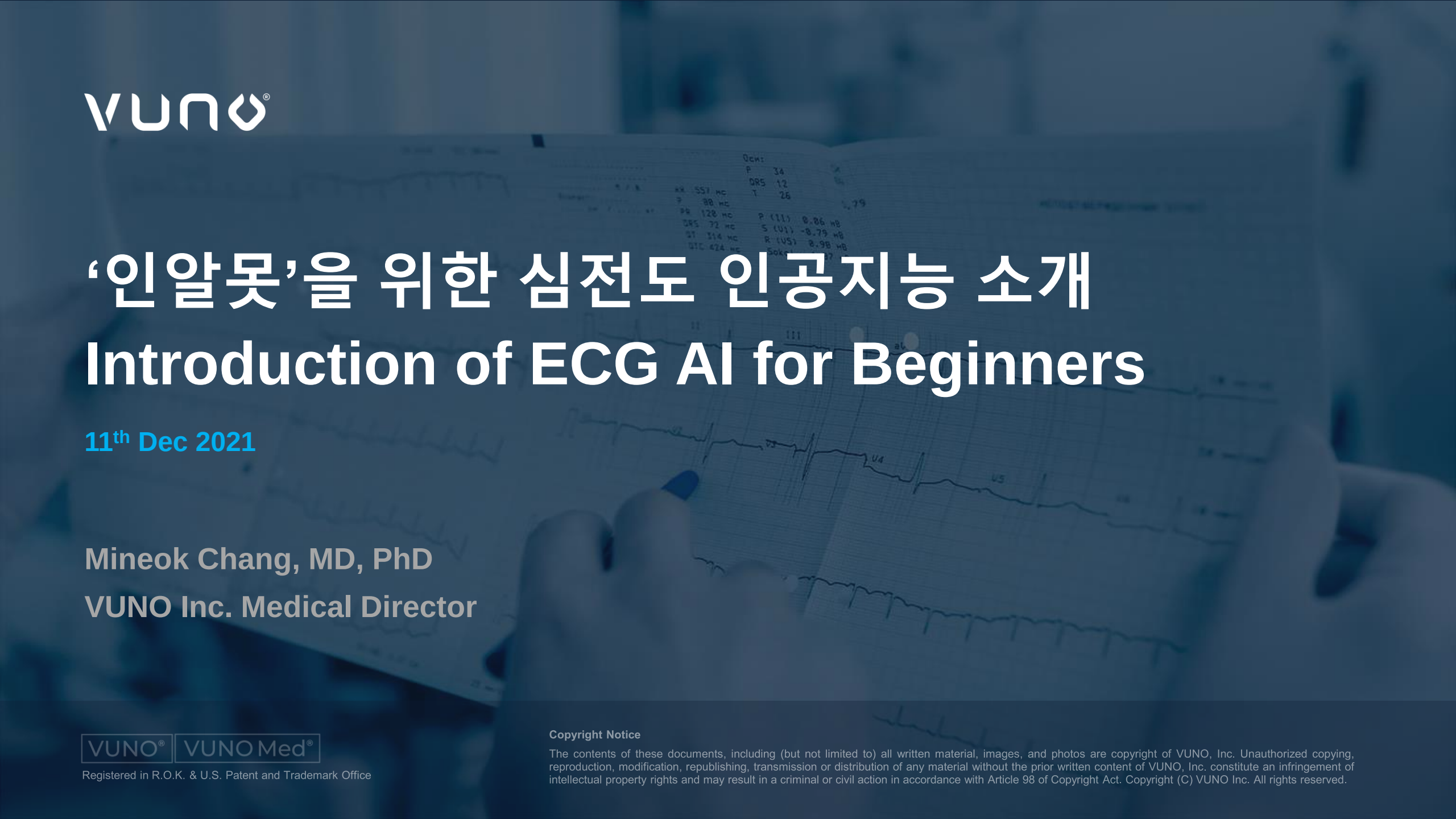


‘인알못’을 위한 심전도 인공지능 소개
Introduction of ECG AI for Beginners

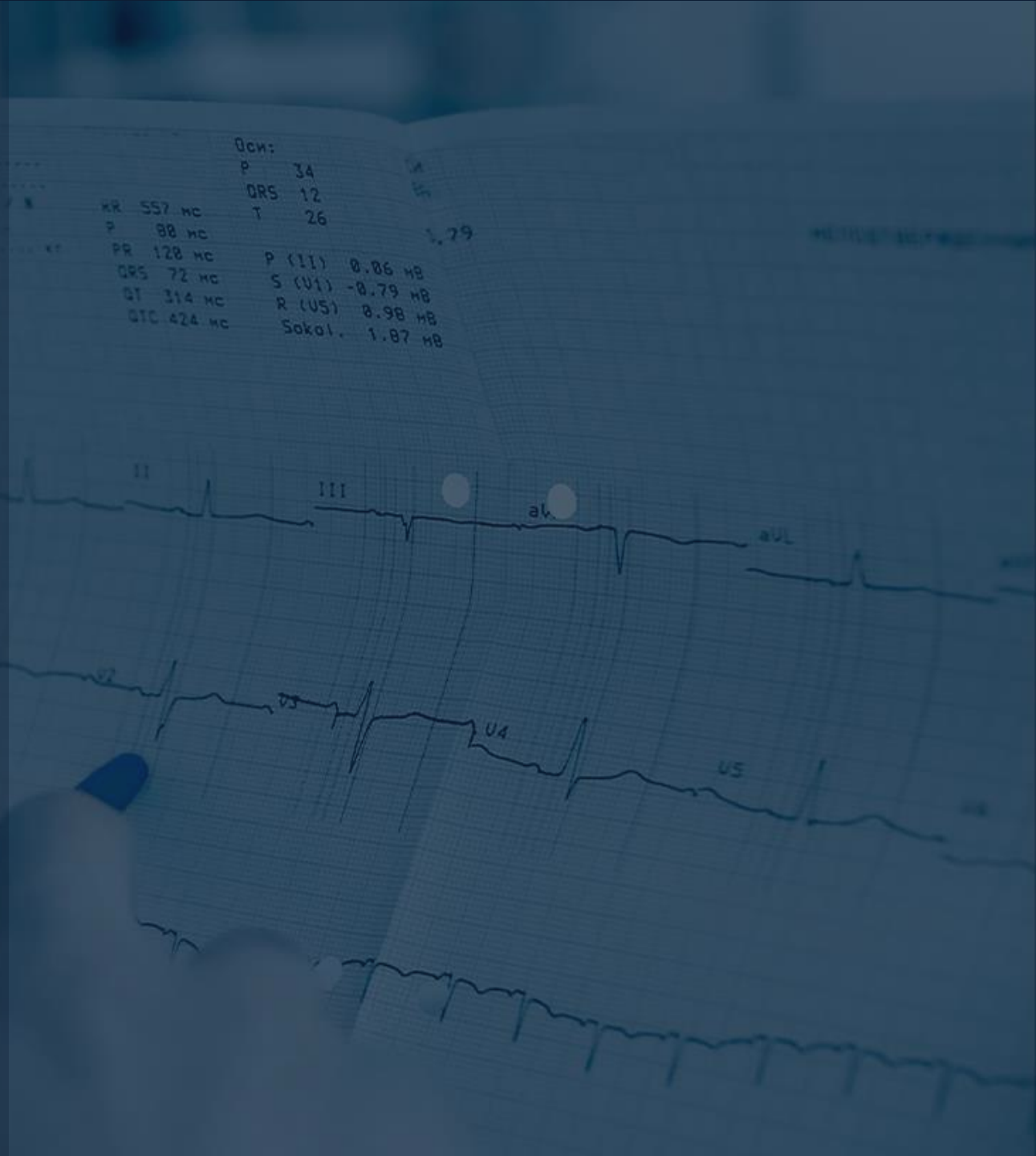
11th Dec 2021

Mineok Chang, MD, PhD

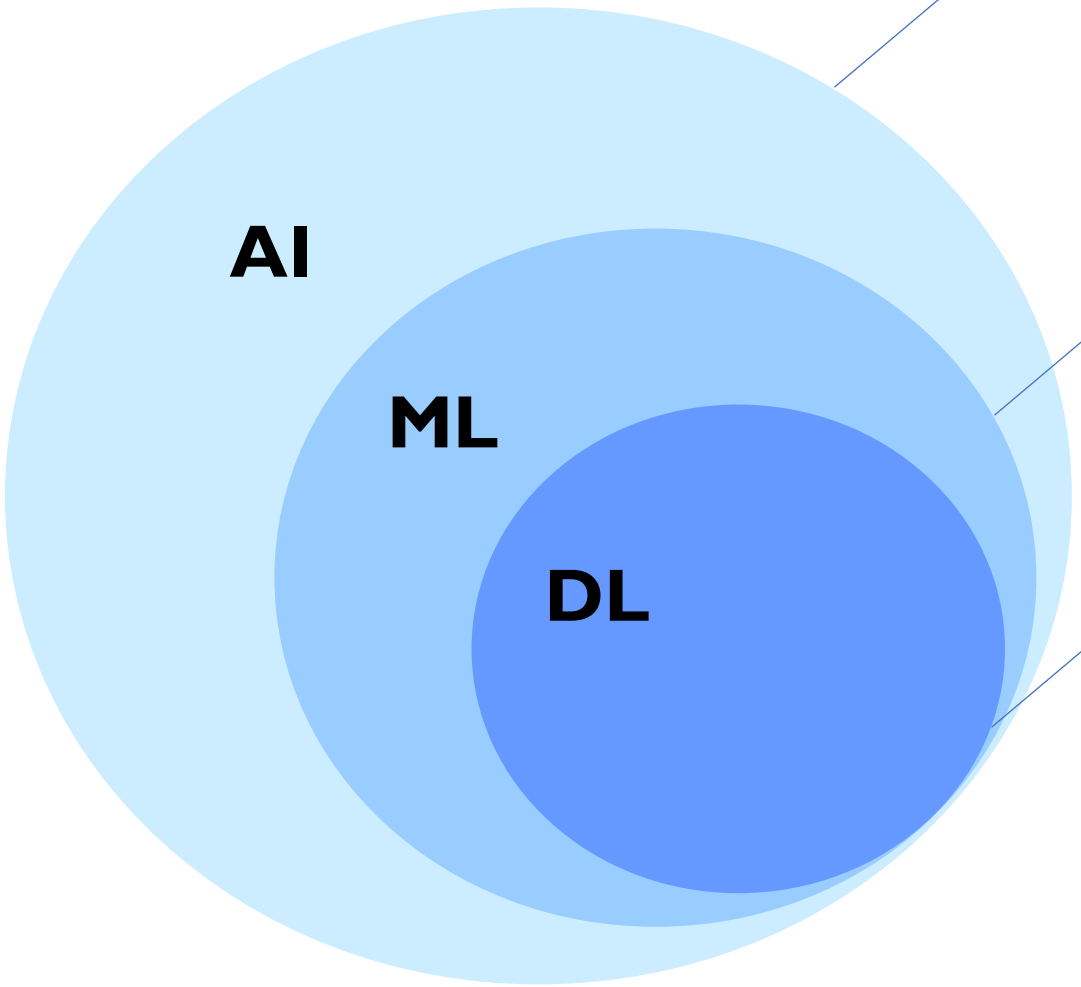
VUNO Inc. Medical Director

Contents

What is AI ?



AI ? ML ? DL ?



