

AI & Machine Learning Workshop

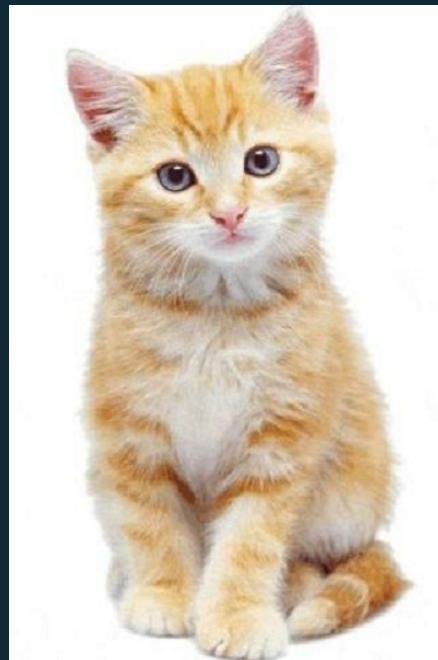
Indonesia Advanced Workshop Series

Yudho Ahmad Diponegoro (diponego@amazon.com)

23 May 2019

Agenda

- Machine Learning : Model, Training, Deploy
- AWS Services for Machine Learning
- Lab 1 : Person Identification and Sentiment Analysis
- Deep Learning and Amazon Personalize
- Lab 2 : Products Recommendation and Integration with Web Application
- Amazon SageMaker
- Lab 3 : Image Classification with SageMaker + Jupyter Notebook



A B C

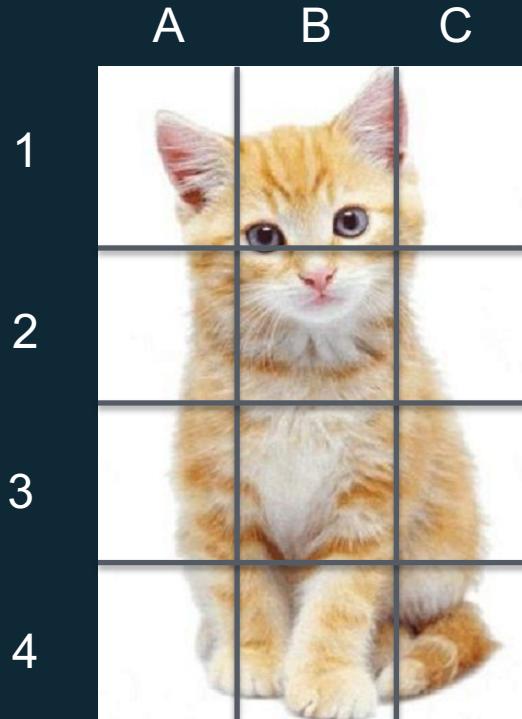
1

2

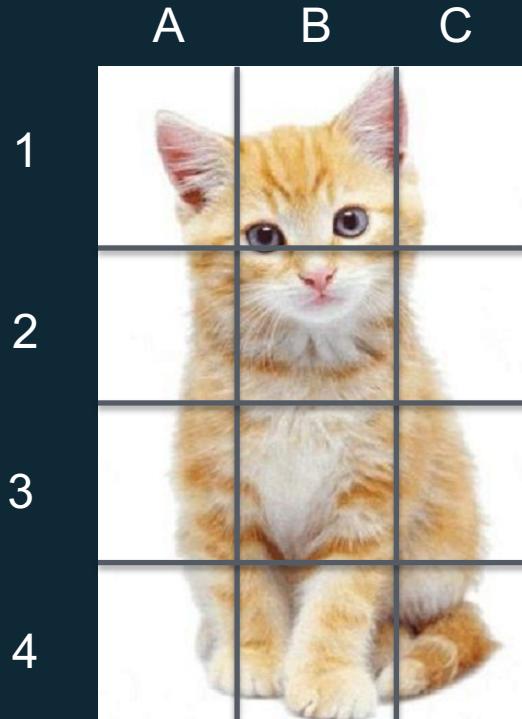
3

4

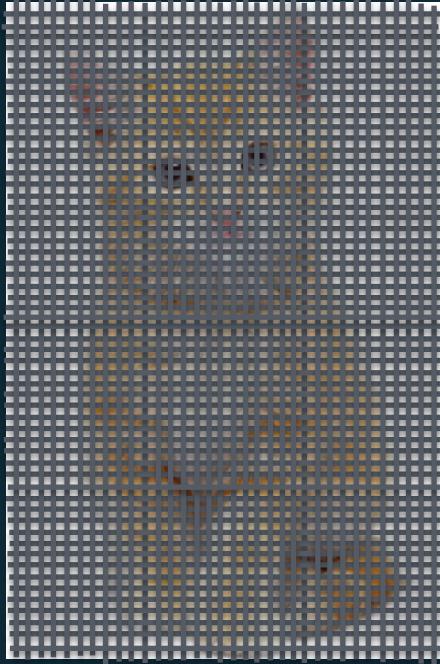




```
def is_cat:  
    ear    = has_ear_at(A1)  
    eyes   = has_eyes_at(B1)  
    legs   = has_legs_at(B4)  
    tail   = has_tail_at(C4)  
  
    if ear and eyes and legs and tail:  
        print("MEOW")
```



```
def is_cat:  
    ear    = has_ear_at(A1)  
    eyes   = has_eyes_at(B1)  
    legs   = has_legs_at(B4)  
    tail   = has_tail_at(C4)  
  
    if ear and eyes and legs and tail:  
        print("MEOW")
```



```
def is_cat:  
    . . . . . very long code . . . . .  
  
    if . . . thousands conditions . . .:  
        print("MEOW")
```



and the code gets much much longer.....





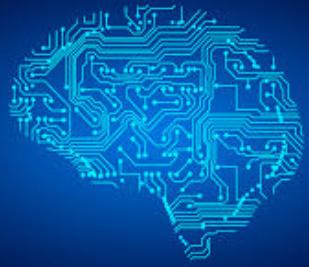
```
function use_array(a, b, c) {
    for (var d = 0; d < a.length; d++) {
        if (a[d] == b && c++ >= 1) {
            return c;
        }
    }
    return -1;
}

function indexOf_keyword(a, b) {
    for (var c = 0; c < a.length; c++) {
        if (a[c].word == b) {
            return c;
        }
    }
    return -1;
}

function dynamicSort(a) {
    var b = -1;
    a = a.substring(1);
    return function(c, d) {
        b = -1;
        if (c[a] > d[a] ? 1 : 0) {
            return b;
        }
        if (b >= a.length) {
            b = -1;
            a += "";
            b += "";
            if (b >= a.length) {
                return 1;
            }
            for (c = a.length - 1; c >= 0; c--) {
                if (a[c] != " ") {
                    break;
                }
            }
            if (c < 0) {
                return 1;
            }
            if (c >= a.length) {
                return -1;
            }
            if (a[c] >= "0" & a[c] <= "9") {
                a = a.substring(0, c + 1);
                a += "0";
                a += a.substring(c + 1);
            }
            a = a.substring(0, c + 1);
            a += "0";
            a += a.substring(c + 1);
        }
        if (b <= a.length - 1) {
            if (a[b] >= "0" & a[b] <= "9") {
                a = a.substring(0, b + 1);
                a += "0";
                a += a.substring(b + 1);
            }
            a = a.substring(0, b + 1);
            a += "0";
            a += a.substring(b + 1);
        }
        if (a[c] > a[d]) {
            return 1;
        } else {
            return -1;
        }
    };
}
```

