

Types of Machine Learning

Machine Learning
CSx824/ECEx242
Bert Huang
Virginia Tech

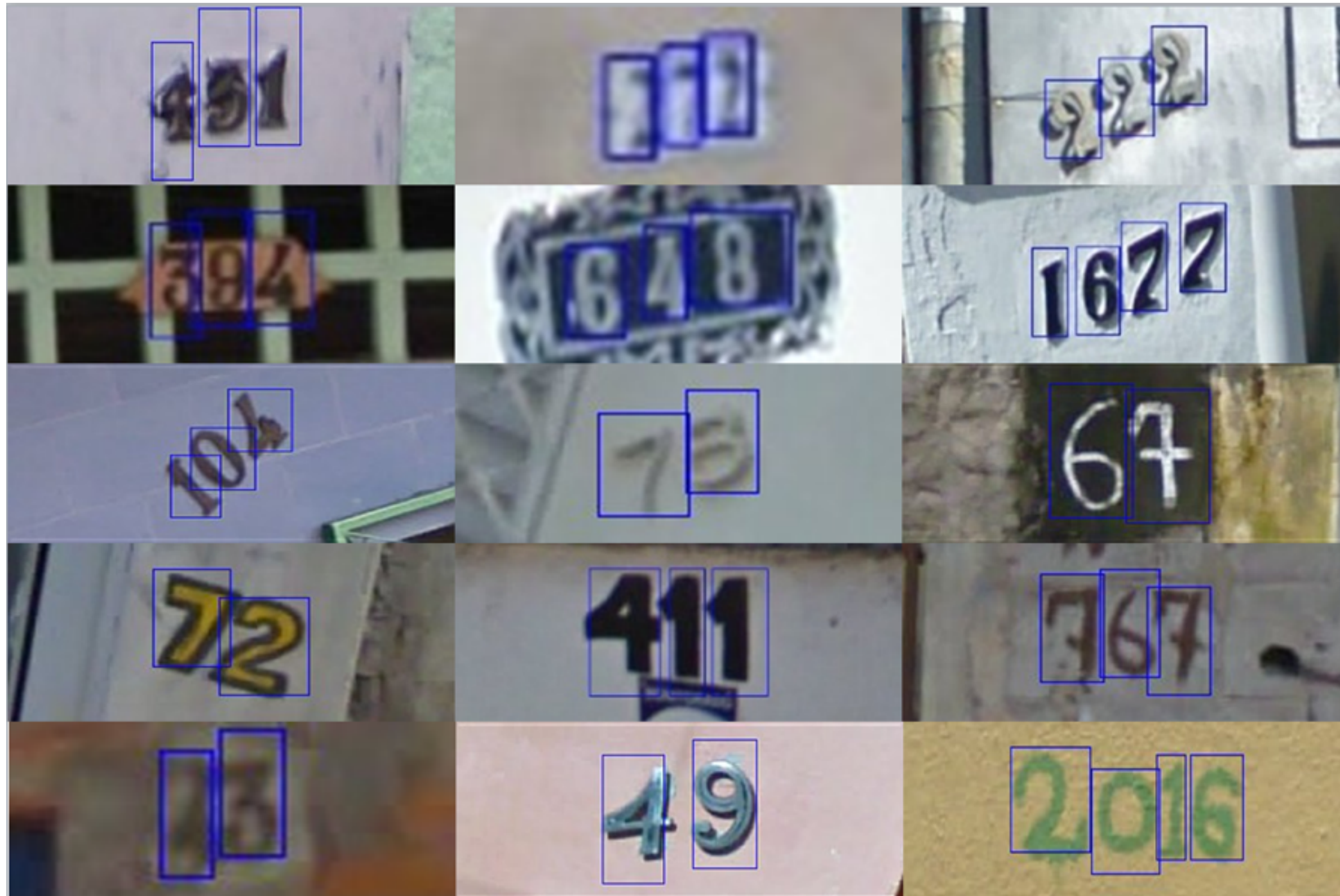
Outline

- Types of learning settings
 - Supervised learning
 - Unsupervised learning
- Types of learning algorithms
 - Batch learning
 - Online learning

