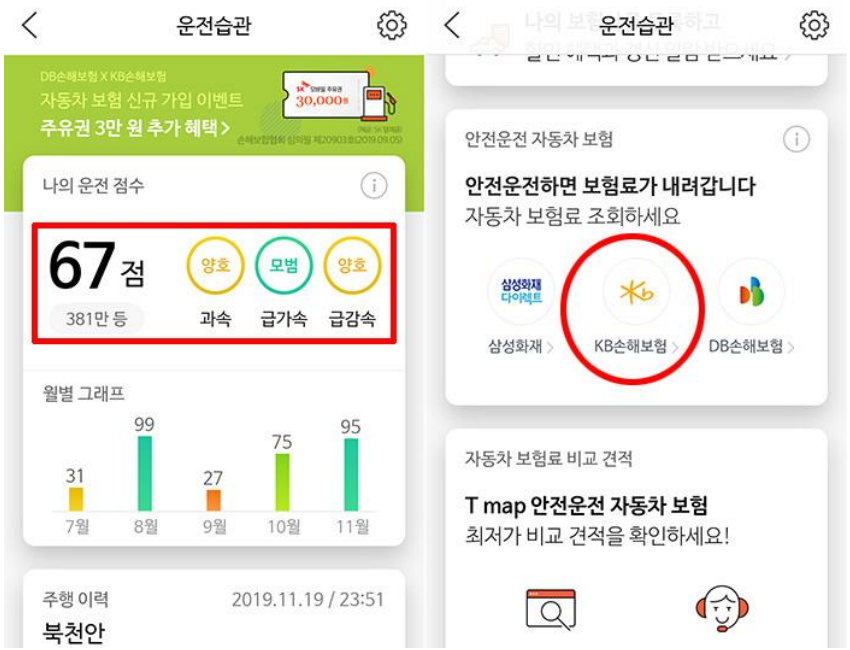
Service Intelligence Week 7.

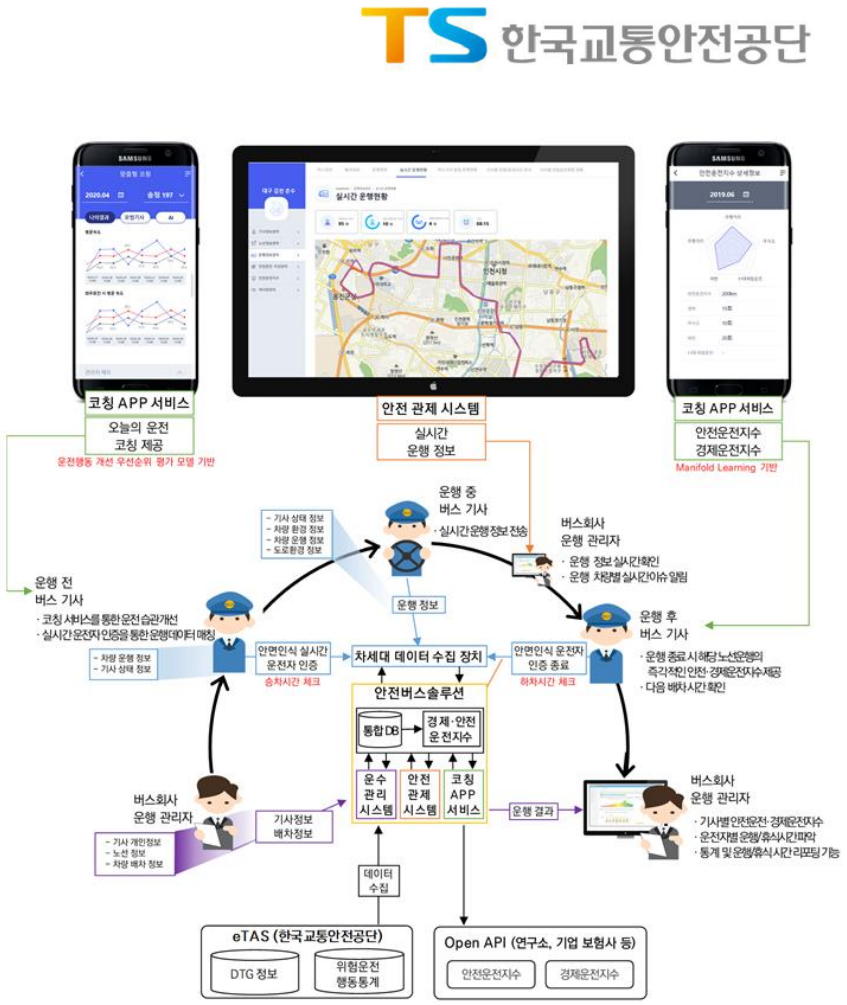
[Learning Behavioral Data of Customers for Personal Process Management]

Chiehyeon Lim

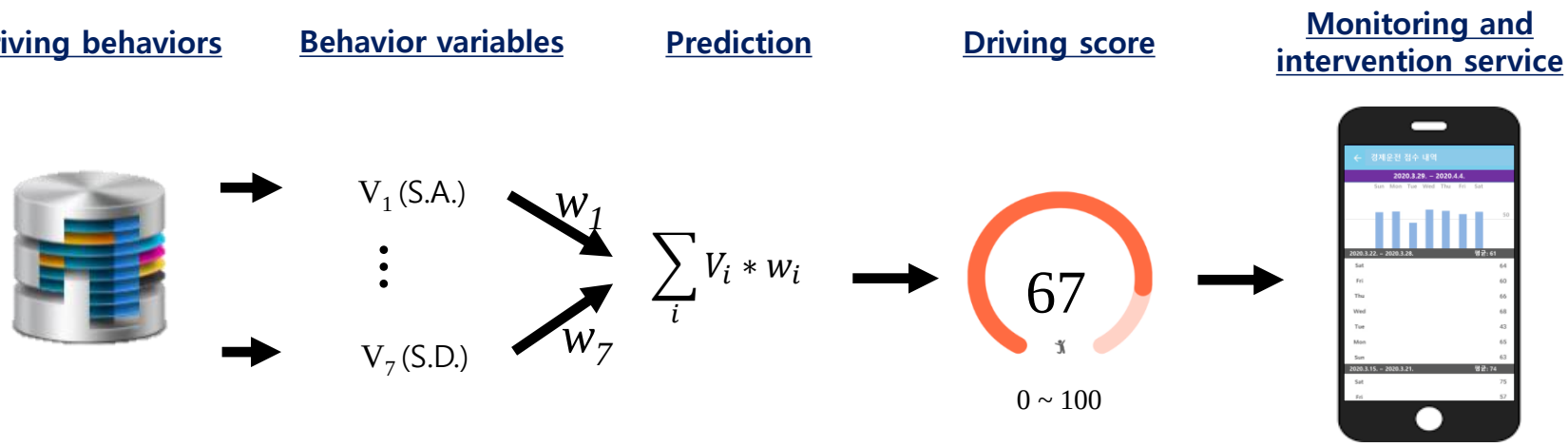
2022. 10. 12

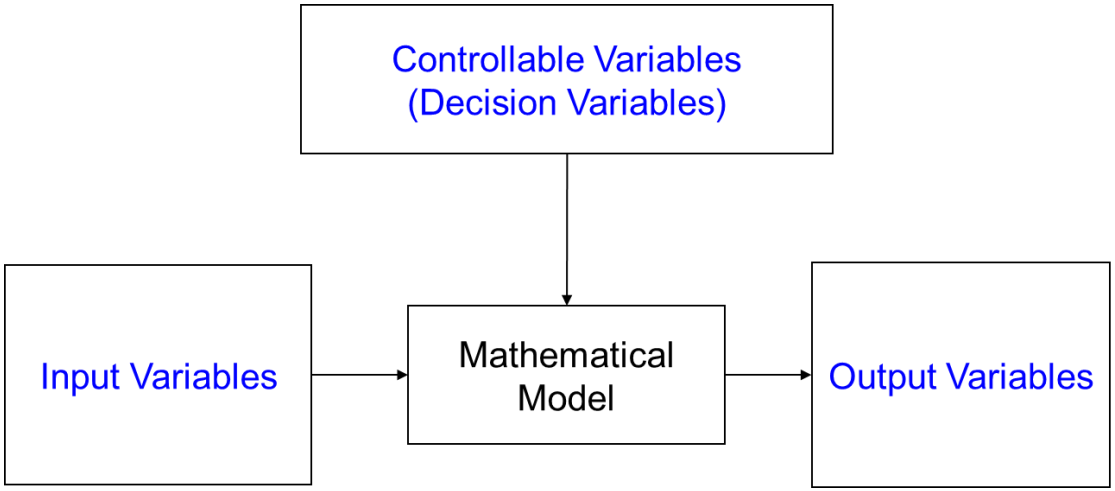
Driving Process Management: Safe Driving Context