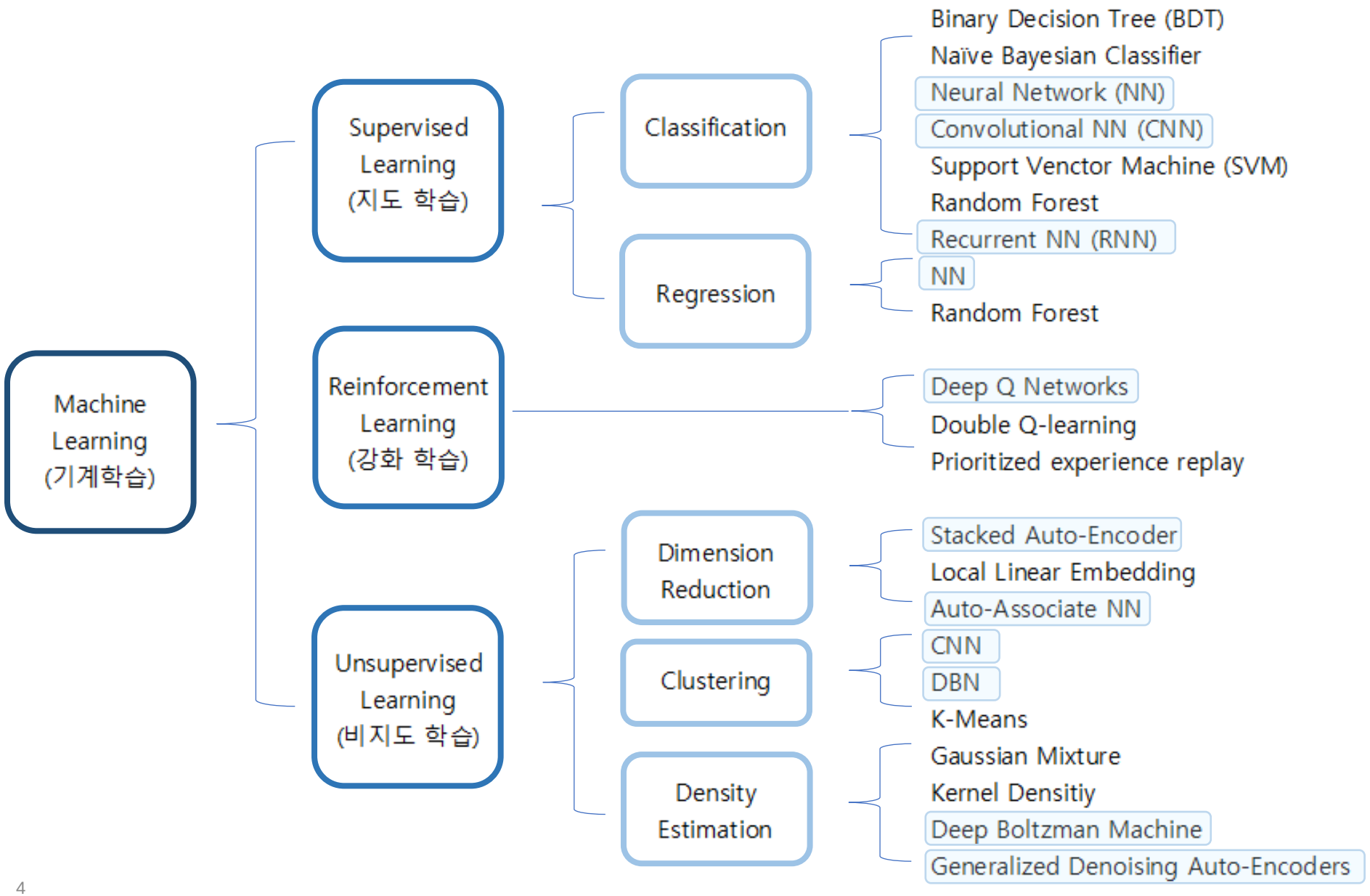
인공지능(Artificial Intelligence)
사람이 해야 할 일을 기계가 대신할 수 있는
모든 자동화에 해당