1st Learning Setting

- Draw data set $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ from distribution \mathbb{D}
- Algorithm A learns hypothesis $h \in H$ from set H of possible hypotheses $A(D) = h$
- We measure the quality of h as the expected **loss**: $E_{(x,y) \in \mathbb{D}} [\ell(y, h(x))]$
 - This quantity is known as the **risk**
 - E.g., loss could be the Hamming loss $\ell_{\text{Hamming}}(a, b) = \begin{cases} 0 & \text{if } a = b \\ 1 & \text{otherwise} \end{cases}$

Example: Digit Classification



<http://ufldl.stanford.edu/housenumbers/>

Example: Airline Price Prediction

The screenshot displays the Kayak website interface for a flight search. The search parameters are: CLT ↔ HNL, Aug 28 Friday to Aug 28 Friday, Economy cabin, 1 traveler. The results are sorted by price (low to high), showing 527 of 533 flights. The top result is a \$367 Honolulu Round Trip advertisement. Below it, two flight options are listed, both priced at \$732. The first is a US Airways flight with a 1-stop itinerary (PHX) and a total duration of 11h 55m. The second is an American Airlines flight with a 1-stop itinerary (DFW) and a total duration of 12h 12m. The left sidebar contains a price prediction graph, a 'Create a price alert' button, and filters for stops (1 stop selected) and times.

KAYAK HOTELS FLIGHTS CARS PACKAGES Login

CLT ↔ HNL | Aug 28 Friday → Aug 28 Friday | Economy cabin | 1 traveler | [Change](#)

Sort by: price (low to high) ▼ | 527 of 533 flights | Round-trip | Segment **NEW**

\$367 Honolulu Round Trip ads
[cheapoair.com/Honolulu-Cheap-Flight](#)
Book Discounted Fares Today & Save! Cheap Fares on Flights to Honolulu.
Search, Select & Save Big · We Make it Easy to Travel · Our Best Price Guarantee · 24/7 Customer Care
Winner - 2014 Customer Focused Innovations Award - CSIA

\$732
US Airways

11:35a CLT → **5:30p** HNL 11h 55m 1 stop (PHX)
9:05p HNL → **1:35p** CLT 10h 30m 1 stop (PHX)

[Select](#) | [Show details](#) ▼ | Economy

\$732
American Airlines

6:10a CLT → **12:22p** HNL 12h 12m 1 stop (DFW)
9:05p HNL → **1:35p** CLT 10h 30m 1 stop (PHX)

[Select](#)

Advice: BUY Confidence: 80%
Prices may rise within 7 days ⓘ

[Create a price alert](#)