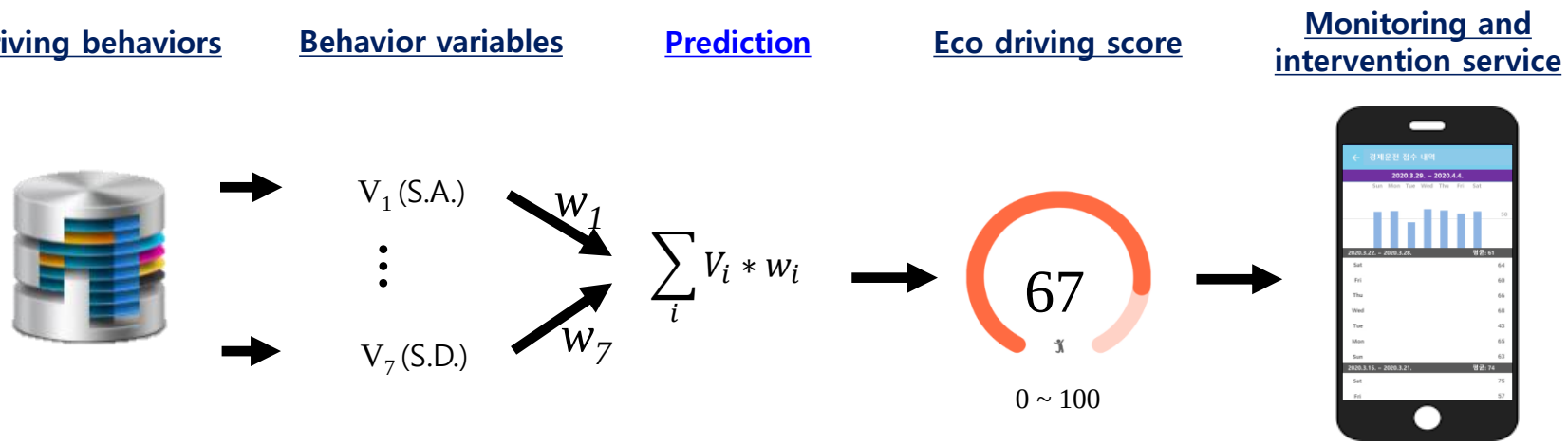


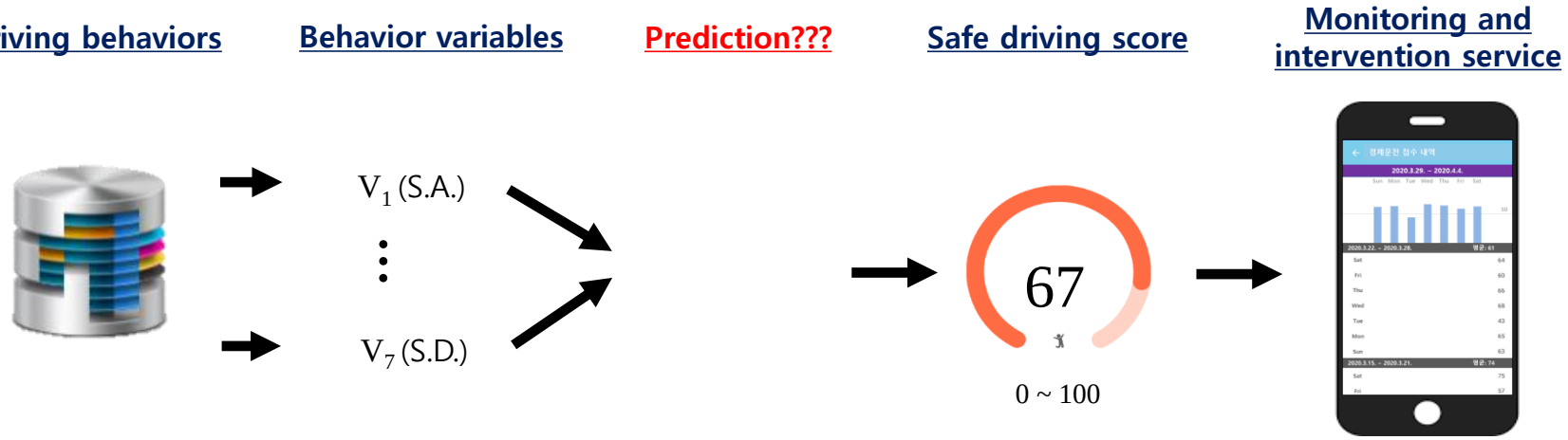
Driving Process Management: Eco Driving Context