Programming

Code

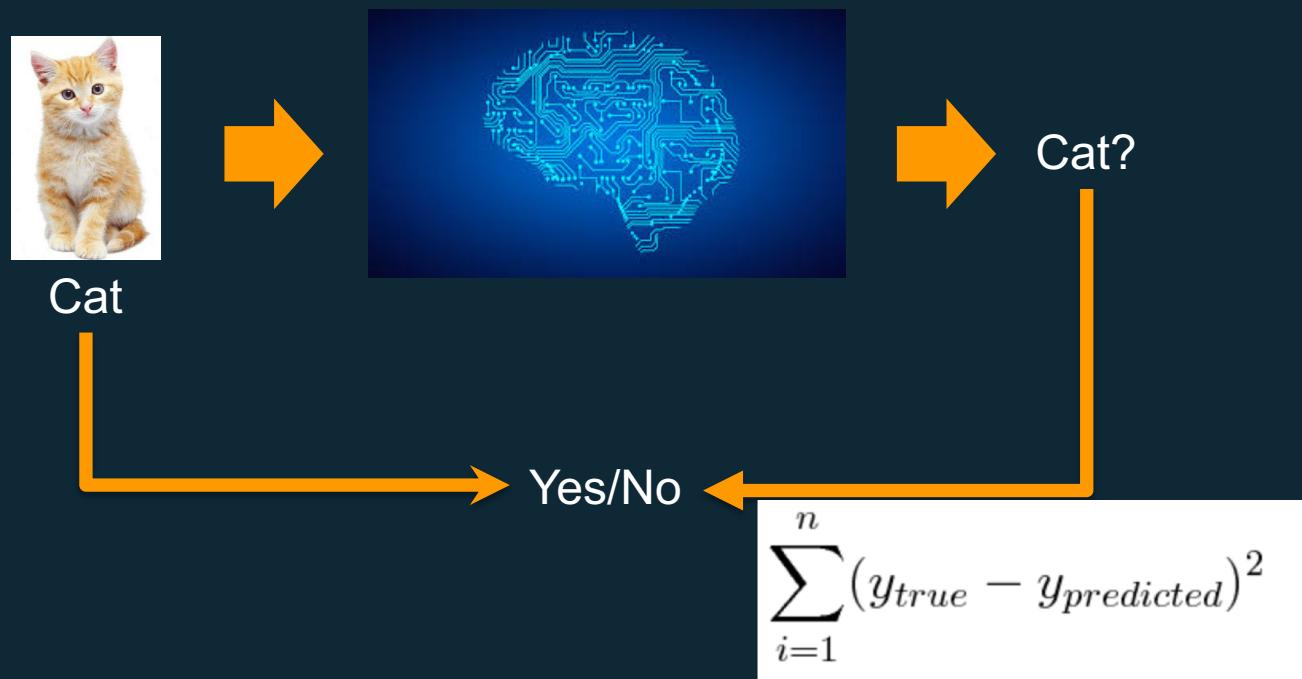


Machine Learning

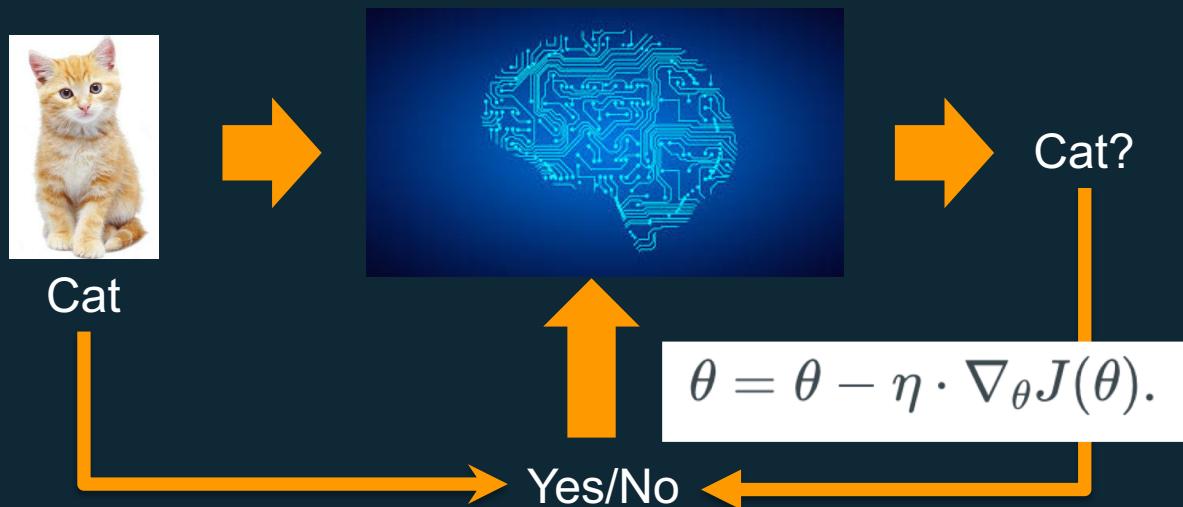
Code

Data

The Concept of Training

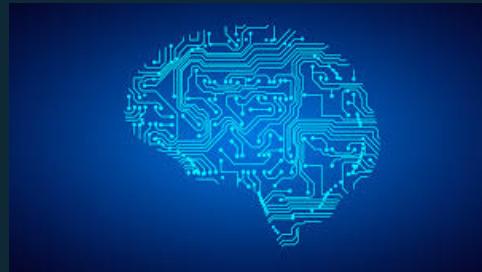


The Concept of Training



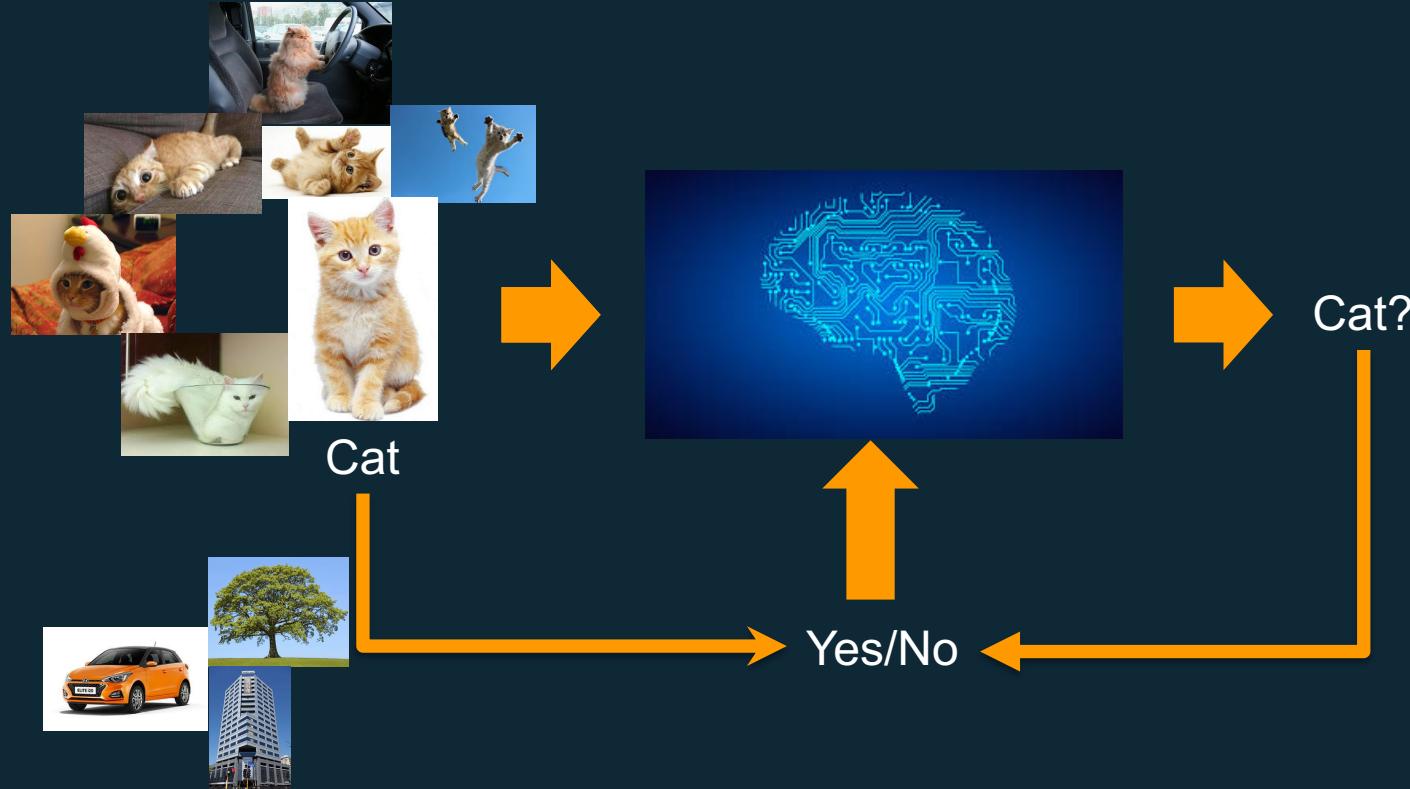
The Concept of Model

Machine Learning Model



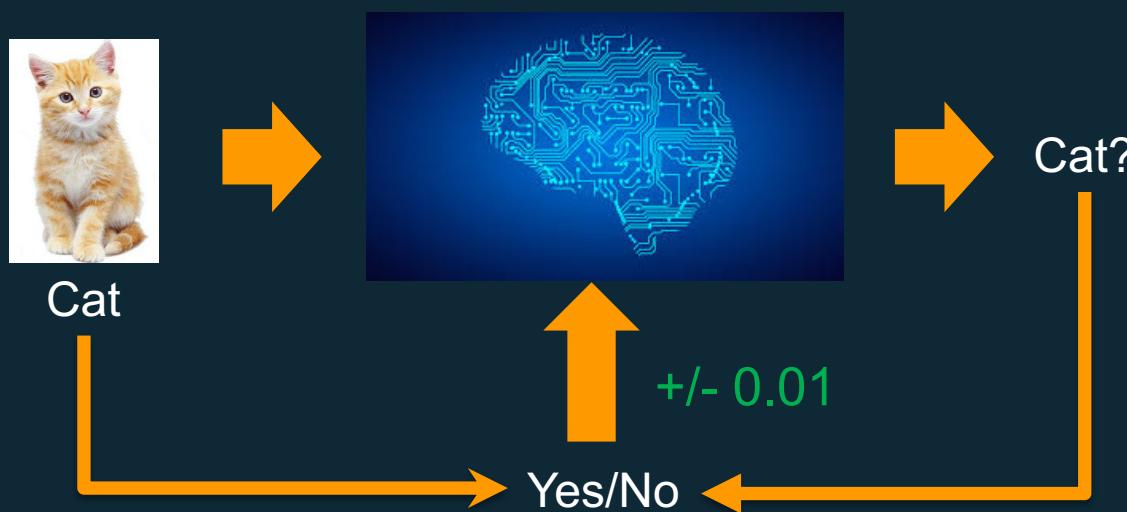
Algorithm + Parameters

A lot of data.....



Hyperparameters

Bias: 0.001
Init: Bayes



Some Machine Learning Algorithm

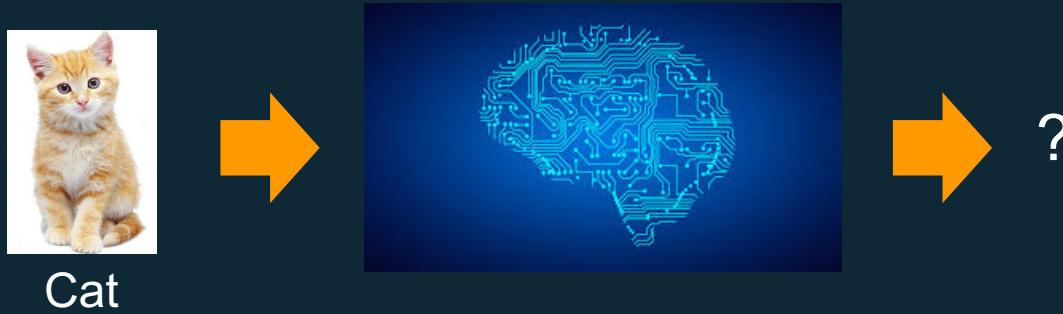
■ Use Built-in Algorithms

- + Common Information
- + BlazingText
- + DeepAR Forecasting
- + Factorization Machines
- + Image Classification Algorithm
- + IP Insights
- + K-Means Algorithm
- + K-Nearest Neighbors (k-NN) Algorithm
- + Latent Dirichlet Allocation (LDA)

+ Linear Learner Algorithm

- + Neural Topic Model (NTM) Algorithm
- + Object2Vec
- + Object Detection Algorithm
- + Principal Component Analysis (PCA) Algorithm
- + Random Cut Forest (RCF) Algorithm
- + Semantic Segmentation
- + Sequence to Sequence (seq2seq)
- + XGBoost Algorithm
- + Use Your Own Algorithms

Inference



Why Machine Learning & AI ?

Find the missing kid quicker!



“Law enforcement needs sophisticated tools to foster victim-oriented policing in the age of the Internet. Law enforcement knew that runaway children are among the most likely to be trafficked. Before using Amazon Rekognition, their only recourse was to manually sift through online data to try to find them; this was time-intensive or not possible. Now with Traffic Jam’s FaceSearch, powered by Amazon Rekognition, investigators are able to take effective action by searching through millions of records in seconds to find victims.”

- Emily Kennedy, CEO and Founder, Marinus Analytics

Amazon's recommender increased sales?



8% - up to 30% of page views on Amazon are from recommendations

Frequently bought together



Total price: EUR 26,12

Add all three to Basket

i These items are dispatched from and sold by different sellers. [Show details](#)

- This item: Beto Floor Pump - Silver EUR 14,25
- BBB Valve Kit BFP-90 Bike pump accessories EUR 5,99
- LS Set of 3 Bicycle Valve Sv-Valve Adaptor Adapter AV DV New EUR 5,88