Stops

nonstop
 1 stop \$732
 2+ stops \$736

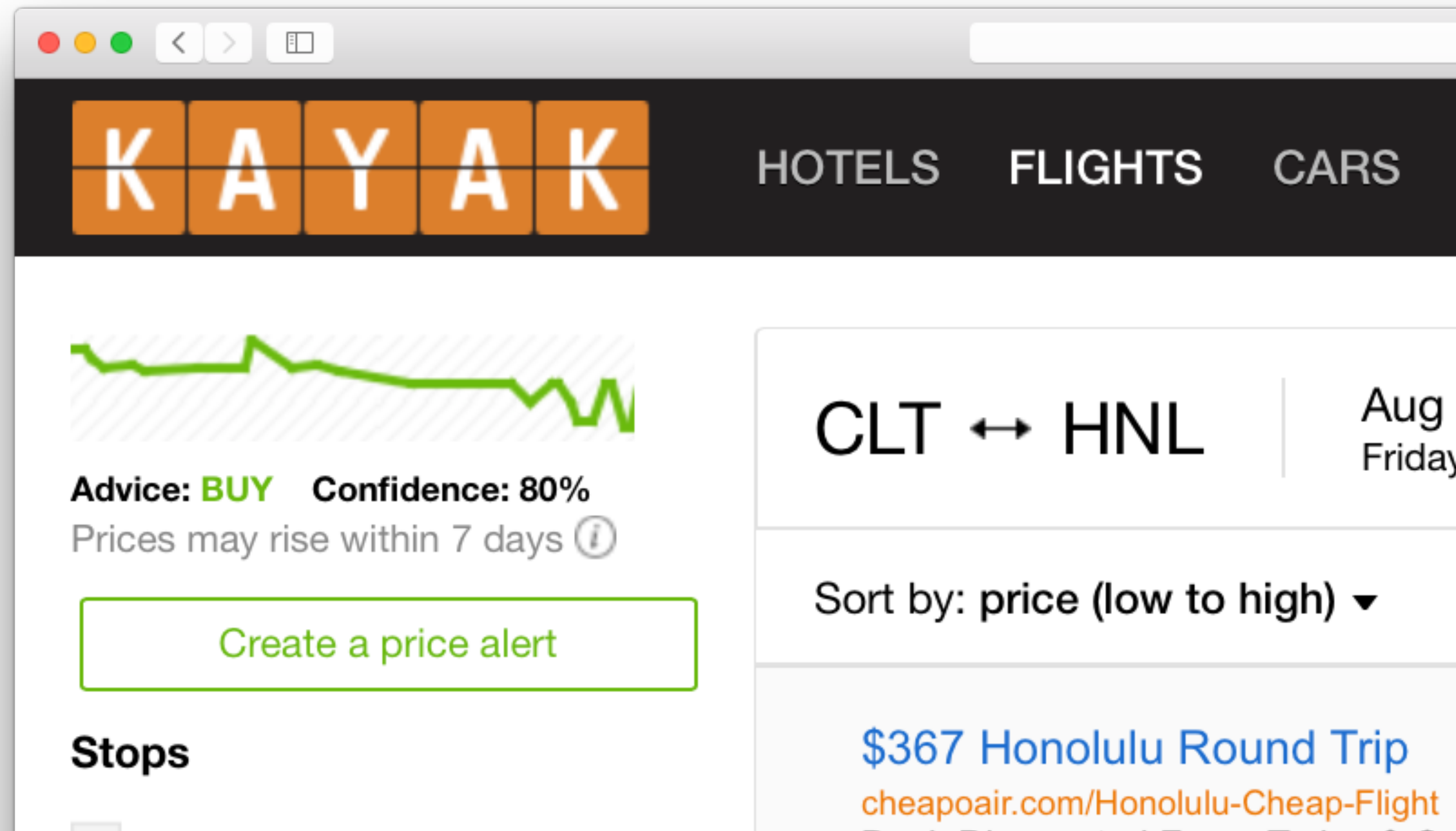
Times

Take-off **Charlotte (CLT)**
Fri 5:00a - 2:30p


Take-off **Honolulu (HNL)**
Fri 2:30p - Sat 12:00a

[Show landing times](#) ▼

Example: Airline Price Prediction



KAYAK HOTELS FLIGHTS CARS



Advice: BUY Confidence: 80%