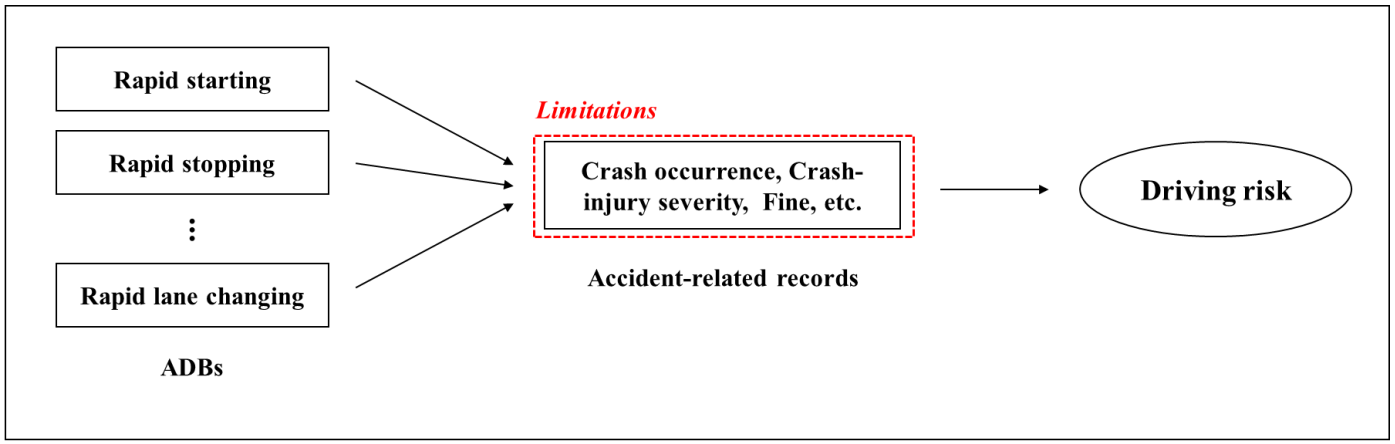


Existence vs. Nonexistence of the Labels

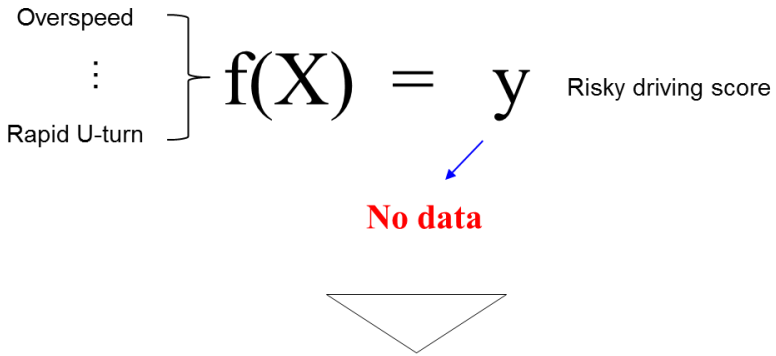




Safe Driving Context: Nonexistence of the Labels



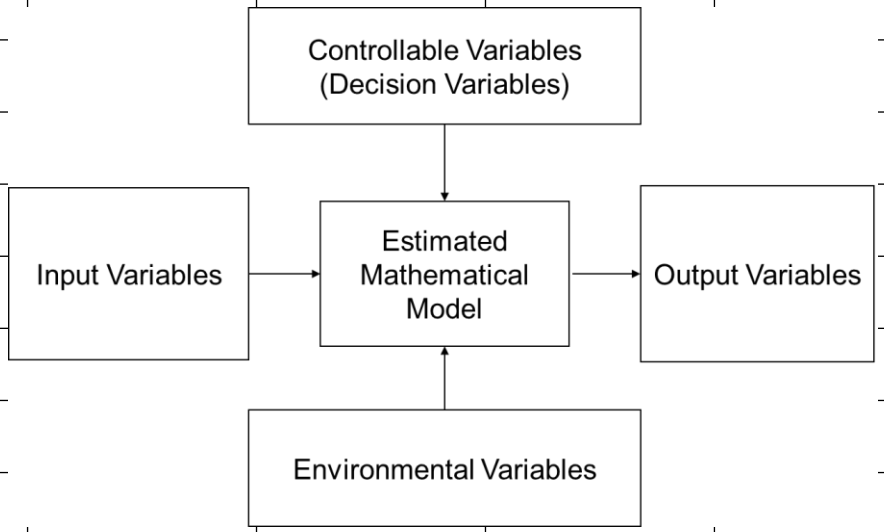
Limitation 1: Accident-related records are imperfect surrogate measures of driving risk
Limitation 2: Accident-related records are unreliable due to recording errors and reporting issues
Limitation 3: Accident-related records are difficult to acquire



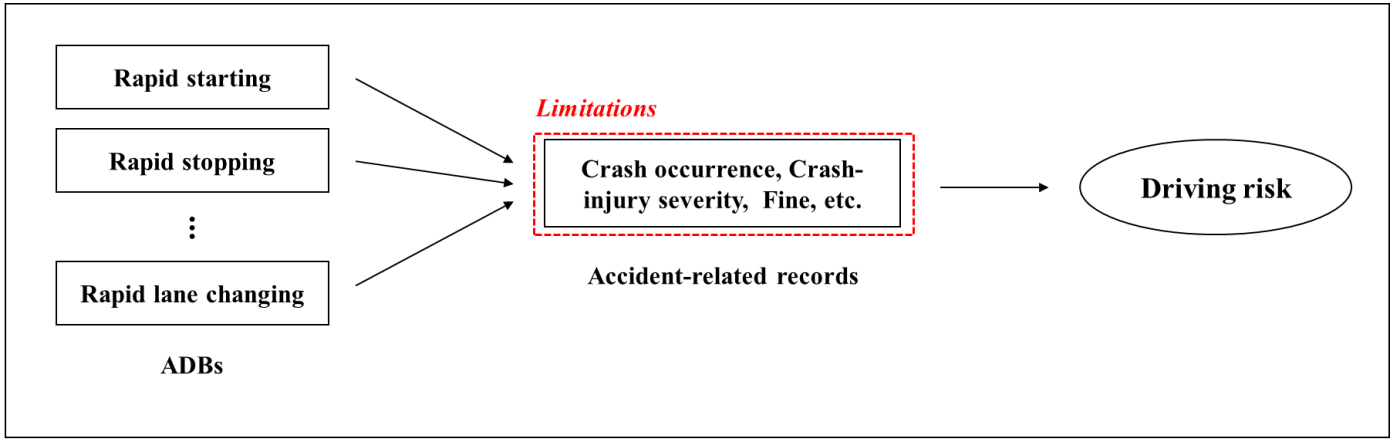
Estimating "f(X) = y" is impossible

A Common Case of PPM: Nonexistence of the Labels

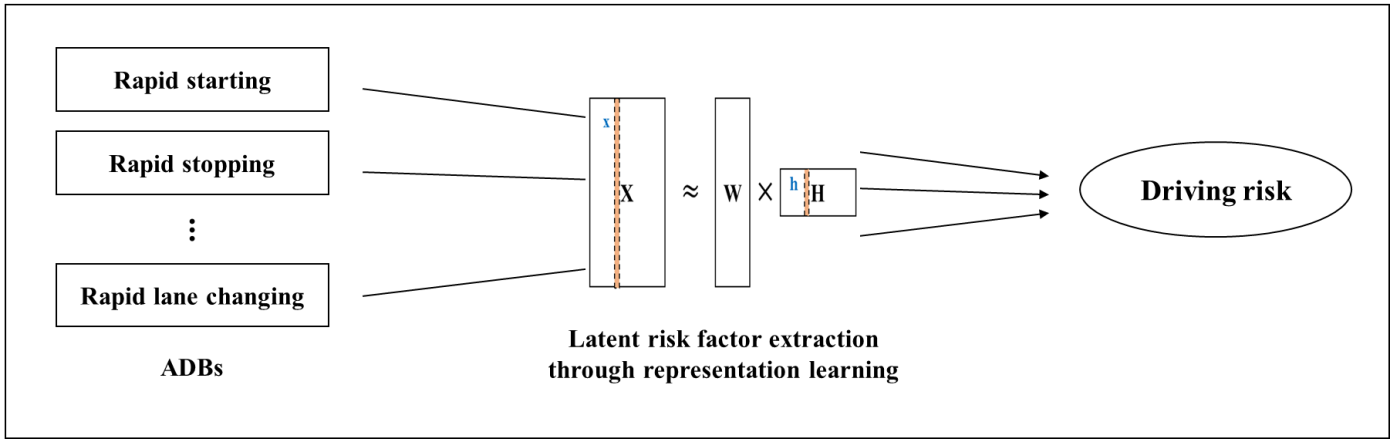
	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2	…					…
Customer 3	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…	…	…	…	…	…
Customer n-1	…	…	…	…	…	…
Customer n	…	…	…	…	…	…