29% sales increase in the year implementing recommender system

Chatbot for registration process

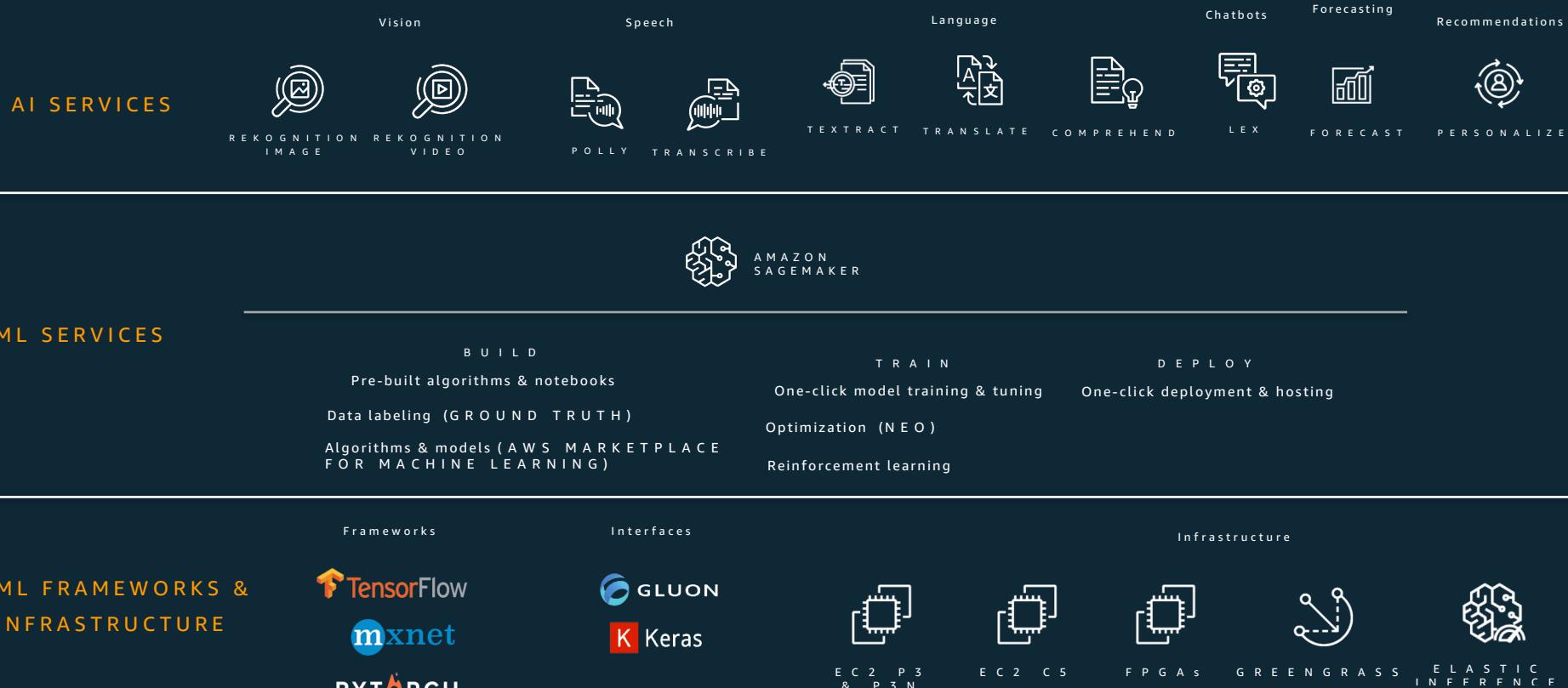


We used Amazon Lex's AI technology to streamline the **registration process** so prospective Heart Walk participants can use their **natural voice to easily register** through heartwalk.org.

Roger Santone, Executive Vice President of Technology at the AHA

AWS for AI & Machine Learning

The Amazon ML stack: Broadest & deepest set of capabilities



Hands-on Time!

Person Identification and Sentiment Analysis

<https://bit.ly/2waY7aq>

Enough face recognition...

Next is recommender engine

Recommender Types

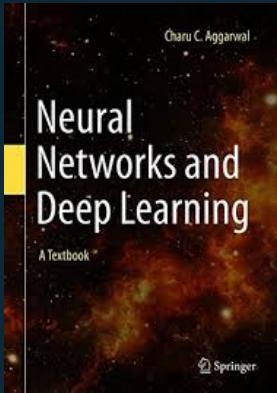
1. Collaborative Filtering

2. Content-based System

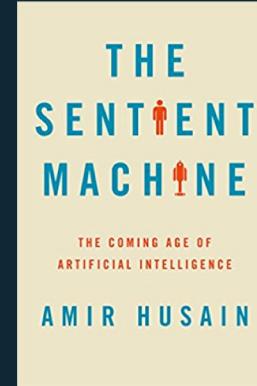
3. Hybrid

Collaborative Filtering

“Rudi likes item **A** and Aris also likes item **A**. Rudi likes item **B**, so Eryan might like item **B** as well.”



A



B

Collaborative filtering

Explicit:

- Asking a user to rate an item
- Asking a user to rank a collection of items from favorite to least favorite
- ...

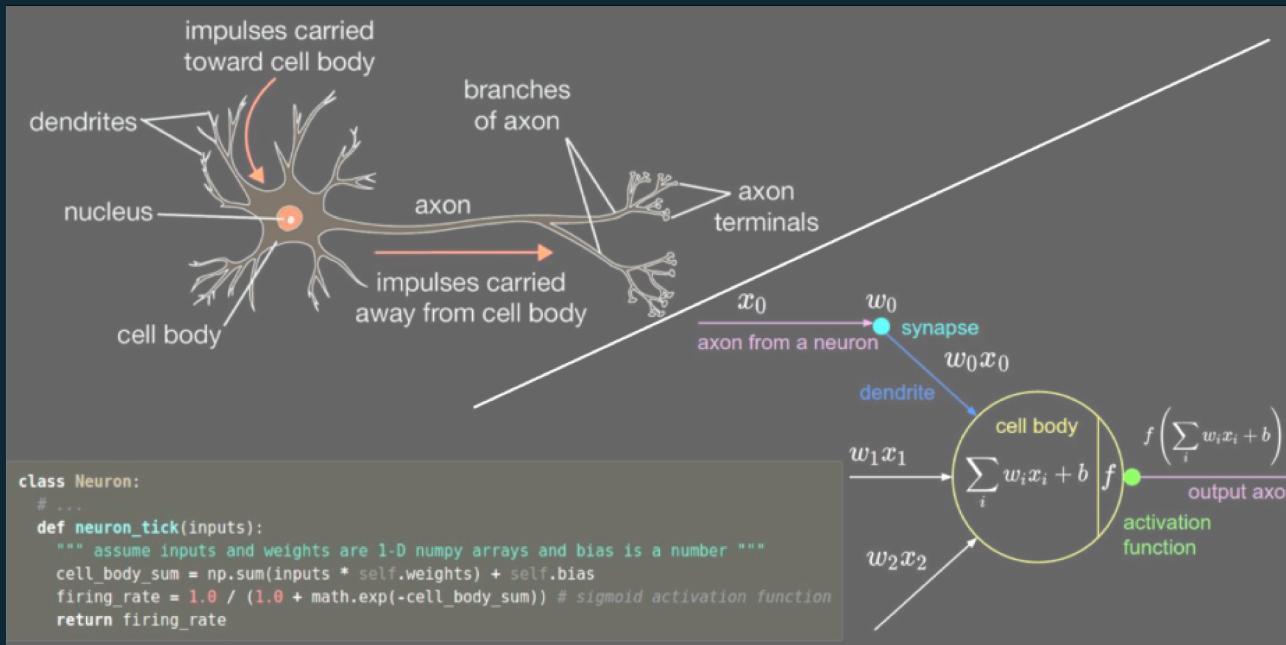
Implicit:

- Observing the items that a user views
- Keeping a record of the items that a user purchases
- ...

Let's step back a bit...

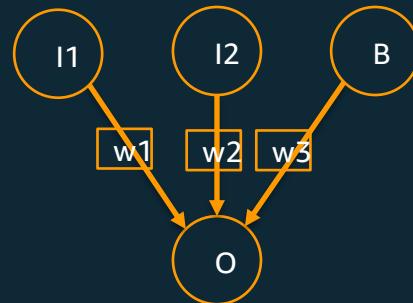
What is the Math behind?

Deep Learning: Biological learning



Source: <http://cs231n.github.io/neural-networks-1/>

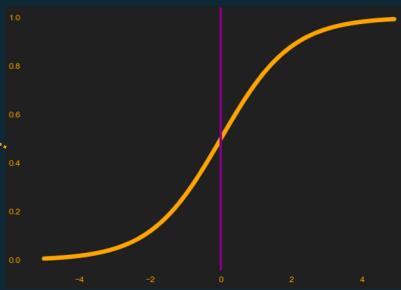
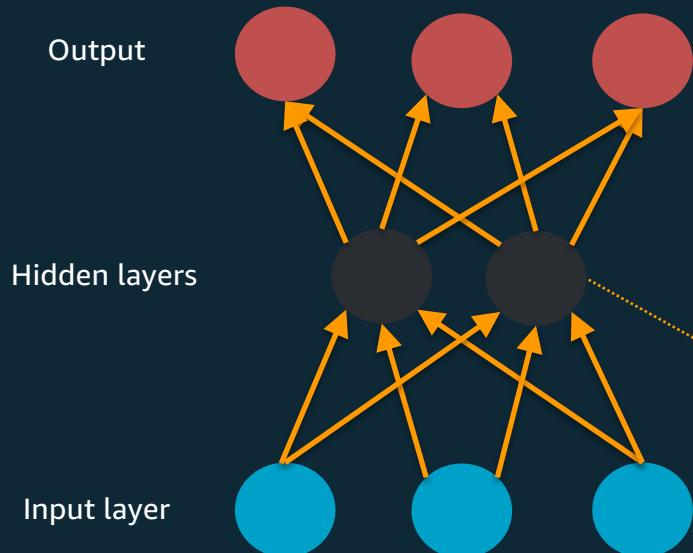
Deep Learning: Perceptron