Prices may rise within 7 days ⓘ

Create a price alert

Stops

CLT ↔ HNL | Aug Friday

Sort by: price (low to high) ▼

\$367 Honolulu Round Trip
cheapoair.com/Honolulu-Cheap-Flight

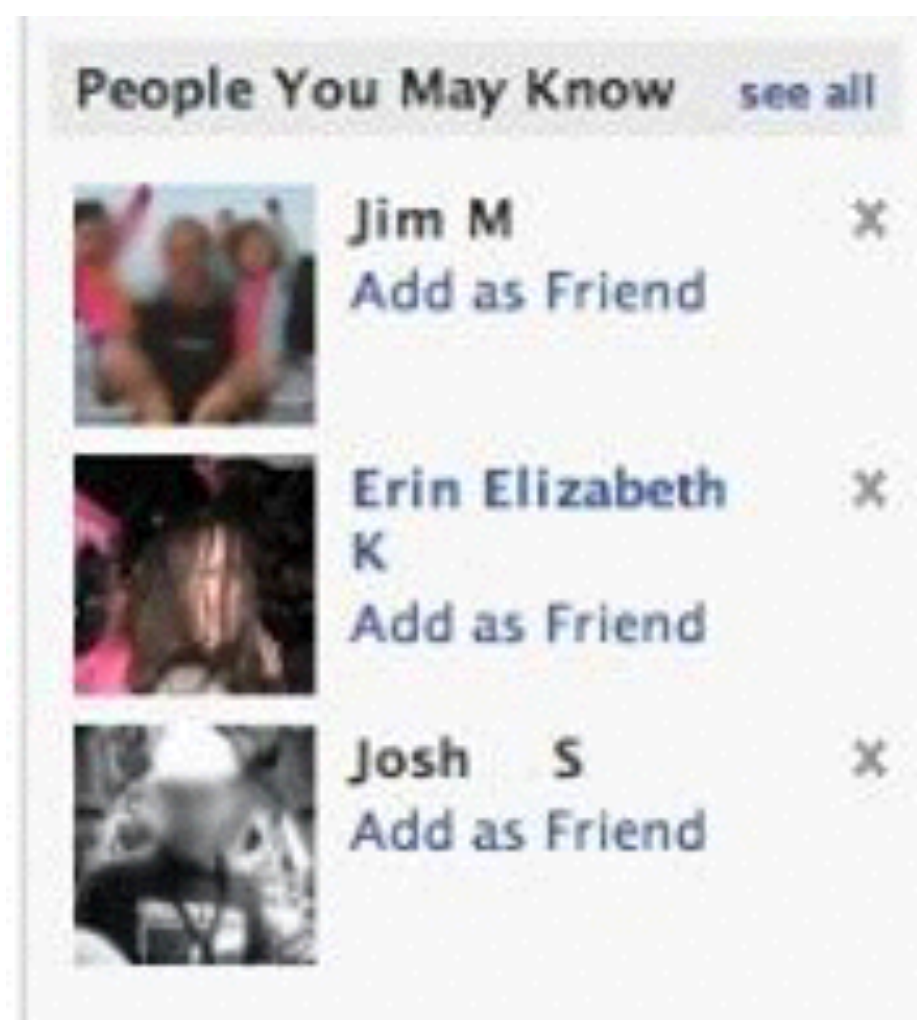
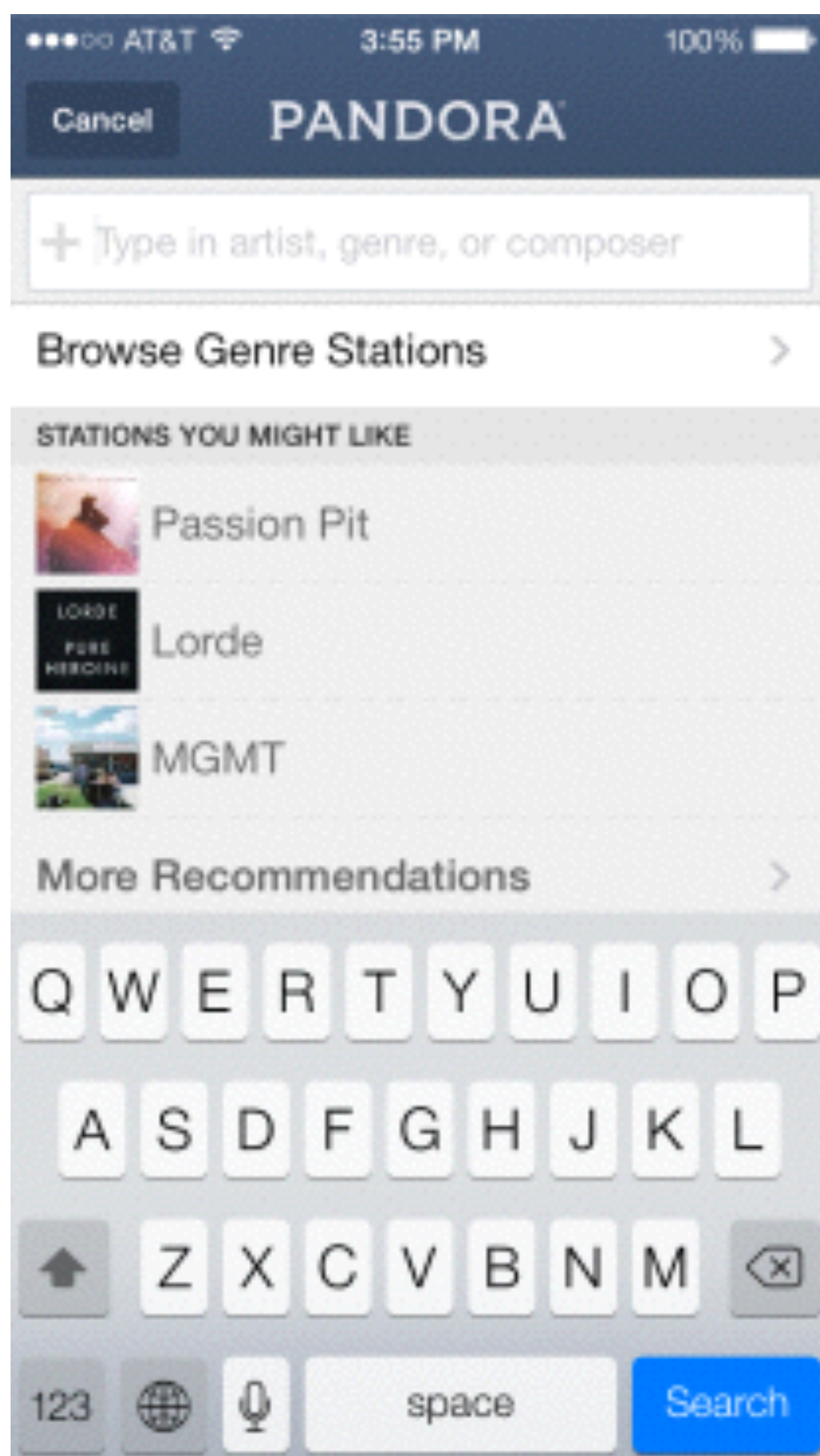
Batch Supervised Learning