Safe Driving Context: Unsupervised Way of Learning is Necessary

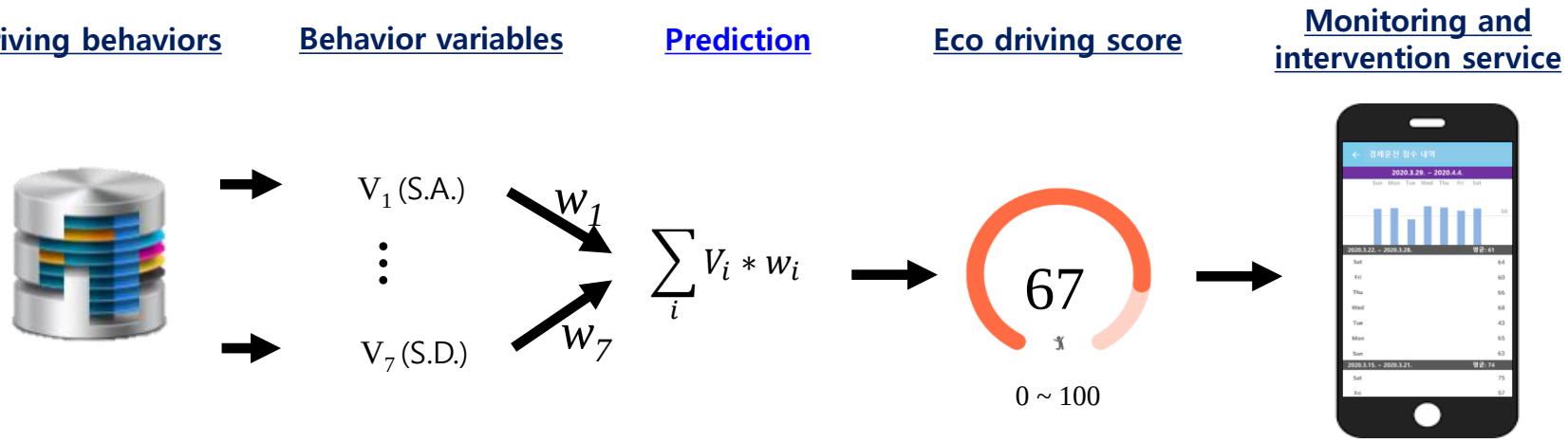


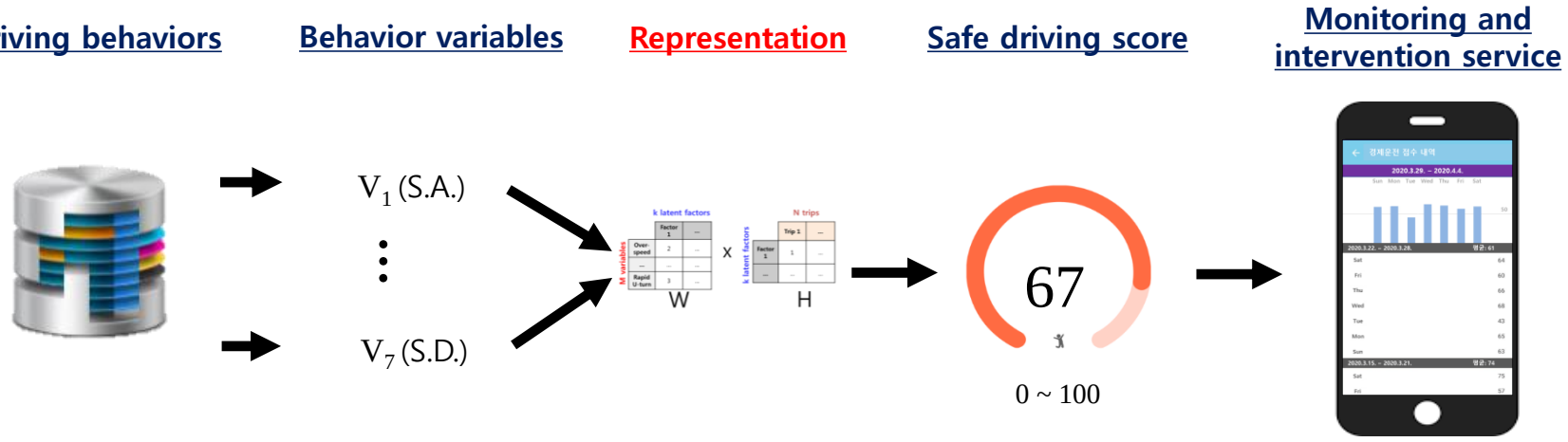
Limitation 1: Accident-related records are imperfect surrogate measures of driving risk
Limitation 2: Accident-related records are unreliable due to recording errors and reporting issues
Limitation 3: Accident-related records are difficult to acquire



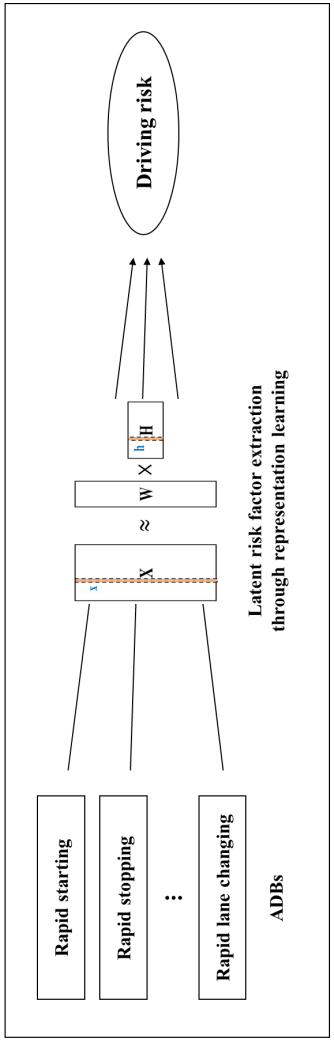
Reference: Seo, H., Shin, J., Kim, K., Lim, C. and Bae, J., "Driving Risk Assessment using Nonnegative Matrix Factorization with Driving Behavior Records," *IEEE Transactions on Intelligent Transportation Systems*, Online First, 2022.

Safe Driving Context: Unsupervised Way of Learning is Necessary



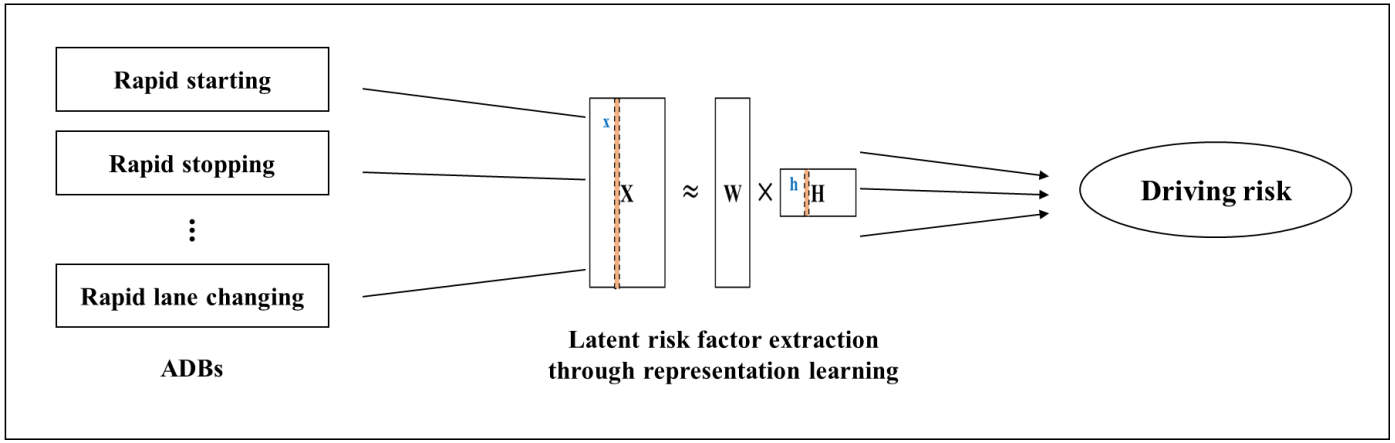


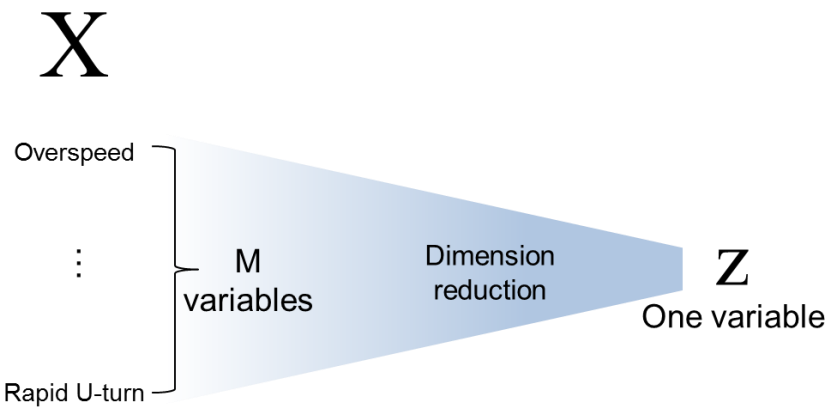
Driving Process Management: Safe Driving Context