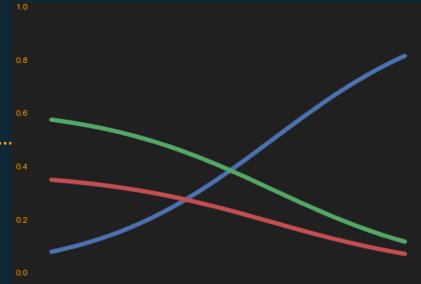
Activation function;
adding non-linearity

$$f(x_i, w_i) = \Phi(b + \sum_i(w_i \cdot x_i))$$
$$\Phi(x) = \begin{cases} 1, & \text{if } x \geq 0.5 \\ 0, & \text{if } x < 0.5 \end{cases}$$

Deep learning

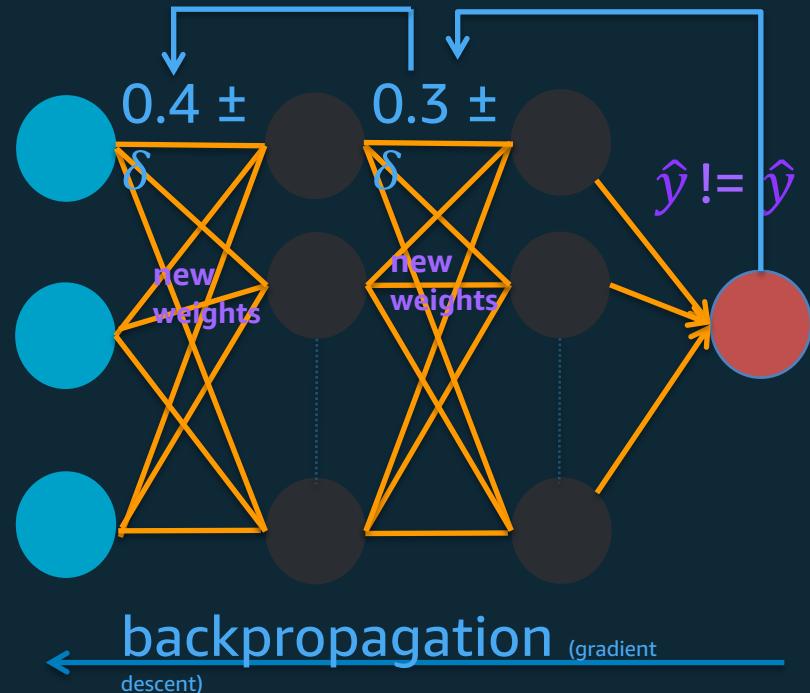
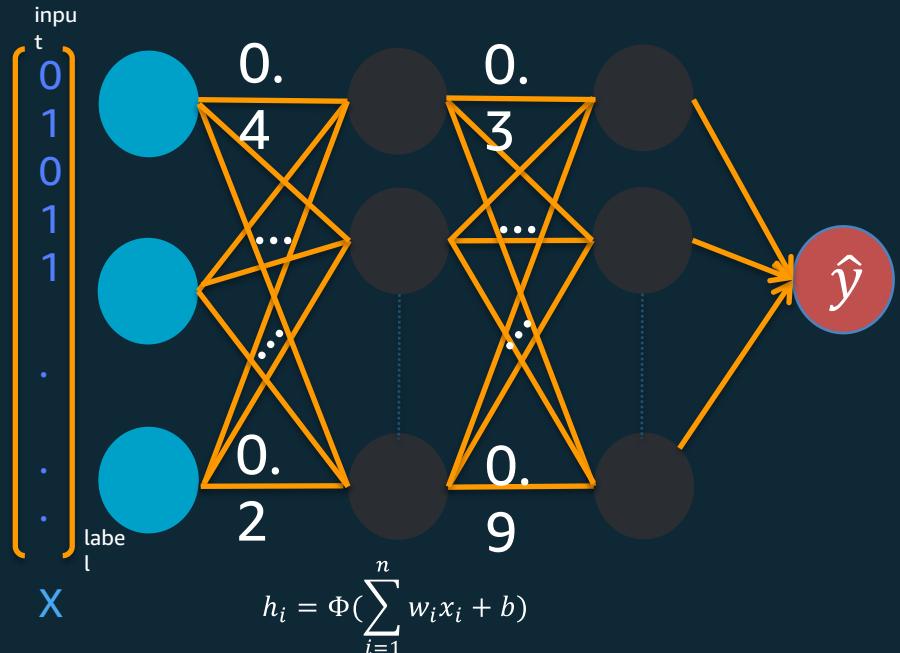


Activation



Discriminator

The "learning" in deep learning



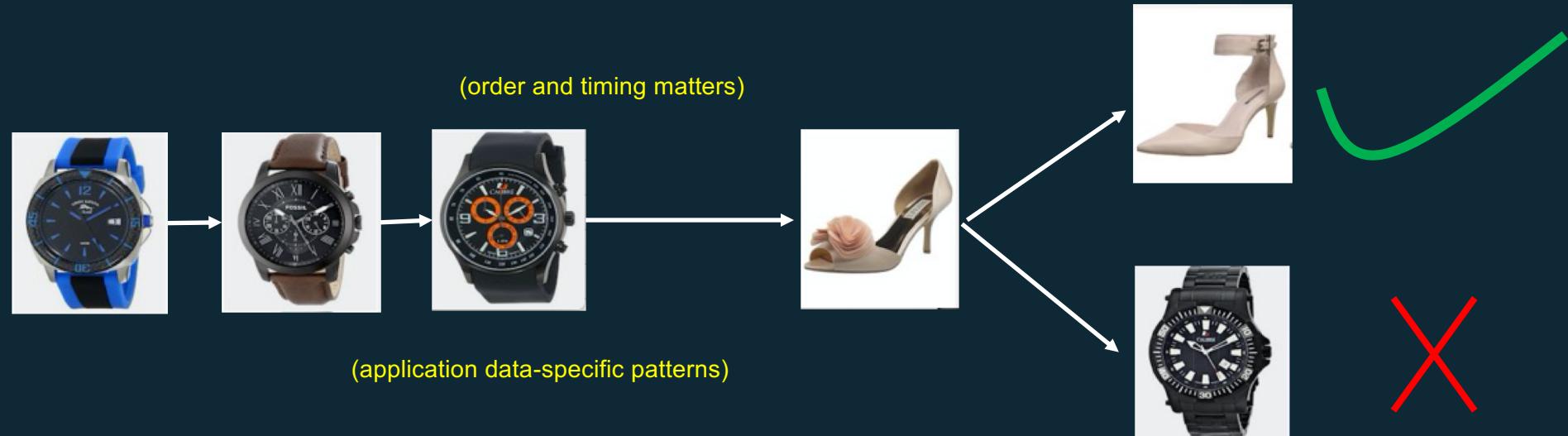
Deep learning is deep.....