- Draw data set $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ from distribution \mathbb{D}
- Algorithm A learns hypothesis $h \in H$ from set H of possible hypotheses $A(D) = h$
- We measure the quality of h as the expected **loss**: $E_{(x,y) \in \mathbb{D}} [\ell(y, h(x))]$
 - This quantity is known as the **risk**
 - E.g., loss could be the Hamming loss $\ell_{\text{Hamming}}(a, b) = \begin{cases} 0 & \text{if } a = b \\ 1 & \text{otherwise} \end{cases}$
classification

Online Supervised Learning

- In step t , draw data point \mathbf{x} from distribution \mathbb{D}
- Current hypothesis h guesses the label of \mathbf{x}
- Get true label from oracle \mathcal{O}
- Pay penalty if $h(\mathbf{x})$ is wrong (or earn reward if correct)
- Learning algorithm updates to new hypothesis based on this experience
 - Does not store history

Example: Recommendation



Recommended for You

These recommendations are based on items you own and more.

[All](#) | [New Releases](#) | [Coming Soon](#)



Cybertext: Perspectives on Ergodic Literature

by Espen J. Aarseth (Aug 6, 1997)
Average Customer Review: ★★★★★ (3)
In Stock

List Price: \$22.96

Price: **\$19.55**

[29 used & new from \\$10.82](#)

[Add to cart](#) [Add to](#)

I own it Not interested Rate it

Recommended because you added **Hamlet on the Holodeck** to your Shopping Cart and more ([Fix this](#))



Narrative as Virtual Reality: Immersion and Interactivity in Lit Media (Parallax: Re-visions of Culture and Society)

by Mark J. P. Sullivan (Oct 3, 2003)

Learning Settings

- Supervised or unsupervised (or semi-supervised, weakly supervised, transductive...)
- Online or batch (or reinforcement...)
- Classification, regression
 - (or structured output, clustering, dimensionality reduction...)
- Parametric or non-parameteric