머신러닝(Machine Learning)
명시적으로 규칙을 프로그래밍하지 않고
데이터로부터 의사결정을 위한 패턴을
기계가 스스로 학습

딥러닝(Deep Learning)
머신러닝의 알고리즘의 하나로,
neural network 형태로 데이터 특징 추출 및
판단까지 기계가 수행

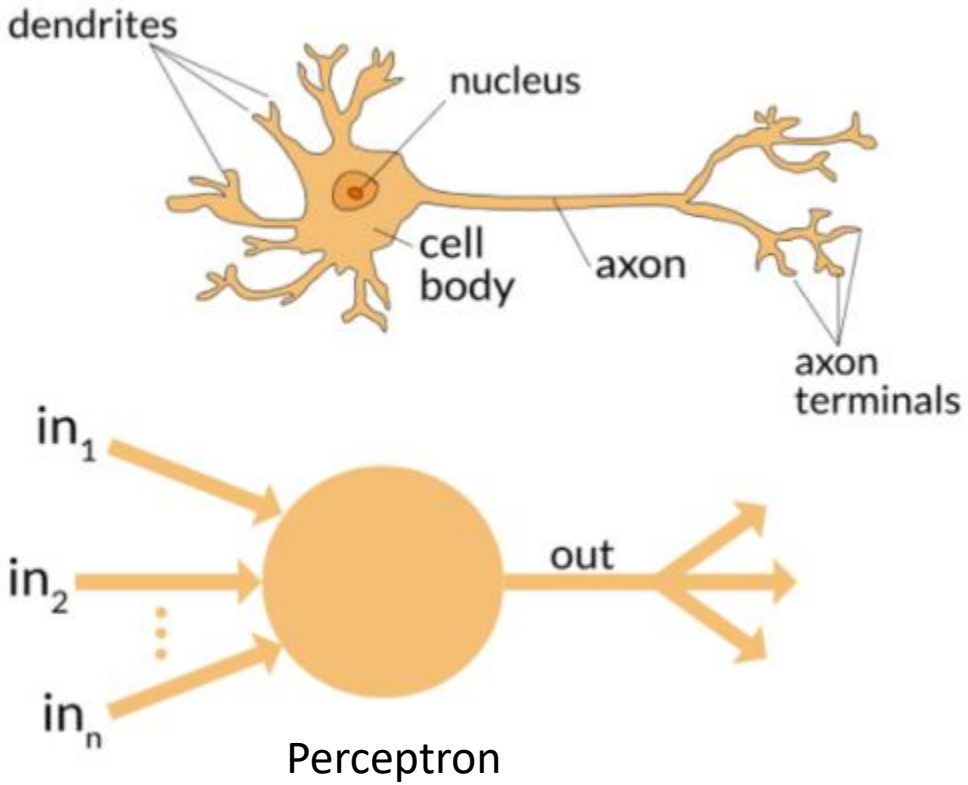
VUNO

Types of Machine Learning

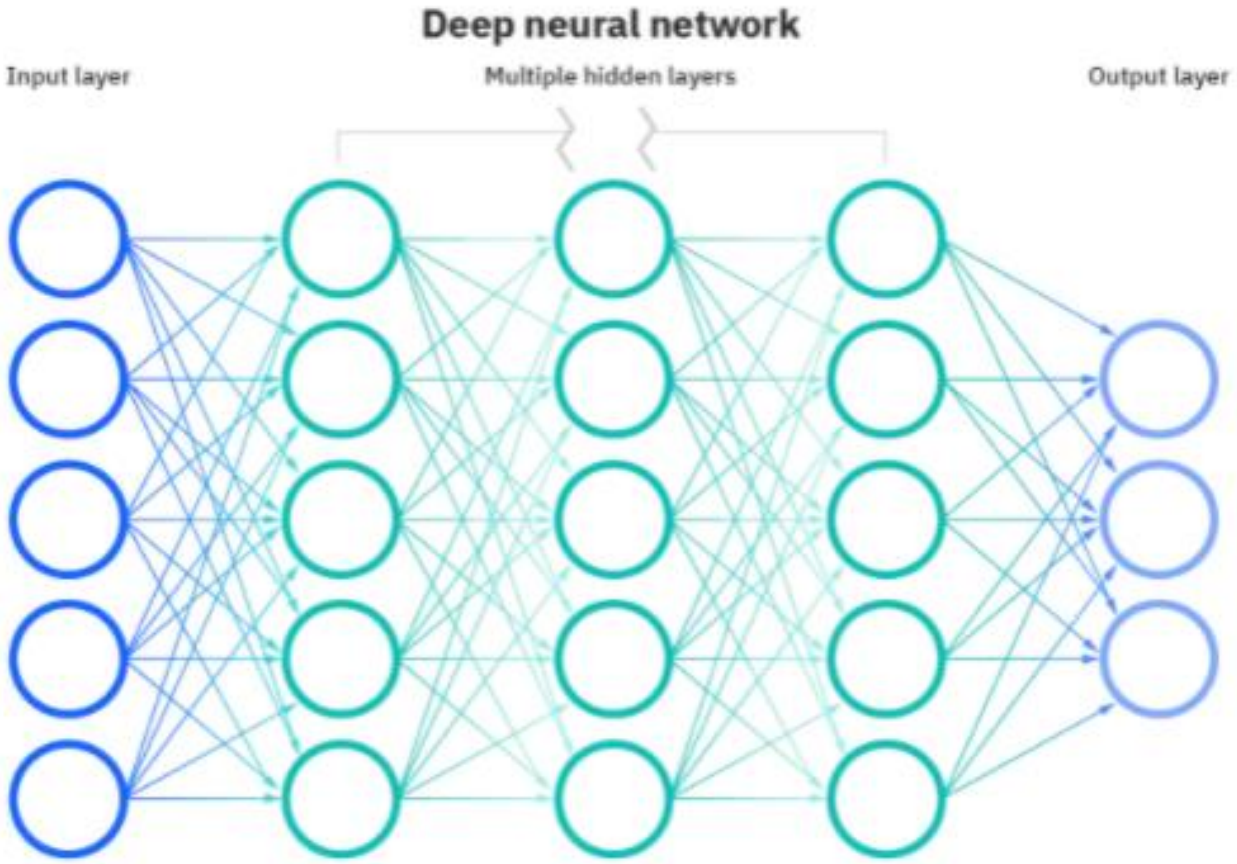


Deep Learning

Artificial Neural Network (ANN) ≒ Neural Network (NN)



Perceptron



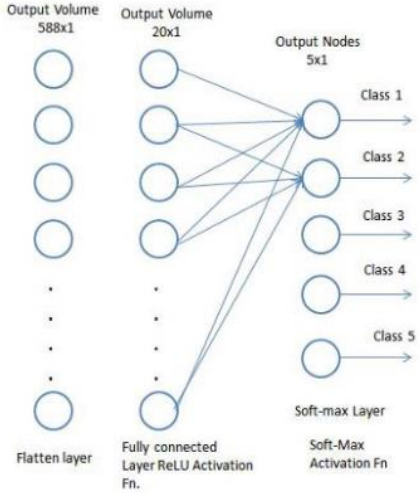
By Geoffrey Hinton

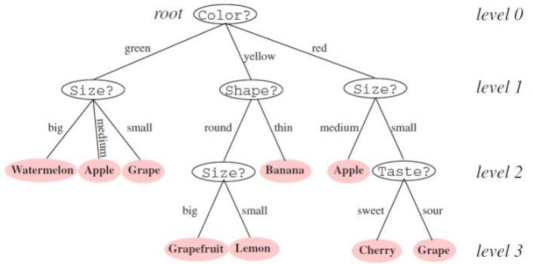
Designing a Learning System = Modeling

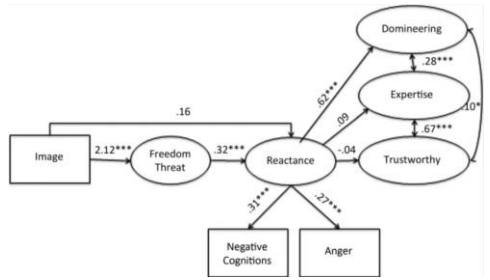
$$y = ax + b$$

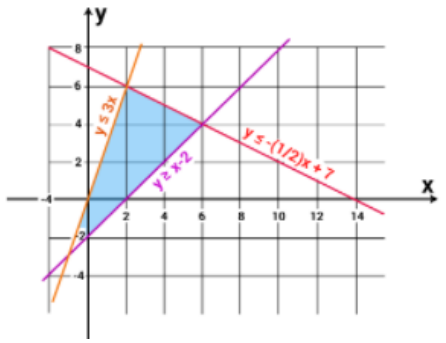
“ Training – Validation (Development) – Test ”

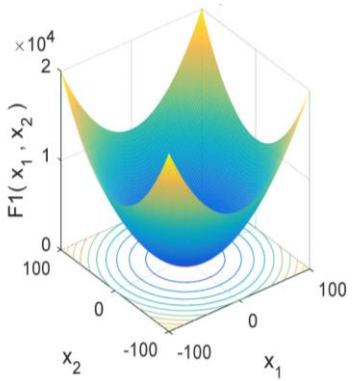
1. Problem Description
2. Choosing the Data (Parameter)
3. Choosing the Target Function
4. Choosing the Learning Architecture
5. Choosing the Learning Algorithm
6. Final Design (Hyperparameter)



