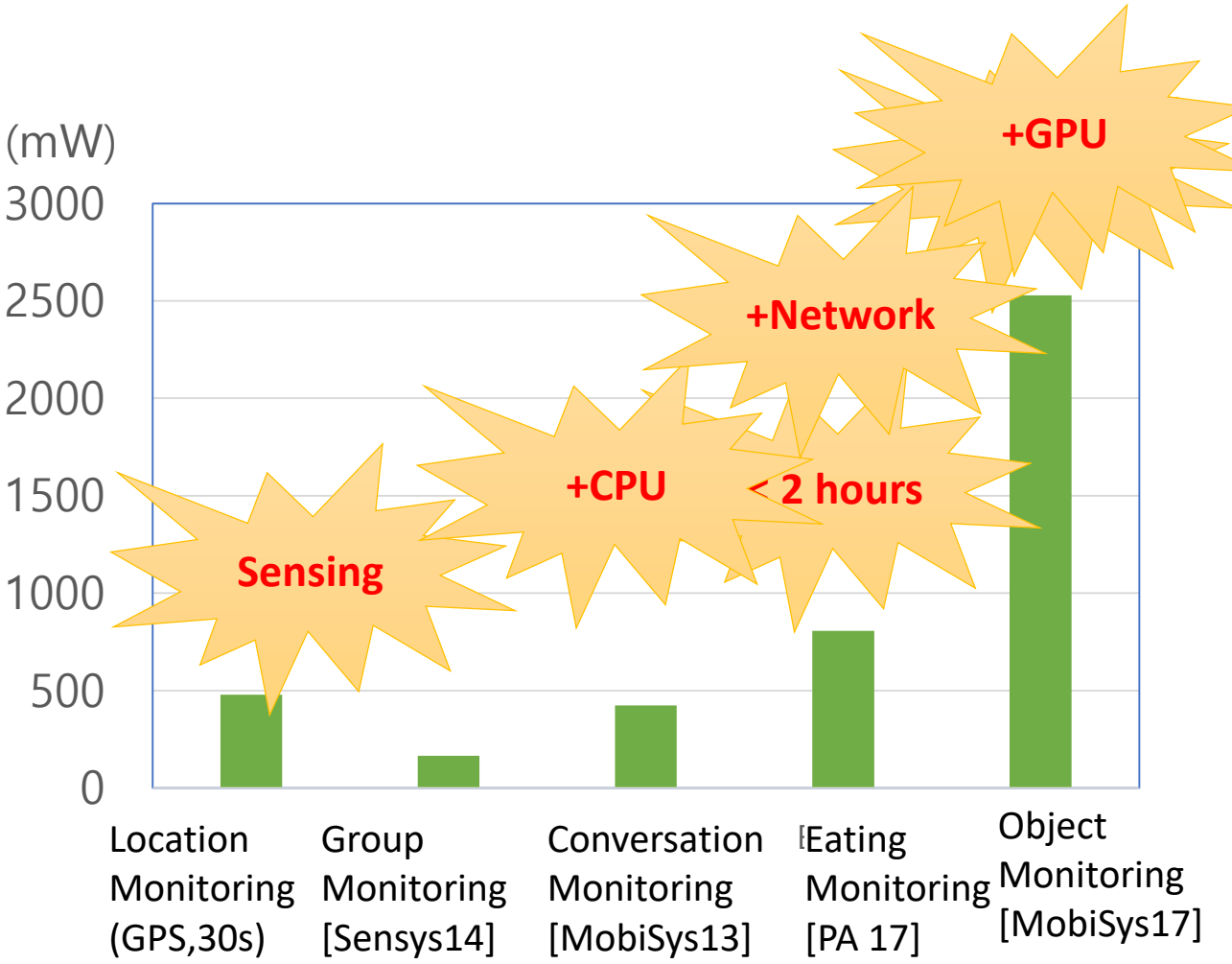
Challenge 2: Application Usability

- The inference results are not 100% correct.
- App design should overcome the inaccuracy.