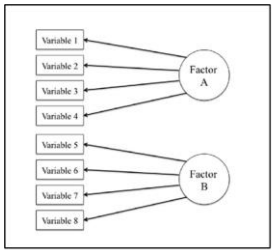


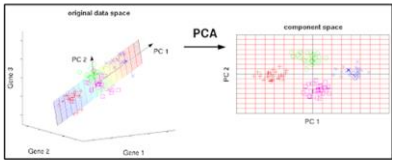
ADB	Bus criteria
Short-term over-speeding	The speed is 20 km/h above the speed limit
Long-term over-speeding	The speed is 20 km/h above the speed limit for 3 min
Rapid acceleration	The acceleration is greater than 6 km/h/s at a speed greater than 6 km/h
Rapid starting	The acceleration is greater than 8 km/h/s at a speed less than 5 km/h
Rapid deceleration	The deceleration is greater than 9 km/h/s at a speed greater than 6 km/h
Rapid stopping	The deceleration is greater than 9 km/h/s at a speed less than 5 km/h
Rapid lane changing	The change in rotation angle is 8°/s at a speed greater than 30 km/h, an acceleration (a deceleration) less than ±2 km/h/s, and an accumulated yaw rate change less than ±2°/s for 5 s
Rapid overtaking	The change in rotation angle is greater than 8°/s at a speed greater than 30 km/h, an acceleration greater than 3 km/h/s, and an accumulated yaw rate change less than ±2°/s for 5 s
Rapid turning	The accumulated change in rotation angle is from 60° to 120° (left or right) for 4 s at a speed greater than 20 km/h
Rapid U-turning	The accumulated change in rotation angle is from 160° to 180° (left or right) for 8 s at a speed greater than 15 km/h

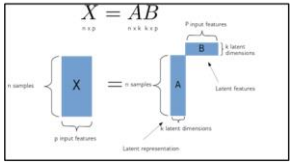
Driving Process Management: Safe Driving Context

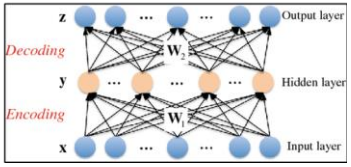




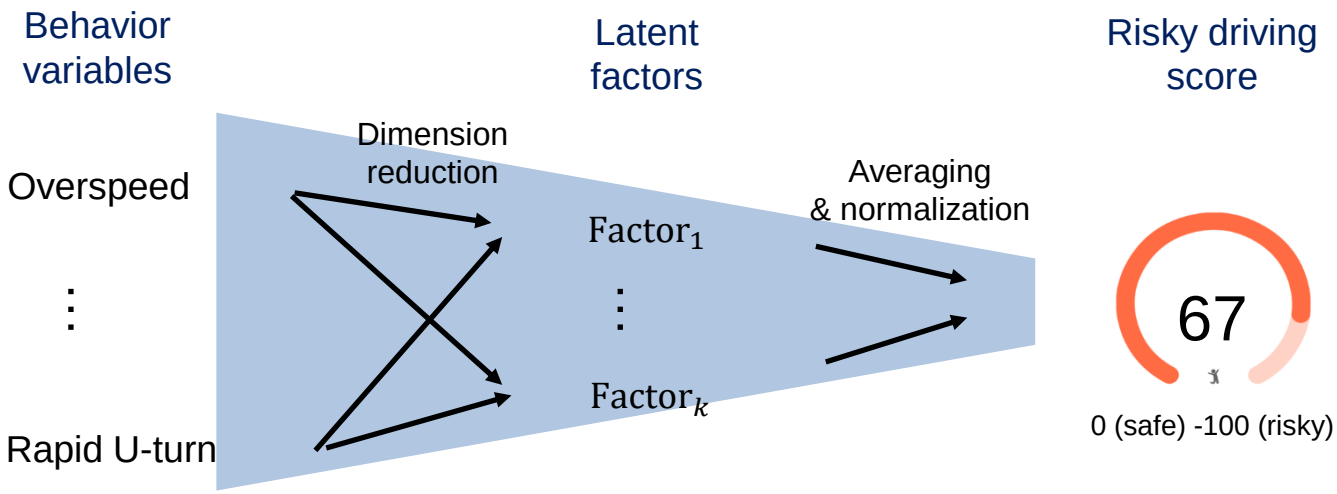


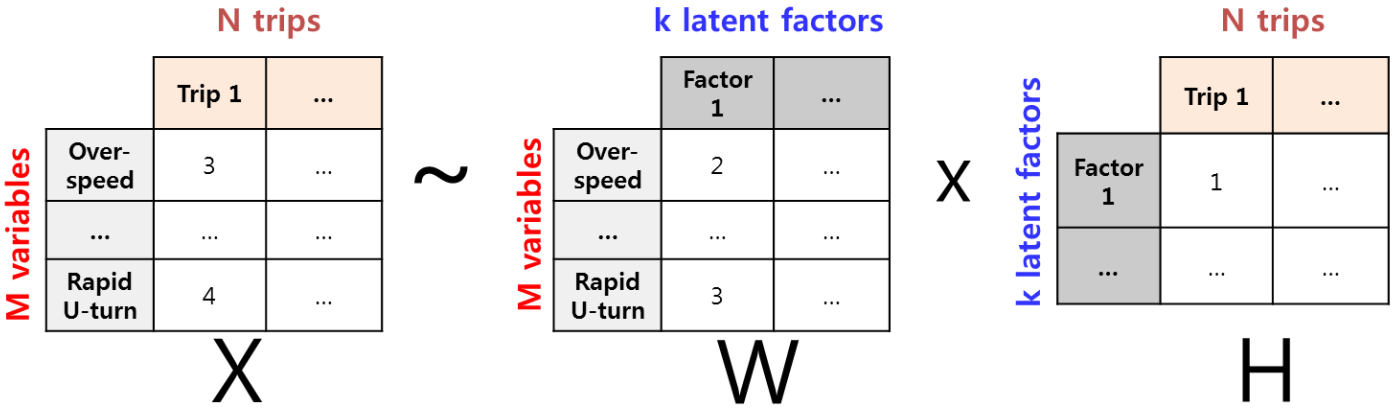




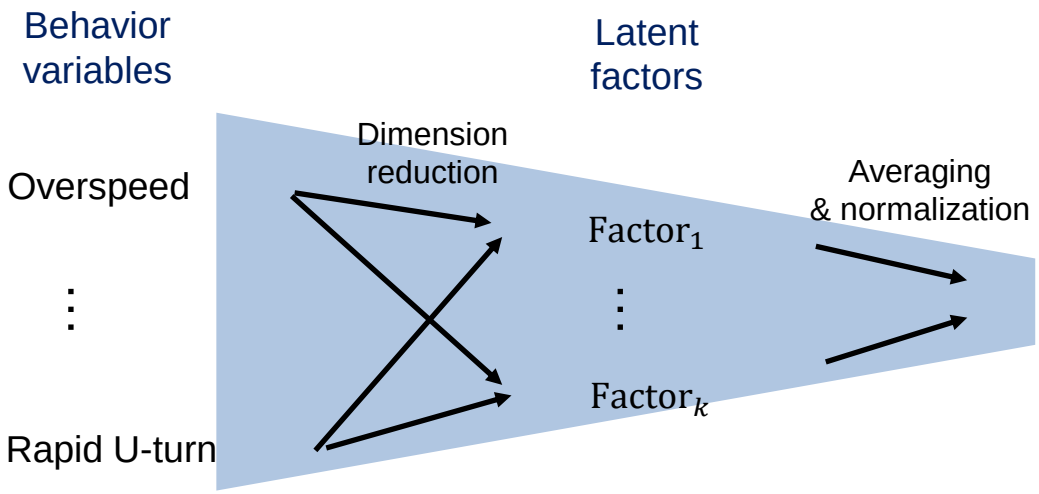


Driving Process Management: Safe Driving Context





Driving Process Management: Safe Driving Context



	Trip 1	...
Overspeed	3	...
Rapid stopping	5	...
Rapid U-turn	4	...