Let's go back to recommendation

Modeling for Personalization

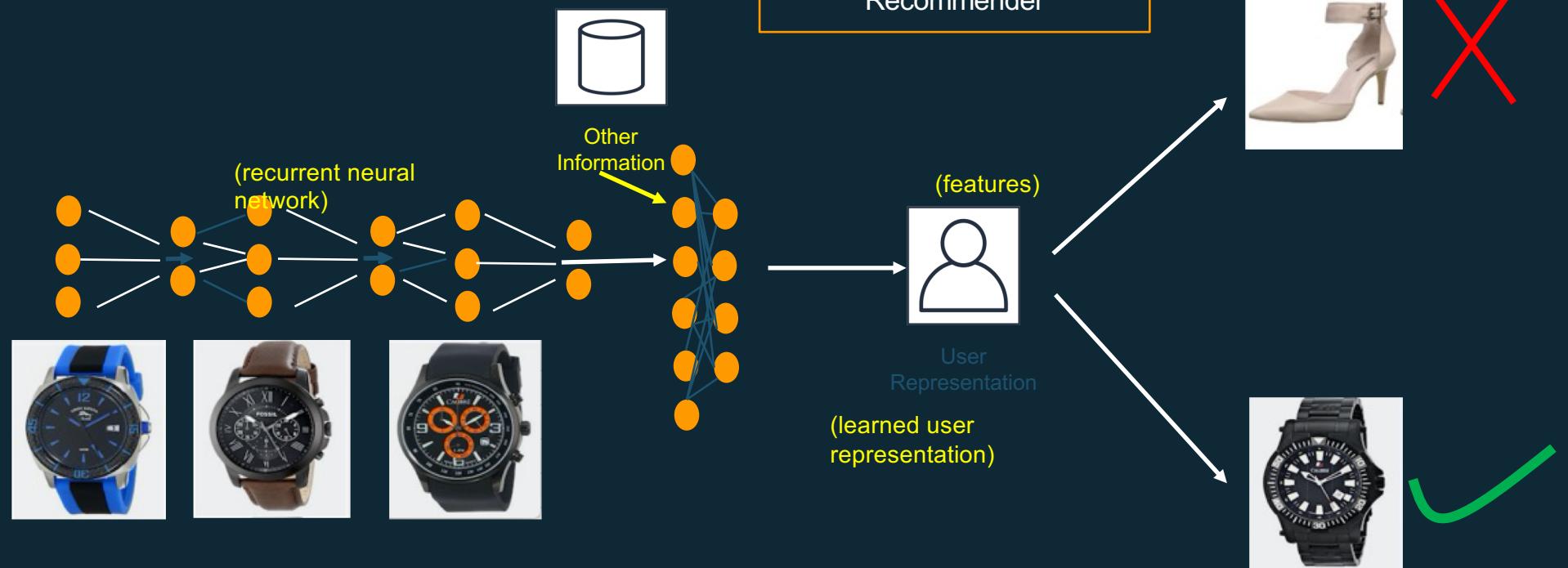
Fundamental insight

The evolution of historical interest and dis-interest is a good indicator of future preferences

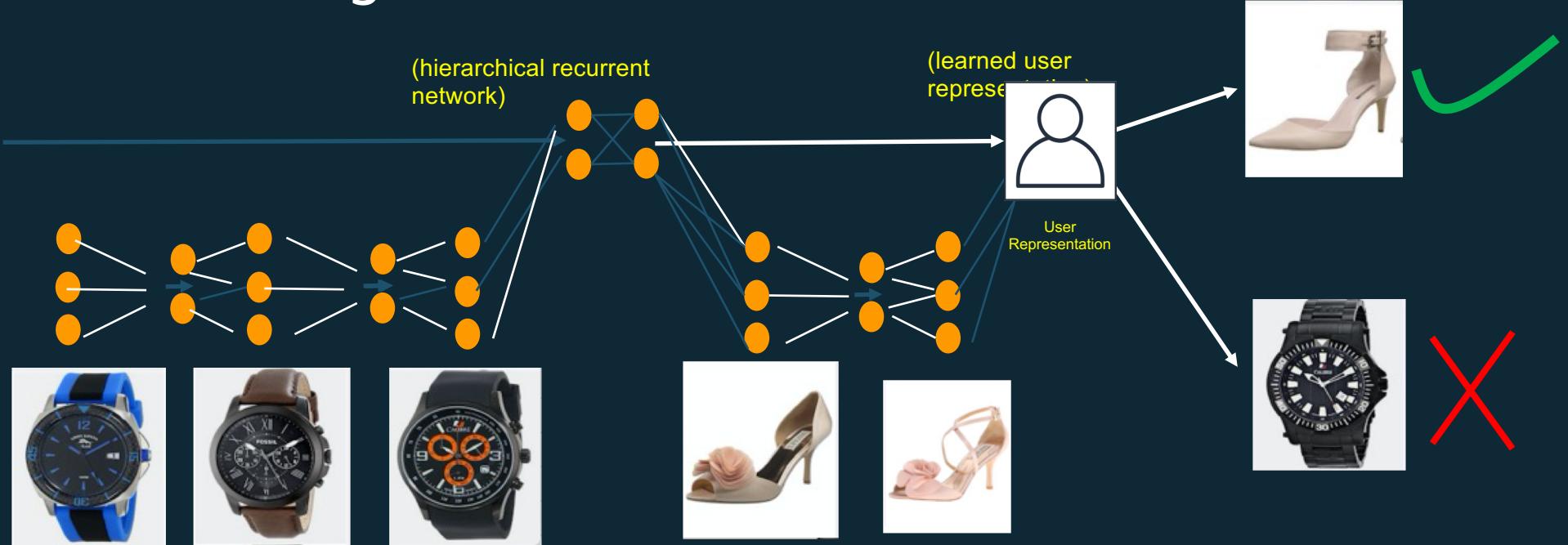


Sequential modeling

Available in Amazon
Personalize as Recurrent
Recommender



Modeling Sessions



Available in Amazon Personalize as
Hierarchical Recurrent Neural Network (HRNN)

State of the Art Performance

- Rolling Average
- PMF [2008]
- DeepRec [2017]
- T-SVD [2009]
- RRN [2017]
- HRNN



Ratings RMSE on Netflix

98 MM interactions, 500k users, 18k items

- Rolling Average



Ratings RMSE on MovieLens

20 MM interactions, 173k users, 131k items

Amazon Personalize

Improve customer experiences with personalization and recommendations



Deliver high quality
recommendations



Real-time



Easy to Use



Works with almost any
product or content

KEY FEATURES

Responsive to
changes in intent

Automated
machine learning

Deep learning
enabled algorithms

Bring existing algorithms
from Amazon SageMaker

HRNN is there!

Amazon Personalize 

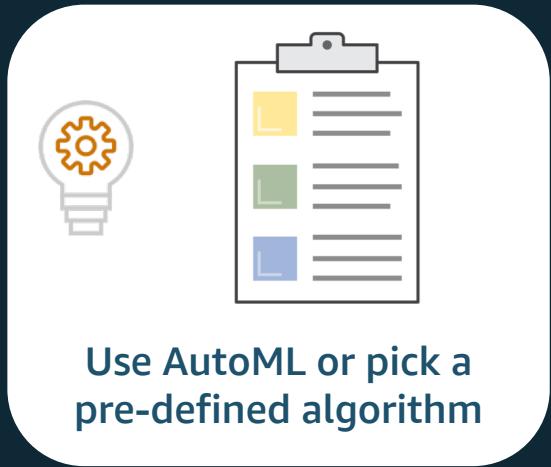
Developer Guide

Documentation - This Guide 

Search 

- What Is Amazon Personalize?
- How It Works
- Setting Up Amazon Personalize
- Getting Started
- Preparing and Importing Data
- Recording Events
- Creating a Solution
 - Using Predefined Recipes
 - DeepFM
 - FFNN
 - HRNN
 - Popularity-Count
 - Personalized-Ranking
 - SIMS
 - Bring Your Own Recipe