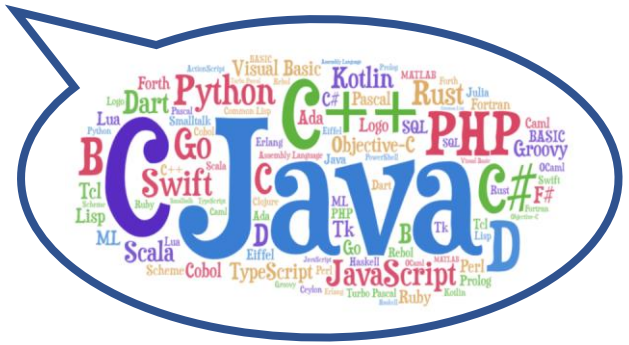






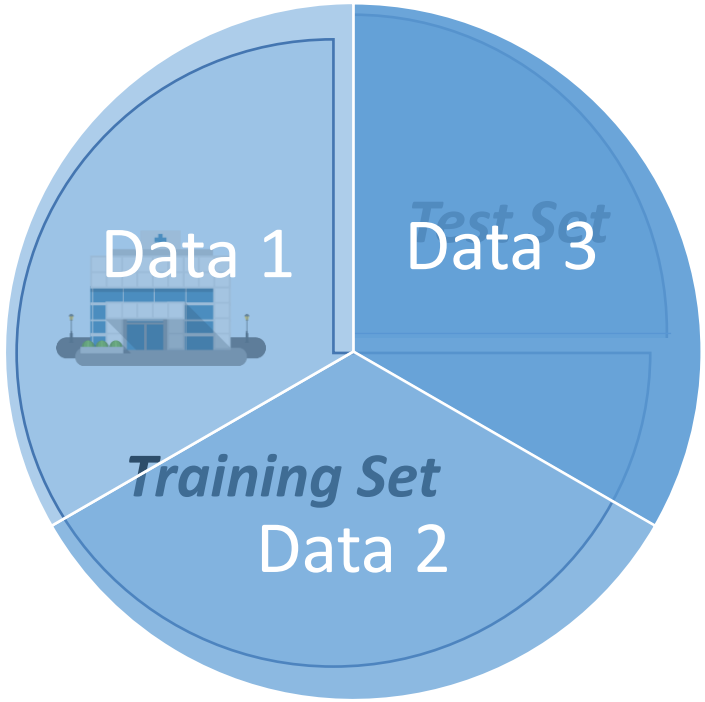


$$D_G^*(\boldsymbol{x}) = \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}$$