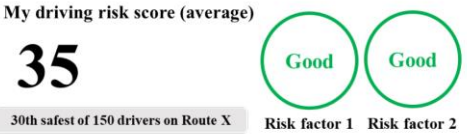
	Trip 1	...
Factor 1	1	...
Factor 2	2.5	...

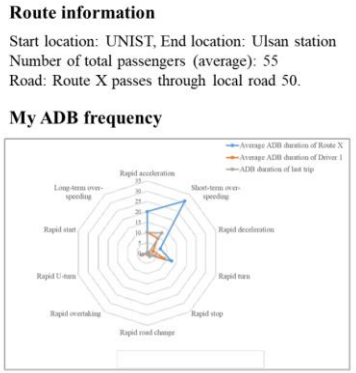
	Trip 1	...
Score	67	...

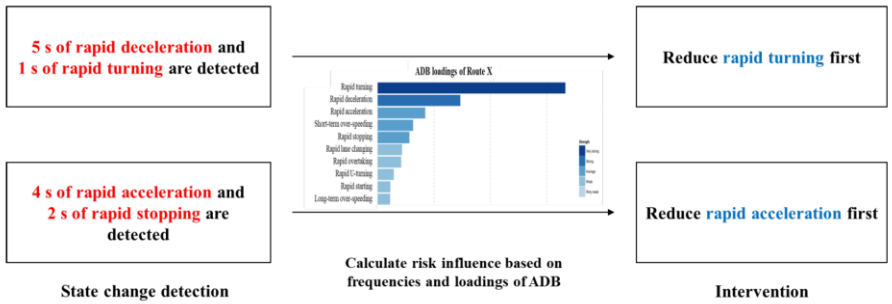
Driving Process Management: Safe Driving Context

LOADINGS OF 10 ADBS ON LATENT RISK FACTORS

Route number	Latent risk factors	Rapid acceleration	Short-term over-speeding	Rapid deceleration	Rapid turning	Rapid stopping	Rapid lane changing	Rapid overtaking	Rapid U-turning	Rapid starting	Long-term over-speeding
A	1	3,157.4	15,836.4	243.8	1,393.9	0.7372	0.8259	0.0001	5.8394	0	0
B	1	355.85	11.414	382.44	2,468.4	21.530	75.958	0	129.78	0	0
	2	242.01	7.5884	263.45	228.97	15.245	42.860	0.0368	86.534	0	0
C	1	5,697.7	0.0004	917.76	1,241.2	18.138	16.452	0.4746	57.896	0	0
	2	733.55	0	547.32	378.12	10.072	2.5321	8.0187	5.2734	0	0
	3	852.56	0.0003	331.79	943.05	49.444	11.642	7.9346	29.365	0	0

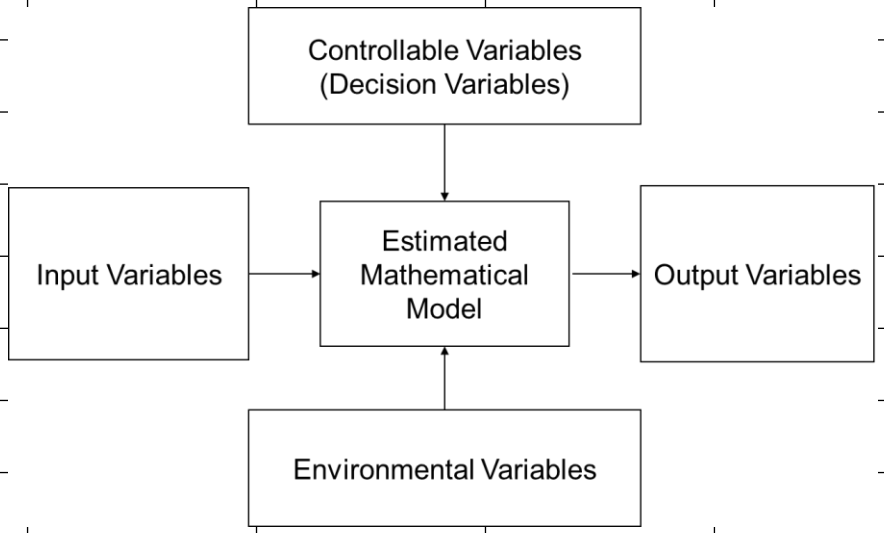






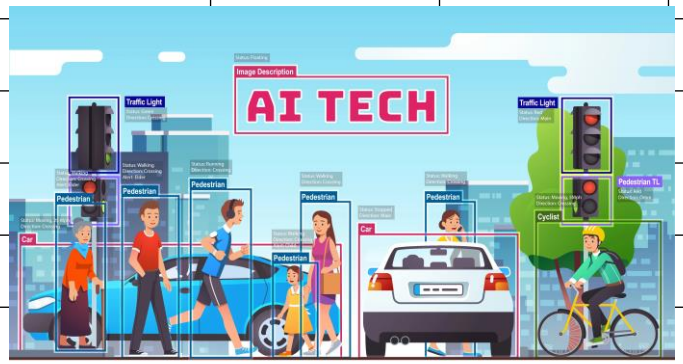
Behavioral Data for Personal Process Management

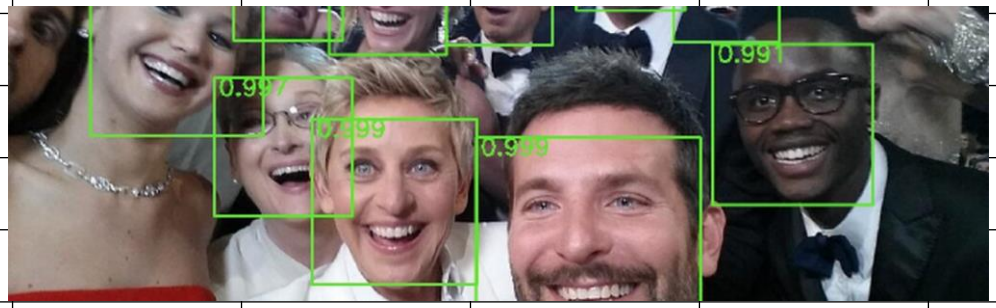
	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2	…					…
Customer 3	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…					…
…	…	…	…	…	…	…
Customer n-1	…	…	…	…	…	…
Customer n	…	…	…	…	…	…



Behavioral Data for Personal Process Management

	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2						
Customer 3						
…						
…						
…						
…	…	…	…	…	…	…
…	…					…
…	…					…
…	…					…
Customer n-1	…					…
Customer n	…	…	…	…	…	…





Existence vs. Nonexistence of the Labels

	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2						
Customer 3						
…						
…						
…						
…						
…						
…						
…						
Customer n-1	…	…	…	…	…	…
Customer n	…	…	…	…	…	…

Input Data
(Pre-processed)