Use the Console/API to train and experiment with models



Use AutoML or pick a pre-defined algorithm

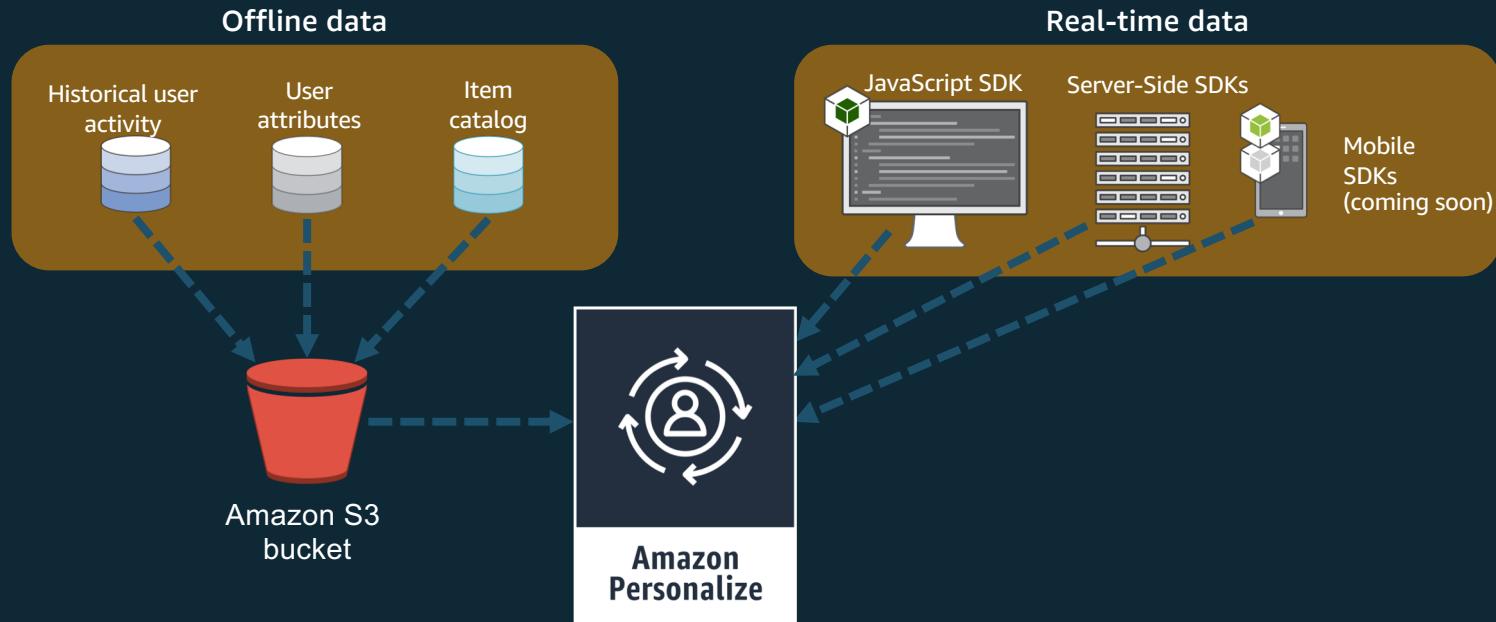
- Choose a pre-existing algorithm (packaged as Personalize Recipes) or use AutoML and Personalize will pick the right recipe for you
- You can train custom deep learning models on your data and compare accuracy metrics with 2 API calls

Deploy the best models (solutions) by launching a campaign



- Launching a campaign will deploy all the infrastructure needed to create a personalize endpoint
- You can use a simple API to `getRecommendations()` or `getPersonalizedResults()`

Real-time data can be consumed by Amazon Personalize



Amazon Personalize Customers



Hands-on Time!

Products Recommendation with Personalize

<https://bit.ly/2YIZUje>

Good that we have Personalize

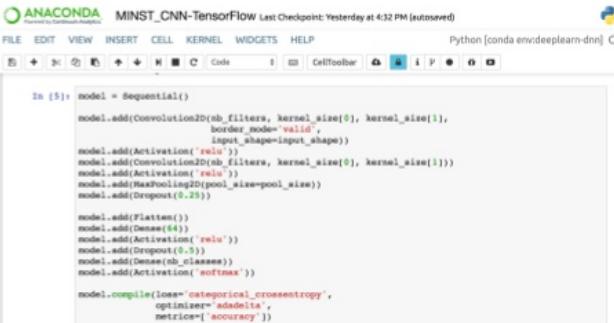
What about other use cases?

Amazon SageMaker
- Available in Singapore -

Amazon SageMaker: Jupyter Notebook

DEEP LEARNING IN THE JUPYTER NOTEBOOK

Defines a simple model in Keras to recognize handwritten digits



```
In [5]: model = Sequential()

model.add(Convolution2D(nb_filters, kernel_size(0), kernel_size(1),
                      border_mode='valid',
                      input_shape=input_shape))
model.add(Activation('relu'))
model.add(Convolution2D(nb_filters, kernel_size(0), kernel_size(1)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=pool_size))
model.add(Dropout(0.25))

model.add(Flatten())
model.add(Dense(48))
model.add(Activation('relu'))
model.add(Dropout(0.5))
model.add(Dense(nb_classes))
model.add(Activation('softmax'))

model.compile(loss='categorical_crossentropy',
              optimizers='adadelta',
              metrics=['accuracy'])
```