Functional Perspective

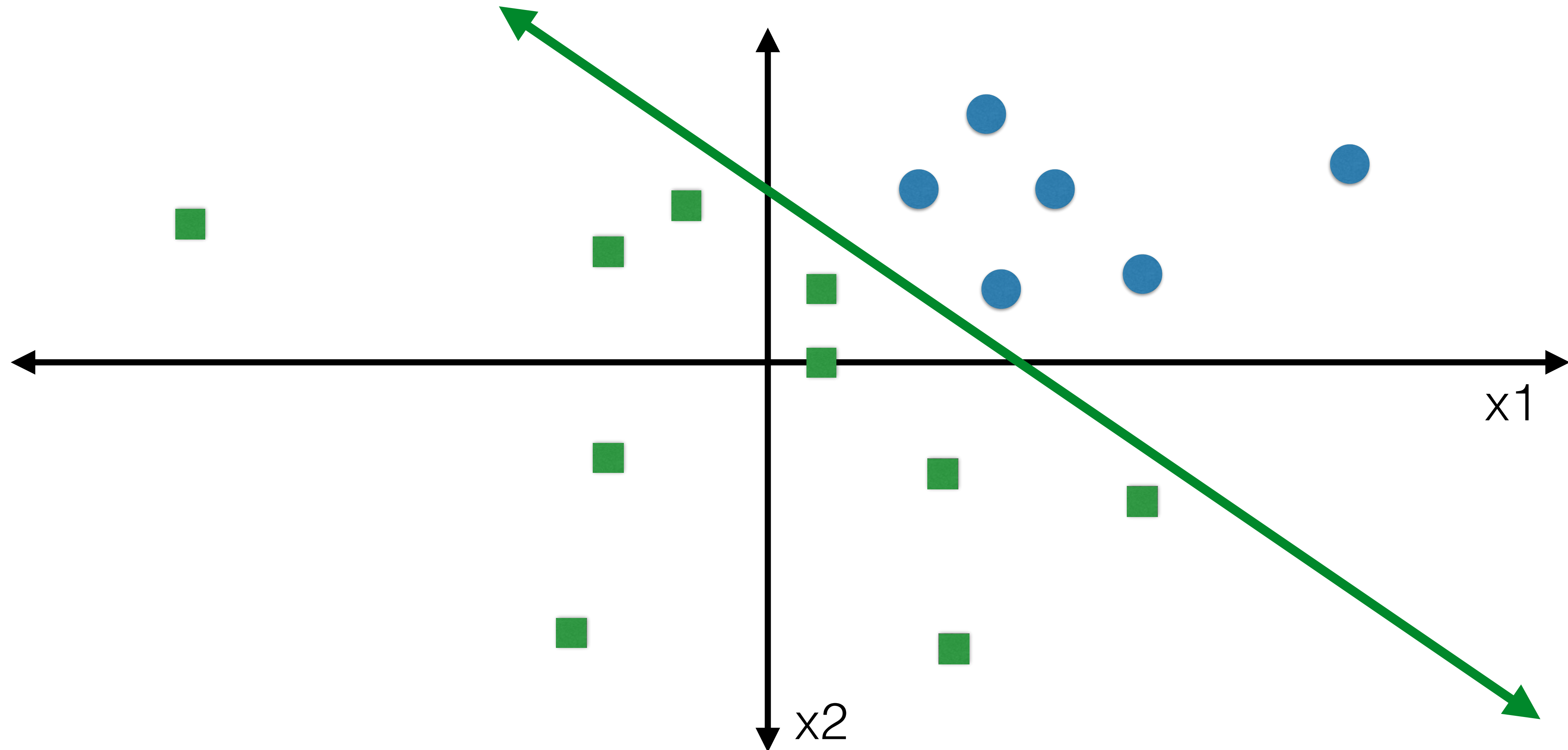
Input	Learning Setting
Batch of Data Points with Labels	Batch Supervised Learning
Batch of Data Points	Batch Unsupervised Learning
Data Point(s) and Previous Model	Online Supervised Learning