Data \ Type	Variable (x_t)	Time window
Behavior data	Behavior 1	$S1_t, S1_{t-1}, \ldots$
	Behavior 2	$S2_t, S2_{t-1}, \ldots$
	…	$S3_t, S3_{t-1}, \ldots$
	Behavior 1	$Sk_t, Sk_{t-1}, \ldots$
	Data smoothing	
	$Mean(\sum_{i=0}^{n} x_{t-i})$	

Model

$$y = f(\mathbf{x}, \boldsymbol{\theta}) + \varepsilon$$

neural networks

Nonlinearity consideration

Decision trees

Rule-based explanation

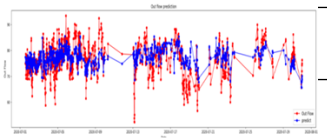
Regression

Linear coefficient estimation

Result

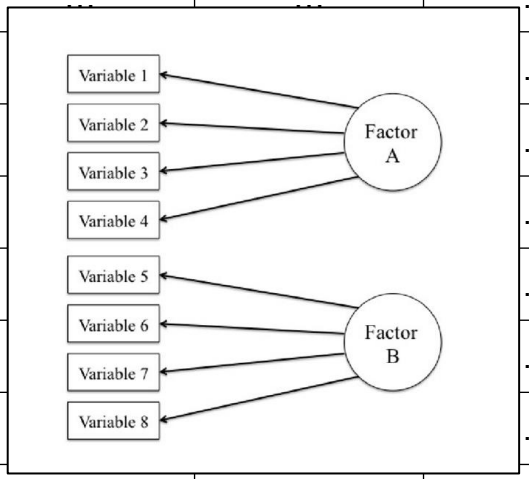
$$\hat{y}(\mathbf{x})$$

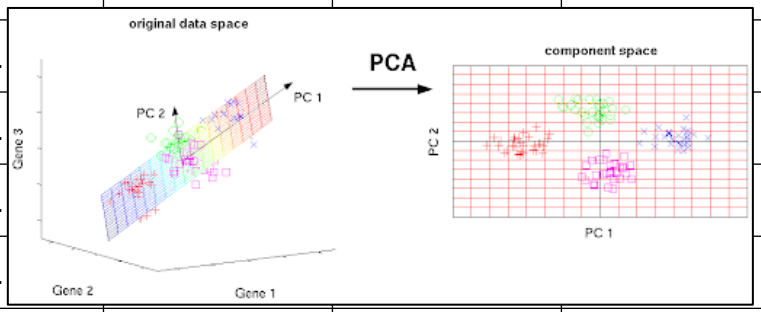
Behavioral outcomes

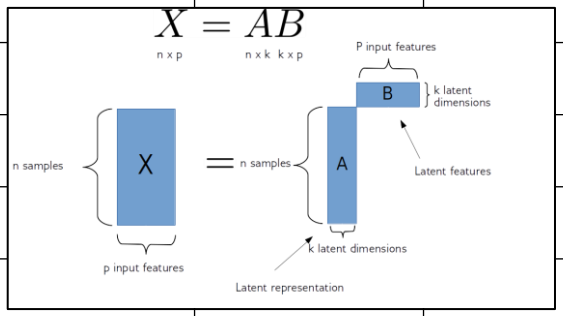


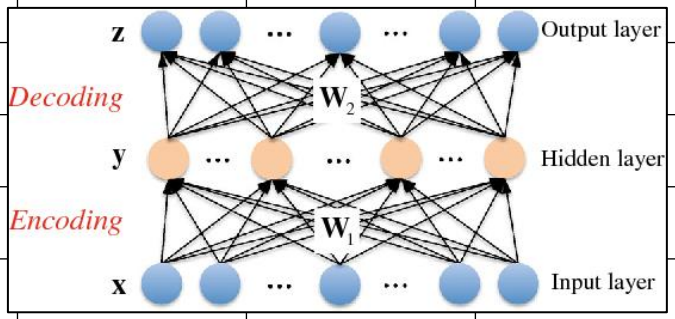
Existence vs. Nonexistence of the Labels

	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2			…	…	…	…
Customer 3			…			
…			…			
…			…			
…			…			
…			…	…	…	…
…						
…						
…						
Customer n-1						
Customer n						



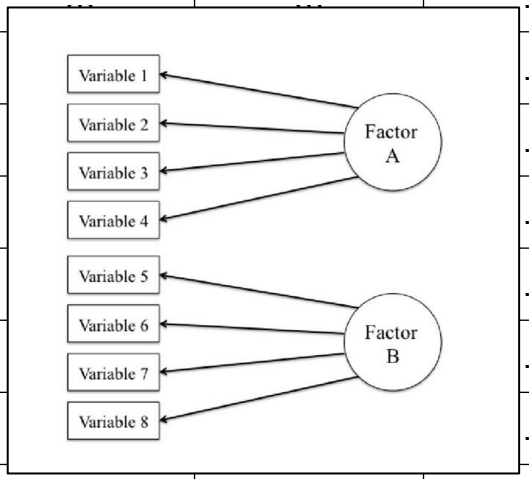


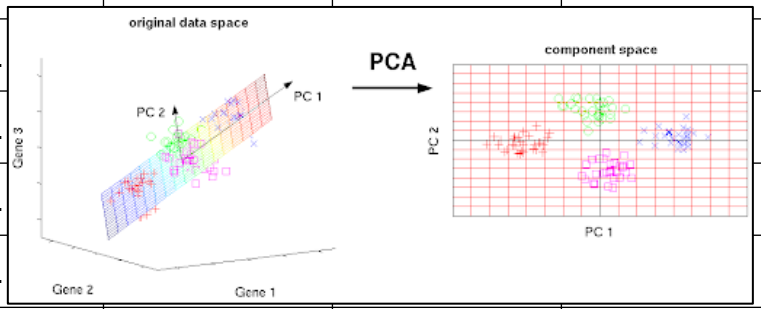


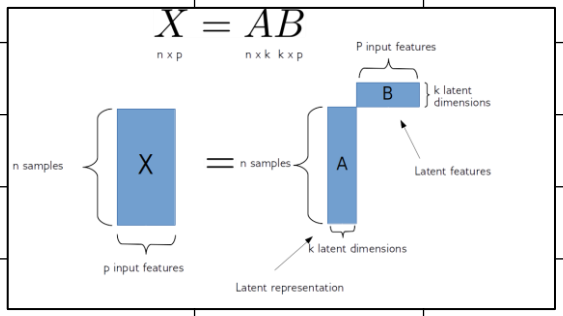


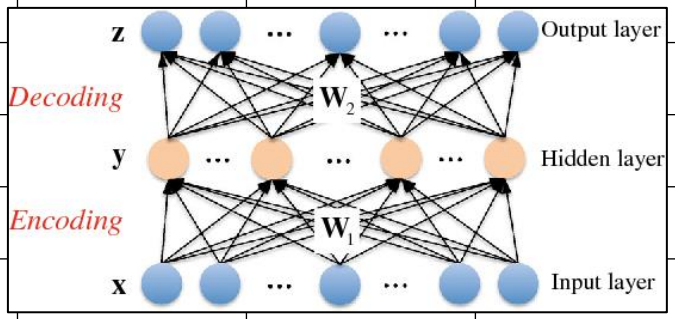
How Can We Validate the Case of Nonexistence of the Labels?