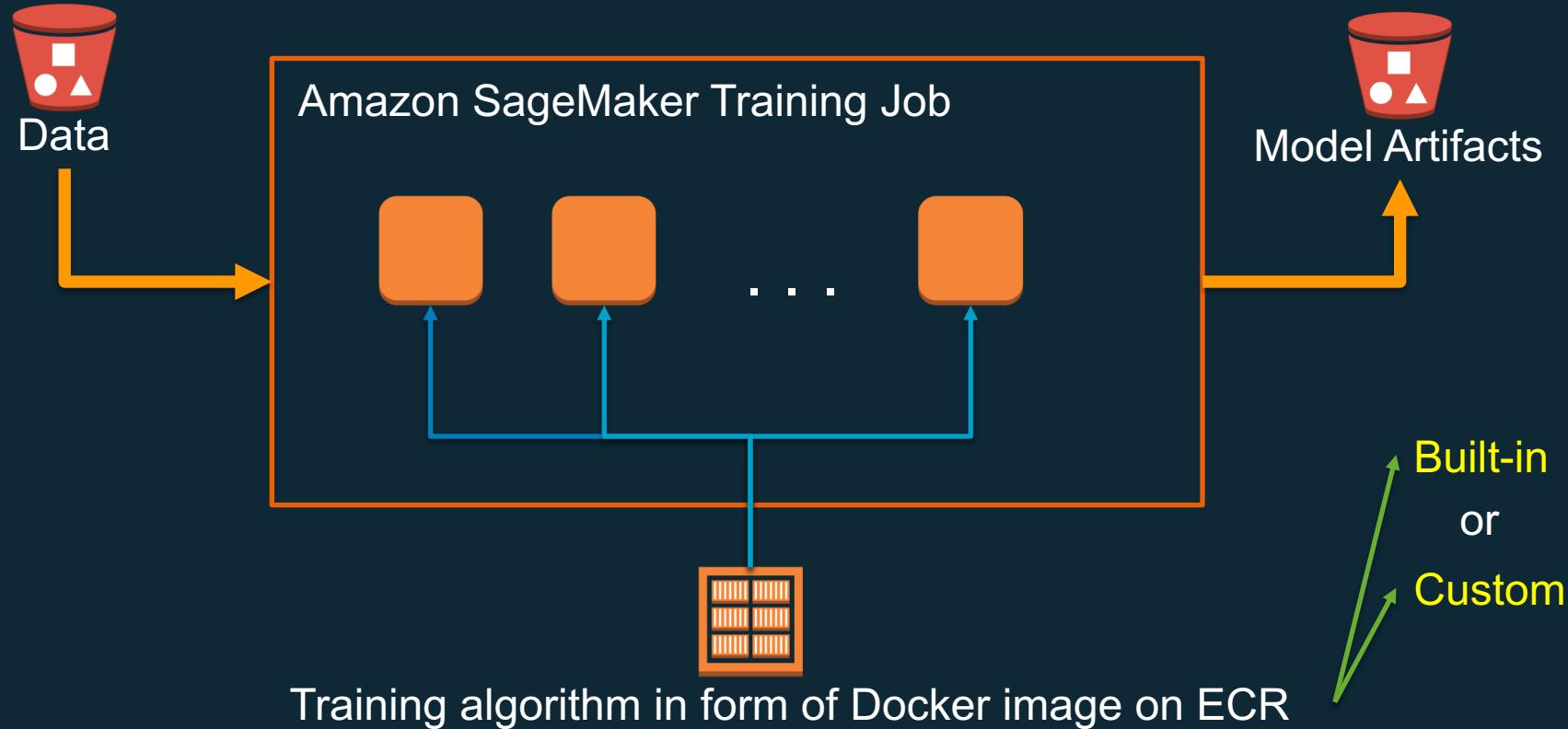
#OpenDataScienceMeans

#AnacondaCON

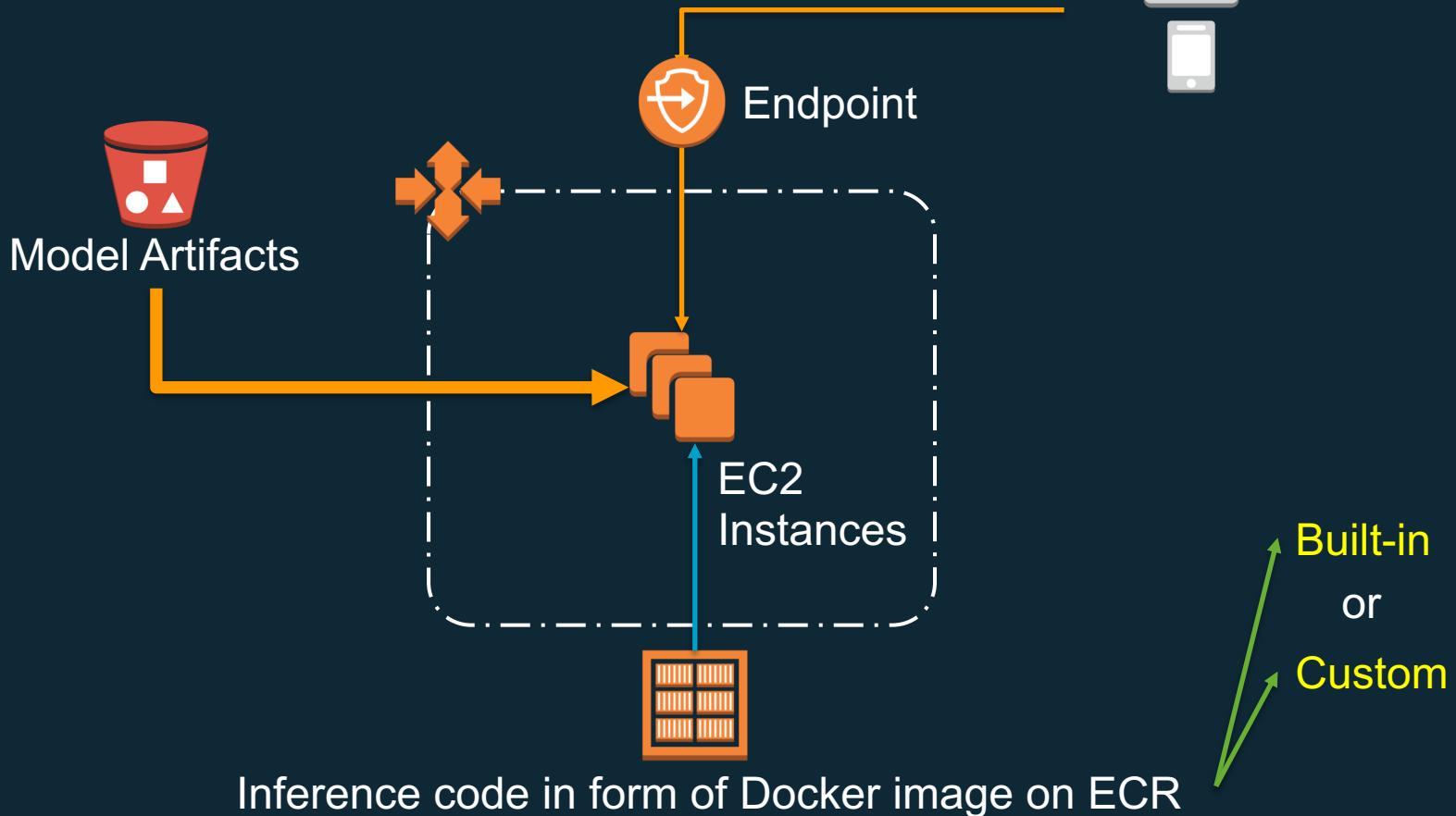


Amazon SageMaker
Training Job

Amazon SageMaker: Training Job



Amazon SageMaker: Hosting

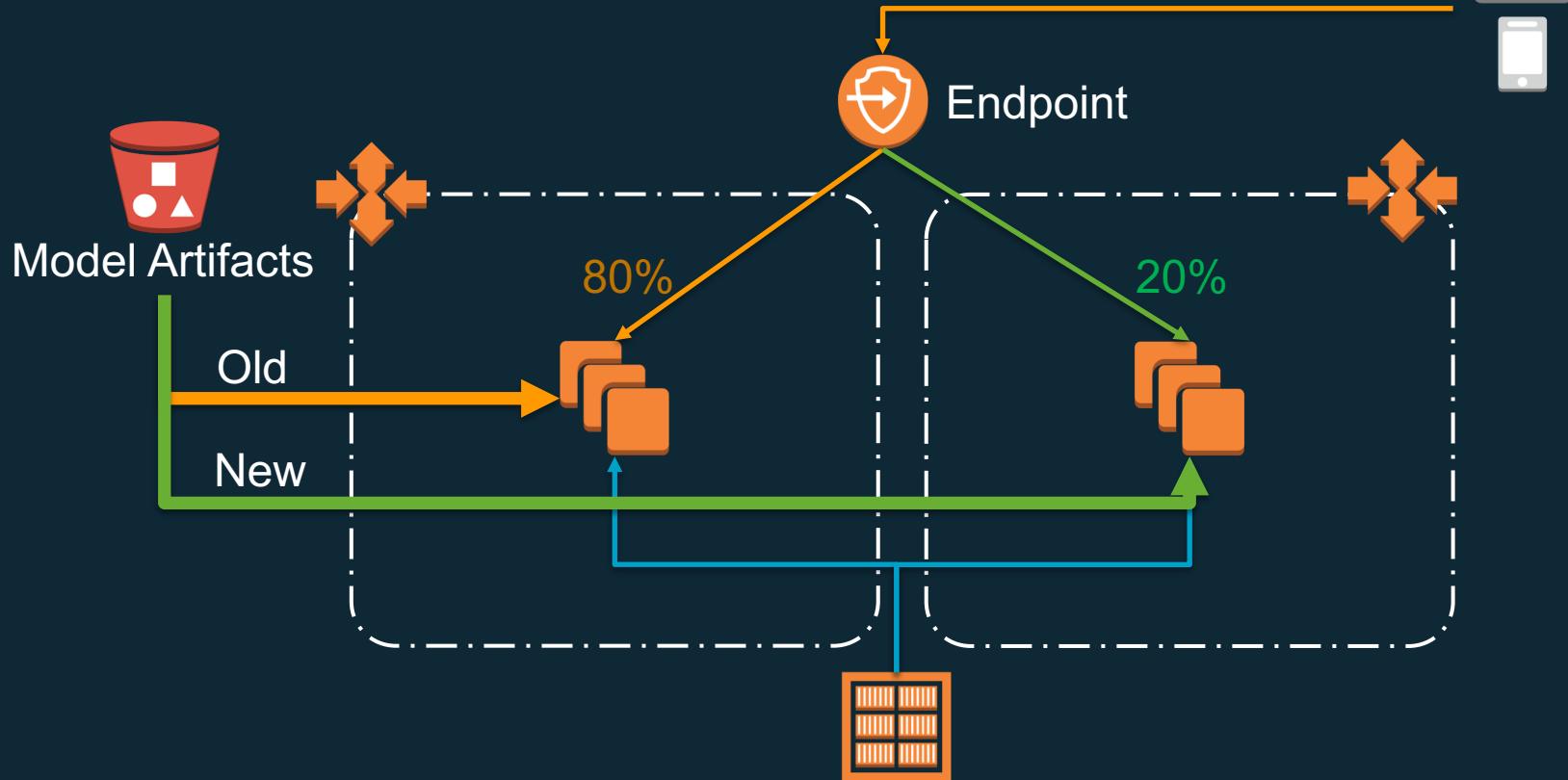


Amazon SageMaker: Hosting Auto Scaling

Scaling policy [Learn more](#) 

Policy name	SageMakerEndpointInvocationScalingPolicy
Target metric	Target value
SageMakerVariantInvocationsPerInstance	<input type="text" value="100"/>
Scale in cool down (seconds)	Scale out cool down (seconds)
<input type="text" value="300"/>	<input type="text" value="300"/>
<input type="checkbox"/> Disable scale in	

Amazon SageMaker: Hosting Model Variant



Inference code in form of Docker image on ECR

Hands-on Time!

Image Classification with SageMaker

<https://github.com/yudho/sagemaker-image-classification-with-own-data-source>

Thank You