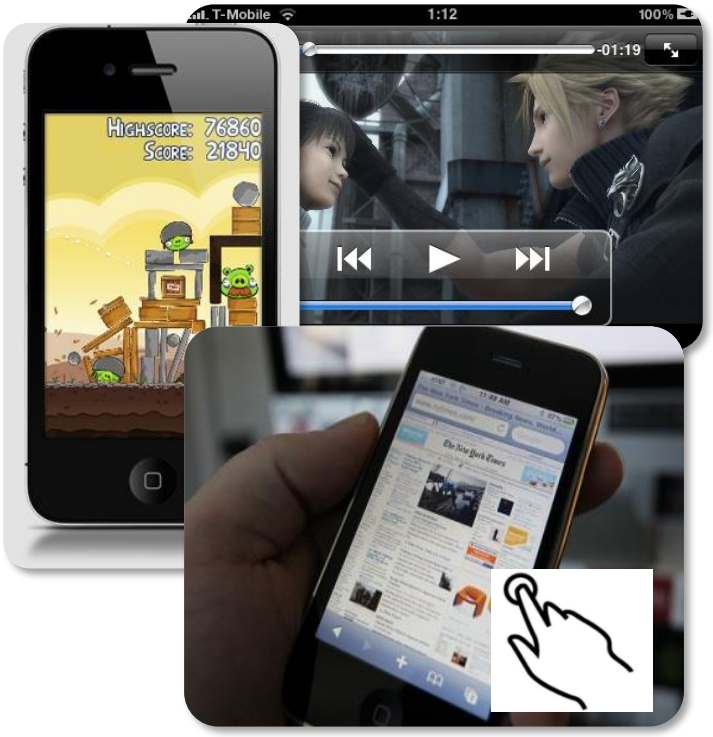
Challenge 3: Power Scarcity





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

Challenge 4: New Operational Mode



Small display, user mobility

→ **A single user-interactive application**

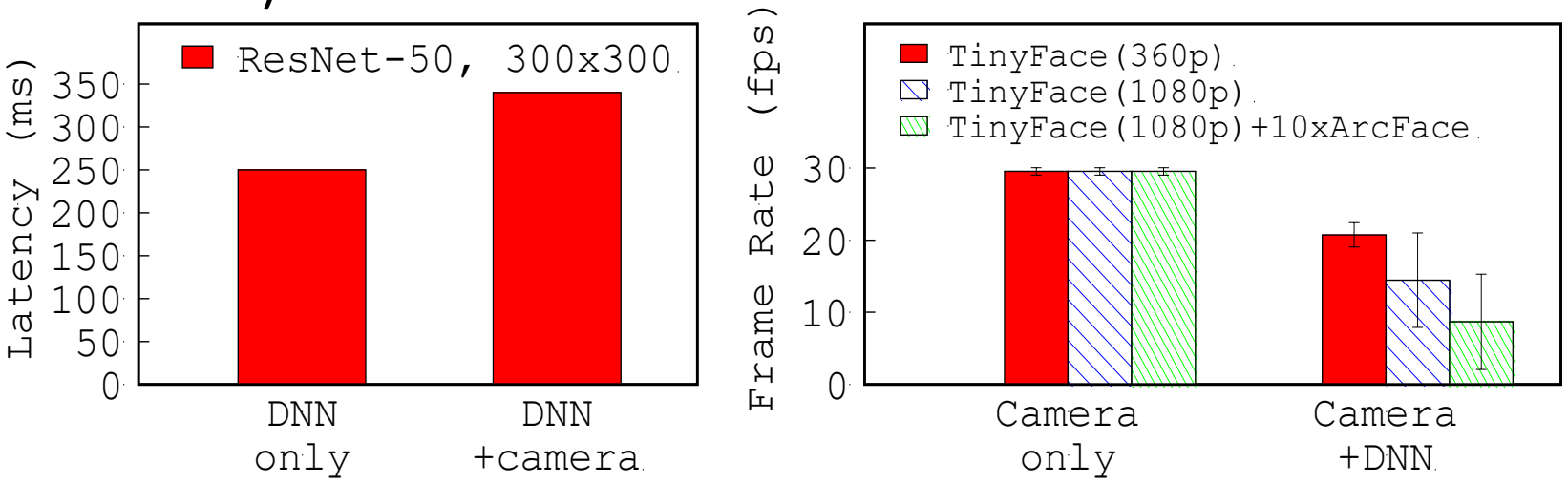
Vs.