Parametric vs. Non-Parametric

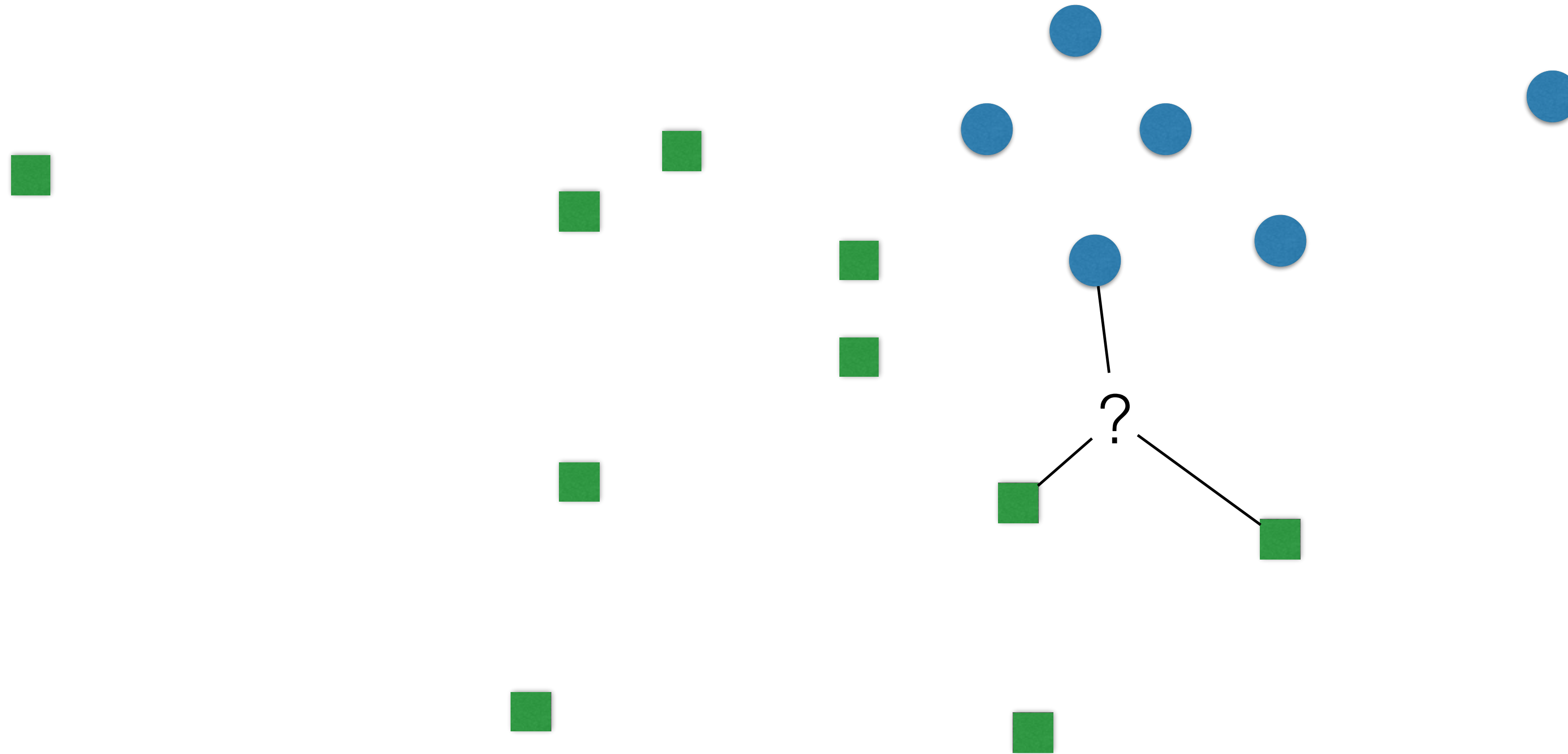
$$A(\mathcal{D}) = h \quad h \in H$$

- If $h \in H$ identifiable by finite-dimensional vector: parametric
 - e.g., linear classifier $h_{\theta}(\mathbf{x}) = \text{sign}(\boldsymbol{\theta}^{\top} \mathbf{x})$
- If $h \in H$ has a flexible number of parameters: non-parametric
 - e.g., k-nearest neighbor

Parametric Classifier: Linear



Non-Parametric Classifier: K-Nearest Neighbors



Learning Settings and Probability

	discriminative	generative
Supervised	$p(y x)$	$p(y, x)$
	$p(y x; \theta)$	$p(y, x; \theta)$
Unsupervised		$p(x)$
		$p(x; \theta)$

θ finite dimensional: parametric
non-parametric otherwise

Concepts

- Supervised and unsupervised learning
- Parametric and non-parametric learning algorithms
- Online and batch learning
- Discriminative and generative
- Output of models: classification and regression

Discussion Questions

- How do you characterize different machine learning algorithms you know about?
- Are learning-algorithm attributes independent? Are there combinations of attributes that fit well together or don't fit well?