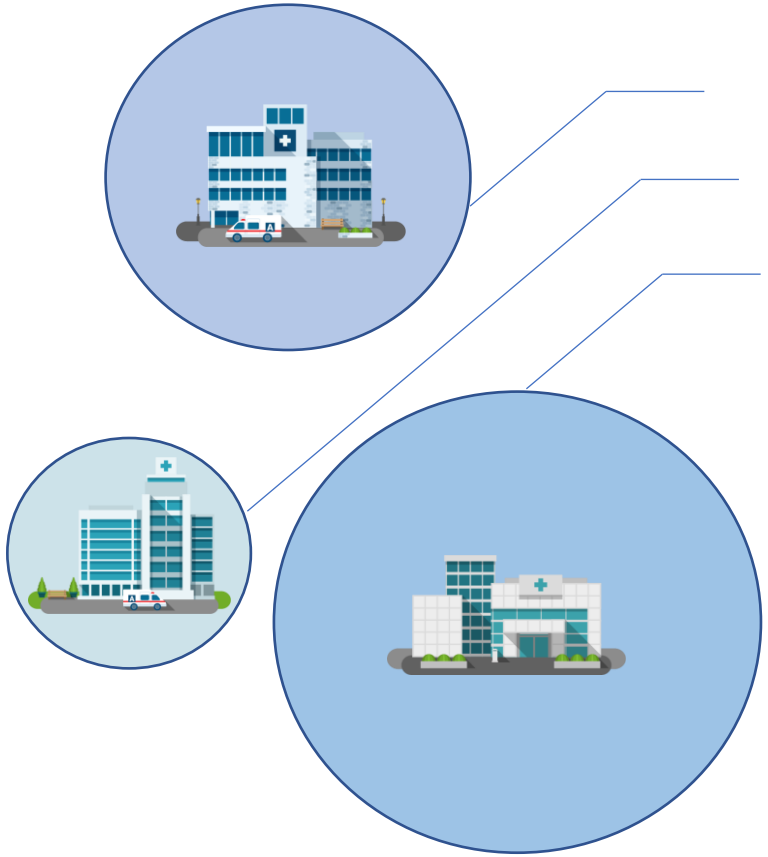


ML in In Medicine

“ Training – Validation (Development) – Test “



Internal Validation

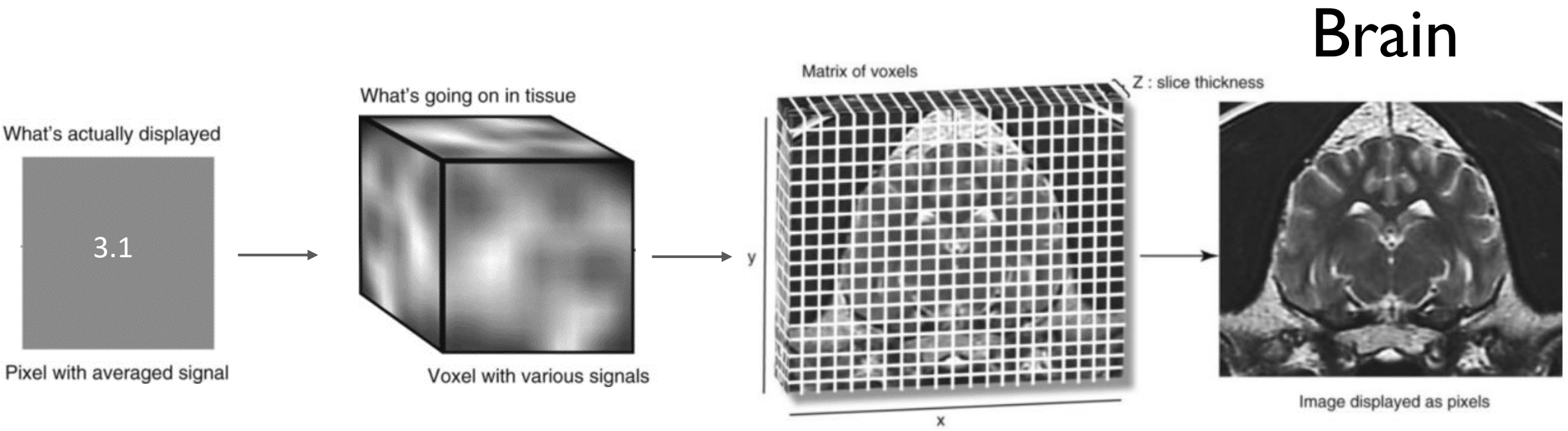


External Validation

Single Center

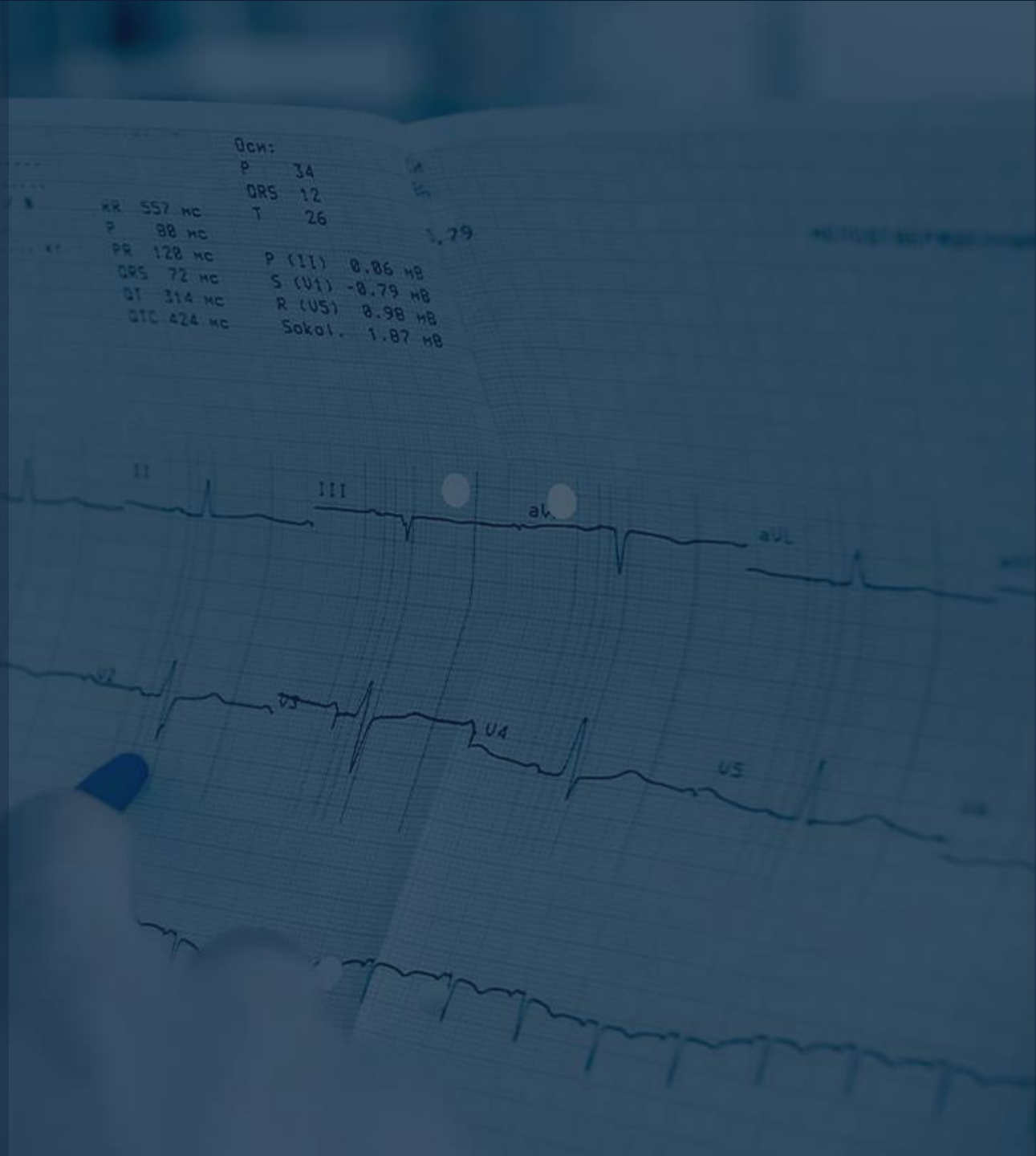