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 Time

Frame Rate of Camera Rendering

DNN Execution and Rendering Performance

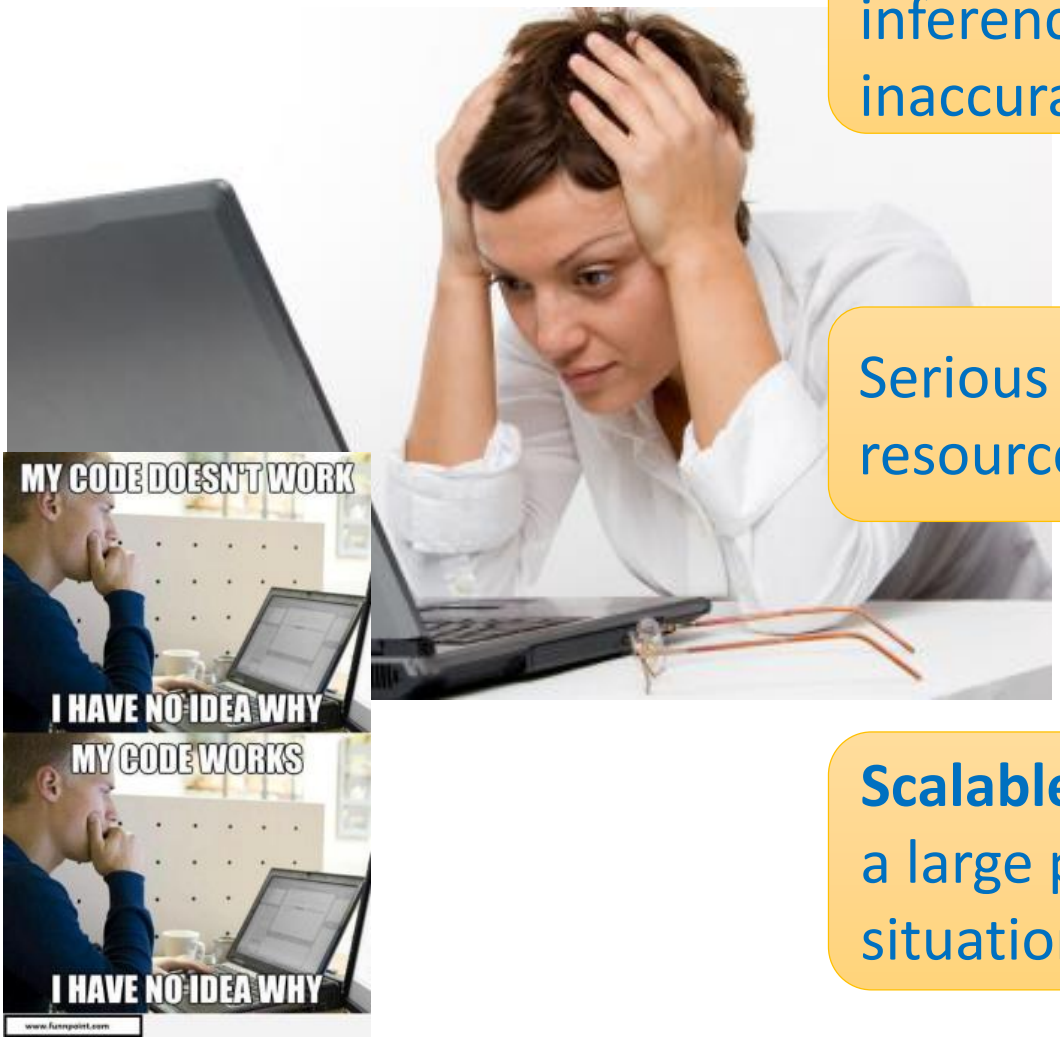
Challenge 6: Poor Scalability

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

+

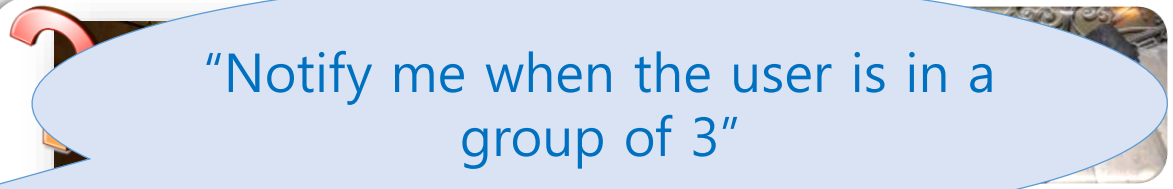
Serious **optimization** in battery and resource usage

+

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

Full-Fledged Mobile Sensing Platform

Life-Immersive Sensing Applications



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

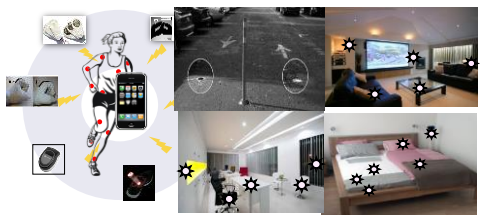
Group-aware Ads **Fall monitoring**

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

4주차	2/3	2/4	2/5	2/6	2/7
오전	데이터마이닝 (심규석)	데이터마이닝 (심규석)	데이터마이닝 (심규석)	데이터마이닝 (심규석)	데이터마이닝 (심규석)
오후	기계학습 (이영기)	기계학습 (이영기)	기계학습 (이영기)	기계학습 (이영기)	기계학습 (이영기)
5주차	2/10	2/11	2/12	2/13	2/14
오전	기계학습 (이영기)	기계학습 (이영기)	기계학습 (이영기)	심층학습 (윤성로)	기계학습 (이영기)
오후	심층학습 (윤성로)	심층학습 (윤성로)	심층학습 (윤성로)	심층학습 (윤성로)	기계학습 (이영기)