	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2			…	…	…	…
Customer 3			…			
…			…			
…			…			
…			…			
…			…	…	…	…
…						
…						
…						
Customer n-1						
Customer n						









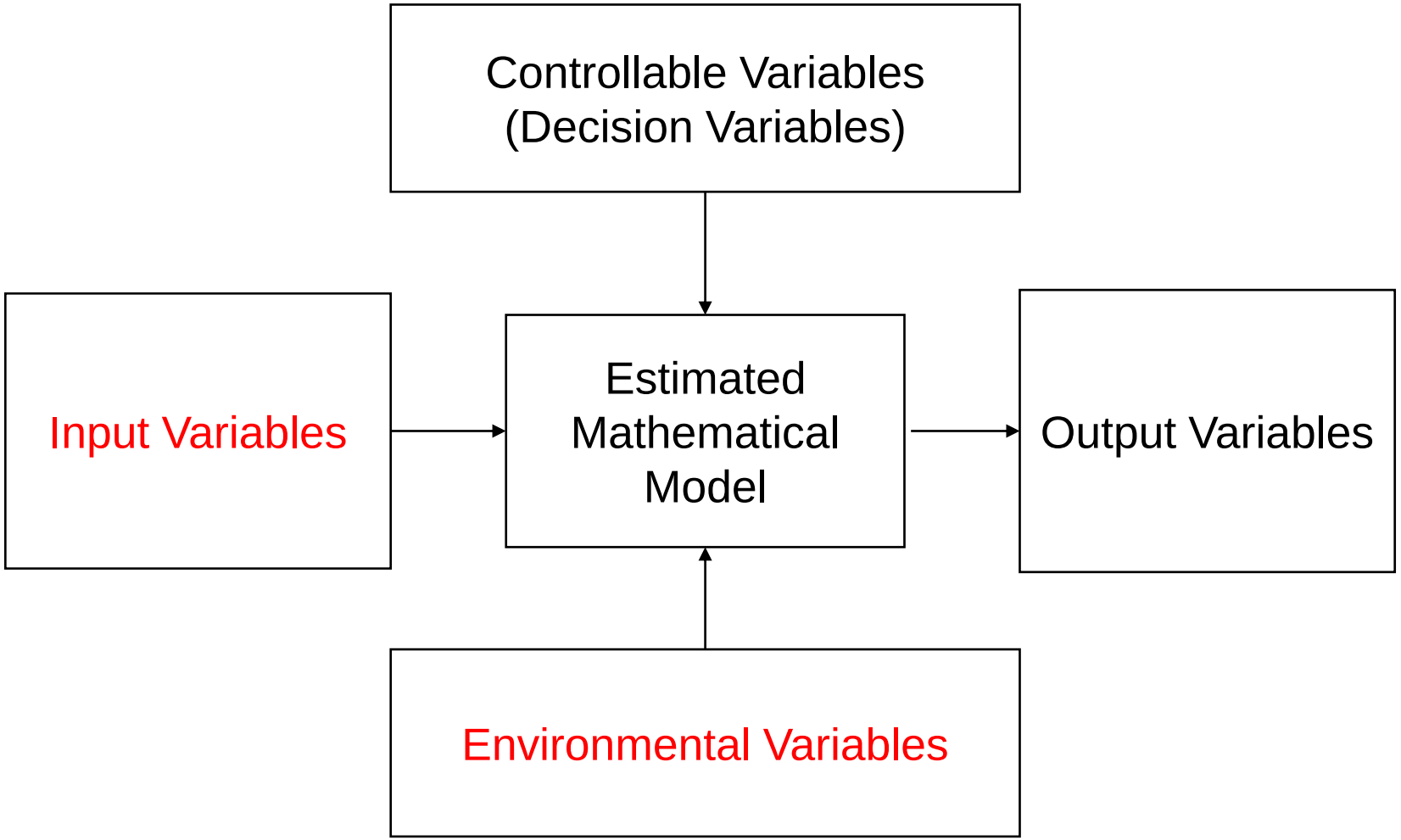
Assignment 5 (by 10.28 11:59 pm)

- Based on the practice demonstration from TA Jung, (1) complete the construction of the behavior feature – outcome label matrix in the context of eco-driving by yourself. You should select significant behavior features (variables) based on your own descriptive analysis and eco-driving domain knowledge study.
- Then, (2) using the behavior feature – outcome label matrix you constructed, develop an eco driving score estimation model. Do it all by yourself, and describe the analysis process and outcome in detail. Interpret the outcome (e.g., describe the behavior features you identified significant, interpret the coefficient/importance values of behavior features to the FPK label).
- (3) Assume you need to use your machine for real-world driving support service (e.g., T map service or Hyundai blueLink service). How can you improve your machine to be used for the service effectively? For example, what kinds of data should you collect and use further to develop a complete driving support service? How would you design a method for using/learning the data? Think beyond these examples in your own creative, unique way!
- (4) You must have your own interested or favorite service WITHOUT a machine for data-driven personal process management (e.g., traditional services that do utilize customers' behavioral data yet). Discuss the requirements of personal process management for the service in detail.
- (5) If you would conduct a study on developing a machine for data-driven personal process management for the service, how would you conduct the research in your own creative, unique way? What kinds of data are you going to collect, analyze, and learn, and what methods are you going to use? Describe your service intelligence development plan in detail. If possible, visualize your plan clearly (e.g., draw an image, construct a mathematical model). To facilitate your thinking, you may want to identify and review a paper related to the service you are interested or concerned.
- Upload your code and a several paragraph essay on the tasks (1)~(5) in the Blackboard.

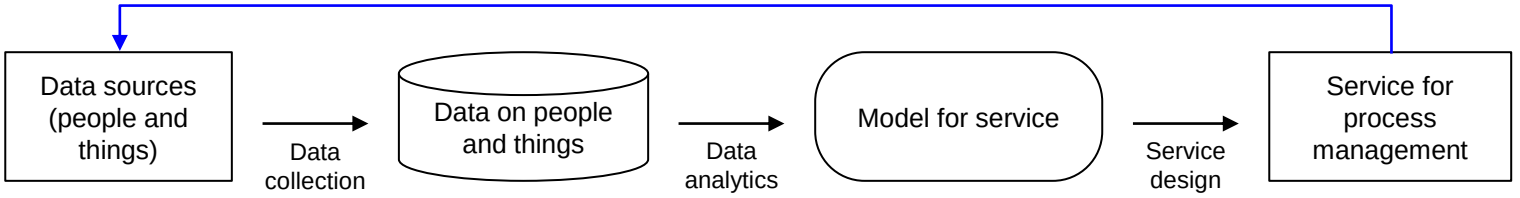
How Can We Address the Case of Existence of Few Labels?

	Behavioral feature 1	Behavioral feature 2	Behavioral feature 3	…	Behavioral feature m	Outcome label
Customer 1	…	…	…	…	…	…
Customer 2	…	…	…	…	…	…
Customer 3	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
Customer n-1	…	…	…	…	…	…
Customer n	…	…	…	…	…	…

How Can We Address Consider the Input and Environmental Variations?



Personal Process Management: Challenges



Collecting the right data for service

Facilitating easy and spontaneous data collection from people and objects

Managing the quality of data for service

Planning data analytics for service

Integrating different data from various sources for service

Following the regulations on data usage

Protecting the customer privacy in using their data

Identifying the information useful to customers

Promoting incentives for customers to accept new information

Designing an effective and efficient data-based service delivery process

Mediating the different stakes of data-related stakeholders for service

Reference: Lim, C., Kim, M., Kim, K., Kim, K. and Maglio, P., "Using Data to Advance Service: Managerial Issues and Theoretical Implications from Action Research," *Journal of Service Theory and Practice*, Vol. 28, No. 1, 99-128, 2018.

Expected Schedule after the Midterm

- Week 8: Midterm Week
- Week 9
  - 10/24 Mon: Term Project Announcement + Service Process Assessment + Industrial Service Intelligence
  - 10/26 Wed: Practices for Service Process Assessment and Industrial Service Intelligence + A Special Lecture
- Week 10
  - 10/31 Mon: Metaheuristics for Service Optimization & Data Envelopment Analysis (DEA) for Service Improvement
  - 11/2 Wed: Practices for Genetic Algorithm and DEA
- Week 11
  - 11/7 Mon: Term Project Proposals I
  - 11/9 Wed: Term Project Proposals II

Eco-driving Practice Demonstration from TA Jung