Multi Center

Medical Image on PACS

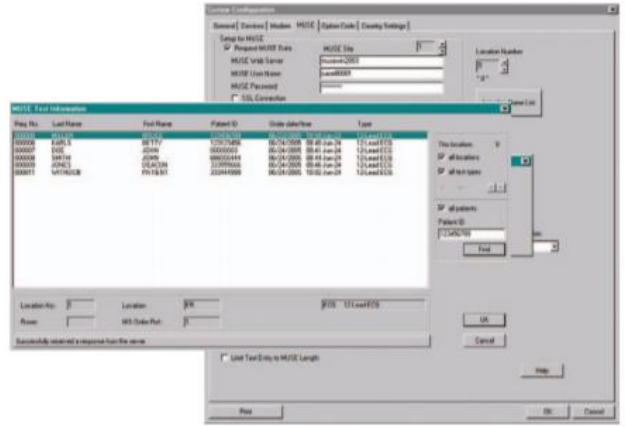


Contents

ECG data & analysis

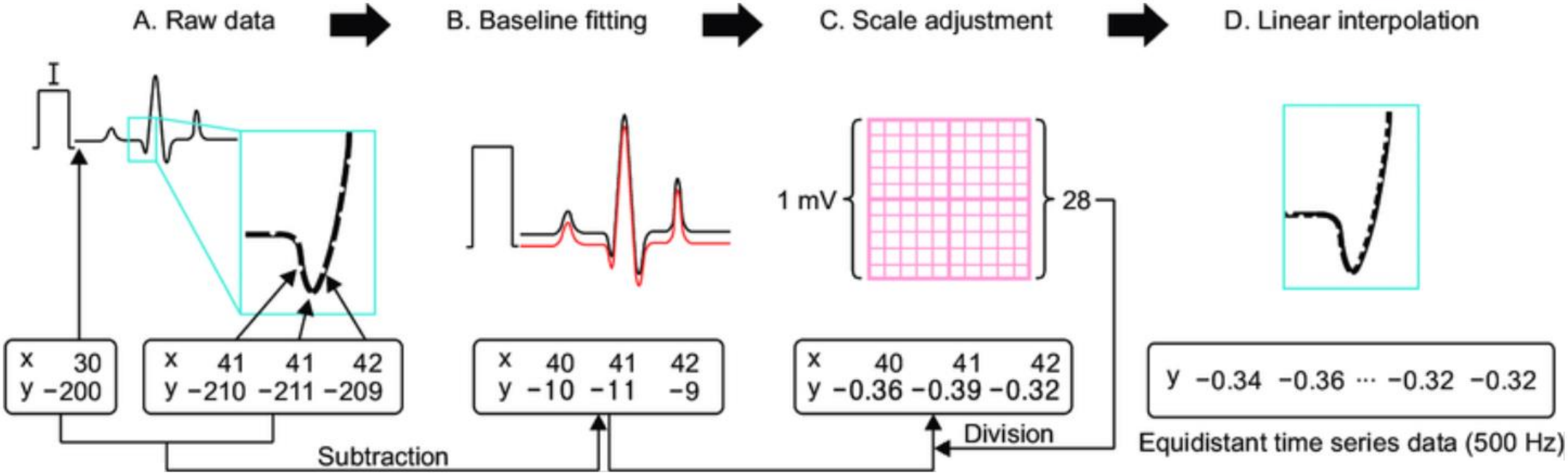


ECG is already Digitalized



What you see is PDF or JPEG on PACS or EMR

MUSE stores ECG raw data (XML, HL7)