Pre-Requisites

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

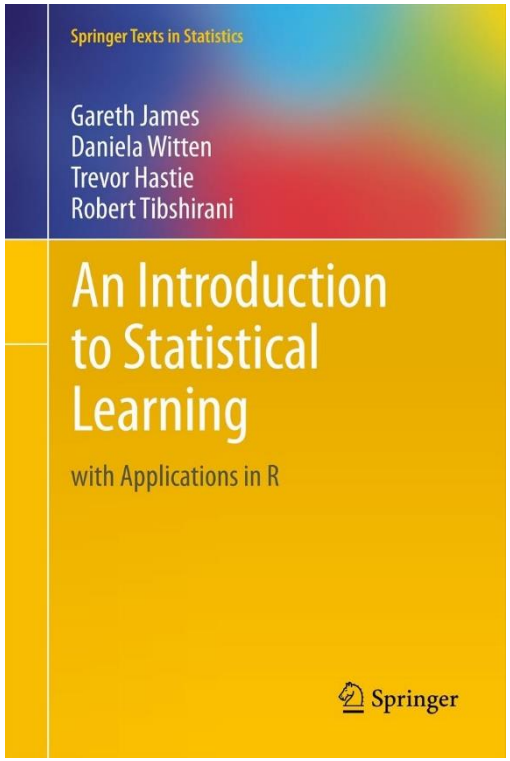
Textbook

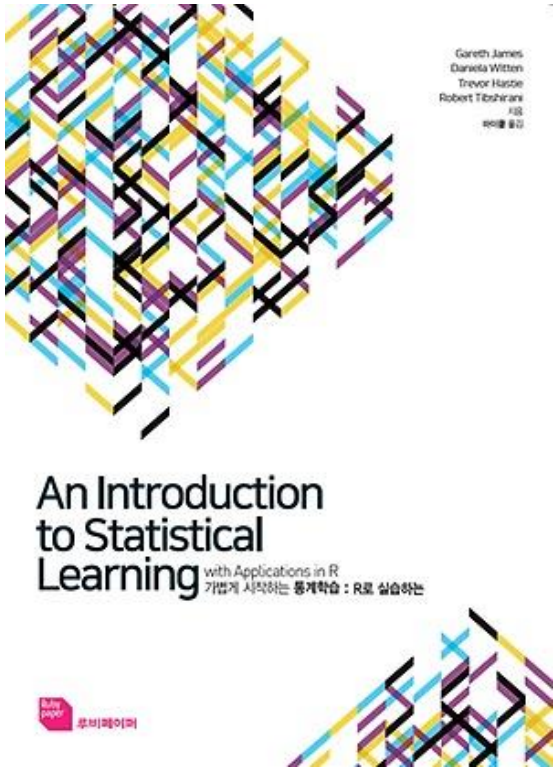
- Week 1

  ✓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

- https://gitlab.com/chphch/ISLR-python
- Version
  - ✓Python >= 3.5.2
  - ✓requirements.txt

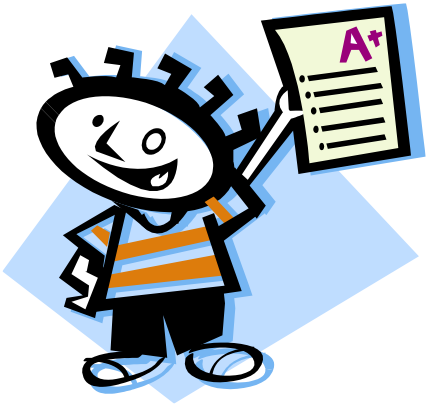
Tentative Lesson Plan

Day	Lecture Topic	Evaluation Milestones
1	Class Intro & Intro to Machine Learning	
2	Linear Regression and Classification	
3	Cross Validation and Variable Selection	
4	Non Linearity and Tree-based Models	
5	Support Vector Machines and Unsupervised Learning	Midterm Quiz (30%)
6	Hidden Markov Models	
7	Introduction to Convolutional Neural Networks	Mini Project Due (30%)
8	Application 1: Face Detection and Recognition	
9	Application 2: Human Activity Recognition	
10	Application 3: Healthcare Applications	Final Exam (40%)

Teaching Assistant

- Jingyu Lee (이진규)
- Changmin Jeon (전창민)
- Hyunwoo Jeong (정현우)
- How to contact TAs?
  - ✓Email to hcs.lab.2019@gmail.com.
  - ✓TAs will check the kakaotalk chatroom.

Assessment



- Project (30%)
- Midterm Exam (30%)
  - ✓Week 1 Friday, Closed Book, Week 1 Materials
- Final Exam (40%)
  - ✓Week 2 Friday, 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!