XML data

```
<RestingECG>
  <PatientDemographics>
    <PatientID>227764</PatientID>
    <PatientAge>72</PatientAge>
    <AgeUnits>YEARS</AgeUnits>
    <Gender>FEMALE</Gender>
    <Race>ASIAN</Race>
    </PatientDemographics>
  <TestDemographics>
    <DataType>RESTING</DataType>
    <Status>UNCONFIRMED</Status>
    <AcquisitionTime>11:33:57</AcquisitionTime>
    <AcquisitionDate>06-19-2019</AcquisitionDate>
    </TestDemographics>
  <LeadData>
    <LeadByteCountTotal>1200</LeadByteCountTotal>
    <LeadTimeOffset>0</LeadTimeOffset>
    <LeadSampleCountTotal>600</LeadSampleCountTotal>
    <LeadAmplitudeUnitsPerBit>4.88</LeadAmplitudeUnitsPerBit>
    <LeadAmplitudeUnits>MICROVOLTS</LeadAmplitudeUnits>
    <LeadHighLimit>32767</LeadHighLimit>
    <LeadLowLimit>-32768</LeadLowLimit>
    <LeadID>II</LeadID>
    <LeadOffsetFirstSample>0</LeadOffsetFirstSample>
    <FirstSampleBaseline>0</FirstSampleBaseline>
    <LeadSampleSize>2</LeadSampleSize>
    <LeadOff>FALSE</LeadOff>
    <BaselineSway>FALSE</BaselineSway>
    <LeadDataCRC32>680213310</LeadDataCRC32>
    <WaveFormData>
```

```
CQAJAAkACAAIAAgACAAIAAgACAAIAAgACAAHAAcACAAIAAgACAAIAAYABgAHAAgACAAH
AAcACAAIAAcABgAGAAgACAAIAAgABgAGAAcABwAGAAYABgAHAAcABgAGAAYABgAGAAQA
BAAFAAUABAAEAAQABQAFAAUABQAFAAUABQAEAAQABAADAAMAAwADAAIAAgACAAMABAAE
AAQABAAEAAEAAQAEAAQAAwACAAIAAgACAAIAAgACAAIAAwADAAMAAwADAAIAAgADAAMA
AgACAAIAAQABAAMAAwADAAIAAgACAAIAAgACAAIAAgACAAIAAQABAAEAAQADAAMAAQAB
AAEAAgACAAIAAgACAAIABAAEAAMA//8BAAMAAwADAAMAAwACAAIAAgACAAIAAgABAAEA
BQAFAAMAAQABAAMABAAEAAEAAQACAAIABAAEAAMAAwADAAMAAwADAAMAAgACAAQABAAD
AAEAAQACAAIAAgACAAIAAgACAAIAAgACAAMAAwADAAIAAgACAAMAAwACAAIAAgABAAEA
AgACAAIAAgABAAAAAAAAAAAAAAAABAAMAAQABAAEAAQAAAAAAAgACAAIAAgAGAAoAEAAV
ACEALAA7AE4AYABxAIEAiwCPAJQAmwCWAHsAXgBEADMAHQAFAPn/9//3//f/+/8CAAUA
BwAHAAcABAD//wEAAQABAAAA/////wAAAAABAAEAAQABAAEAAQABAP//AQABAAIAAgAC
AAEAAQABAAEAAgACAAEAAQABAAAAAAAABAAIAAgACAAEAAQABAAIAAgADAAMAAwADAAMA
AwADAAIAAgACAAIAAgACAAIAAQABAAEABAAEAAIAAgACAAAAAAAABAAIAAgACAAIAAgAC
AAMAAwADAAMAAwAFAAUABQAEAAIAAgADAAQABQAFAAQABAAEAAQABAAEAAQABAAFAAUA
BQAFAAUABgAGAAYABQAFAAUACAAGAAQABgAGAAYABgAHAAkACQAHAAkACwAIAAgACgAK
AAoACgANAAsACwALAAoACgAMAAwADAANAA4ADgAOAA4AEAANAA0ADgAQABAADwAPABAA
EAAQABAAEAAQABAAEgATABMAEQARABAAEAAQABAAEAAQAA8ADwARABIAEgARABEAEQAS
ABIAEAAQABAAEAAQABAAEAAQABAAEAAOAA4ADgANAA0ADAALAAsACwALAAkACQAJAAkA
CQALAAwADAANAA4ADgAOAAsACwALAAsACwAMAAwACwALAAsACwALAAsACgAIAAgACQAJ
AAoACgAJAAcACgAKAAoACQAIAAgACAAIAAkACQAIAAgACQAJAAYABgAHAAkACQAIAAgA
CAAIAAgABgAGAAgACAAIAAgACAAIAAgACAAIAAYABgAIAAgACAAIAAgACQAJAAgACAAI
AAgACAAIAAgACAAIAAkACQAHAAcABwAIAAgACAAIAAgACAAIAAgACAAIAAgACAAIAAgA
CAAGAAYACAAIAAYABgAIAAgABwAHAAcABwAIAAgACAAHAAcABwAHAAgACAAKAAoACgAK
AAoACgALAAsACwALAAsACwALAAsADAAMAAwA
```

```
    </WaveFormData>
```

Inference

x

XML data

```
CQAJAAkACAAIAAgACAAIAAgACAAIAAgACAAHAAcACAAIAAgACAAIAAYABgAHAAgACAAH
AAcACAA[redacted]AAgACAAIAAgABgAGAAcABwAGAAYABgAHAAcABgAGAAYABgAGAAQA
BAAFAAUABAAEAAQABQAFAAUABQAFAAUABQAEAAQABAADAAMAAwADAAIAAgACAAMABAAE
AAQABAAEAAEAAQAEAAQAAwACAAIAAgACAAIAAgACAAIAAwADAAMAAwADAAIAAgADAAMA
AgACAAIAAQABAAMAAwADAAIAAgACAAIAAgACAAIAAgA[redacted]AEAAQADAAMAAQAB
AAEAAgACAAIAAgACAAIABAAEAAMA//8BAAMAAwADAAMAAwACAAIAAgACAAIAAgABAAEA
BQAFAAMAAQABAAMABAAEAAEAAQACAAIABAAEAAMAAwADAAMAAwADAAMAAgACAAQABAAD
AAEAAQACAAIAAgACAAIAAgACAAIAAgACAAMAAwADAAIAAgACAAMAAwACAAIAAgABAAEA
AgACAAIAAgABAAAAAAAAAAAAAAAABAAMAAQABAAEAAQAAAAAAAgACAAIAAgAGAAoAEAAV
ACEALAA7AE4AYABxAIEAiwCPAJQAmwCWAHsAXgBEADMAHQAFAPn/9//3//f/+/8CAAUA
BwAHAAcABAD//wEAAQABAAAA/////wAAAAABAAEAAQABAAEAAQABAP//AQABAAIAAgAC
AAEAAQABAAEAAgACAAEAAQABAAAAAAAABAAIAAgACAAEAAQABAAIAAgADAAMAAwADAAMA
AwADAAIAAgACAAIAAgACAAIAAQABAAEABAAEAAIAAgACAAAAAAABAAIAAgACAAIAAgAC
AAMAAwADAAMAAwAFAAUABQAEAAIAAgADAAQABQAFAAQABAAEAAQA[redacted]AFAAUA
BQAFAAUABgAGAAYABQAFAAUACAAGAAQABgAGAAYABgAHAAkACQAHAAkACwAIAAgACgAK
AAoACgANAAsACwALAAoACgAMAAwADAANAA4ADgAOAA4AEAANAA0ADgAQABAADwAPABAA
EAAQABAAEAAQABAA[redacted]ARABAAEAAQABAAEAAQAA8ADwARABIAEgARABEAEQAS
ABIAEAAQABAAEAAQABAAEAAQABAAEAAOAA4ADgANAA0ADAALAAsACwALAAsACwALAAkA
CQAJAAkACQALAAwADAANAA4ADgAOAAsACwALAAsACwAMAAwACwALAAsACwALAAsACgAI
AAgACQAJAAoACgAJAAcACgAKAAoACQAIAAgACAA[redacted]AgACQAJAAYABgAHAAkA
CQAIAAgACAAIAAgABgAGAAgACAAIAAgACAAIAAgACAAIAAYABgAIAAgACAAIAAgACQAJ
AAgACAAIAAgACAAIAAgACAAIAAkACQAHAAcABwAIAAgACAAIAAgACAAIAAgACAAIAAgA
CAAGAAYACAAIAAYABgAIAAgABwAHAAcABwAIAAgACAAHAAcABwAHAAgACAAKAAoACgAK
AAoACgALAAsACwALAAsACwALAAsADAAMAAwA
    </WaveFormData>
```

$y = ax + b$

Model (Code)

```
def model_fn (params):

    inputs = tf.keras.Input(name='waveform')

    x = layers.Conv2D(
            filters=64,
            kernel_size=3,
            strides=1,
            padding='SAME',
            use_bias=False,
            kernel_initializer='he_normal')(x)
    x = layers.BatchNormalization()(x)
    x = layers.Activation('relu')(x)

    x = block(64)(x)
    x = block(64)(x)

    x = block(128, upscale=True)(x)
    x = block(128)(x)
    x = block(128)(x)

    x = block(256, upscale=True)(x)
    x = block(256)(x)
    x = block(256)(x)

    x = layers.GlobalAveragePooling2D()(x)

    model = tf.keras.Model(inputs=inputs, outputs=outputs)
    model.summary()

    model_spec = {}
    model_spec['model'] = model
    model_spec['loss'] = get_loss(params)
    model_spec['optimizer'] = get_optimizer(params)

    return x, model_spec
```

y

Output

0 ~ 1

Probability

With cut-off

MODEL_OUTPUT
0.655296028
0.542461872
0.644664586
0.902734876
0.801081657
0.576097906
0.660251319

VUNO

Performance of Model

Confusion Matrix		
	Actual Positive	Actual Negative
Model Positive	TP	FP (type I error)
Model Negative	FN (type 2 error)	TN

Sensitivity, Recall (재현율)	$\frac{TP}{TP + FN}$
Specificity	$\frac{TN}{FP + TN}$
PPV, Precision (정밀도)	$\frac{TP}{TP + FP}$
NPV	$\frac{TN}{TN + FN}$
F1 (조화평균)	$2 * \frac{Precision * Recall}{Precision + Recall} = \frac{2TP}{2TP+FP+FN}$
Accuracy	$\frac{TP + TN}{TP + FP + FN + TN}$

Data Imbalance matters

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

$$\text{F1 score} = \frac{2TP}{2TP+FP+FN}$$

Ex. Glioblastoma prevalence : 0.003%

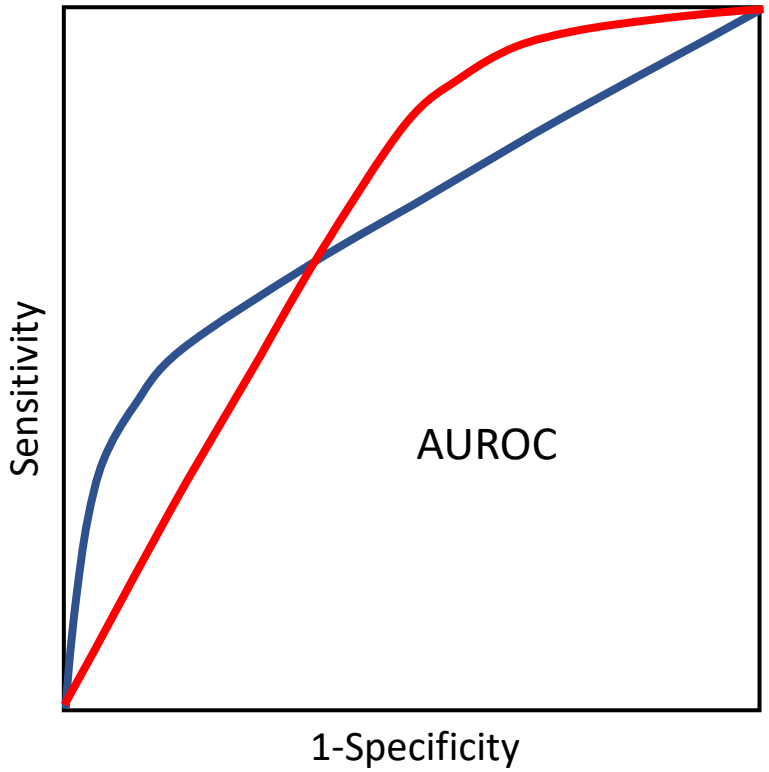
Confusion Matrix		
	Actual Positive	Actual Negative
Model Positive	0 (TP)	0(FP)
Model Negative	3(FN)	100,000(TN)

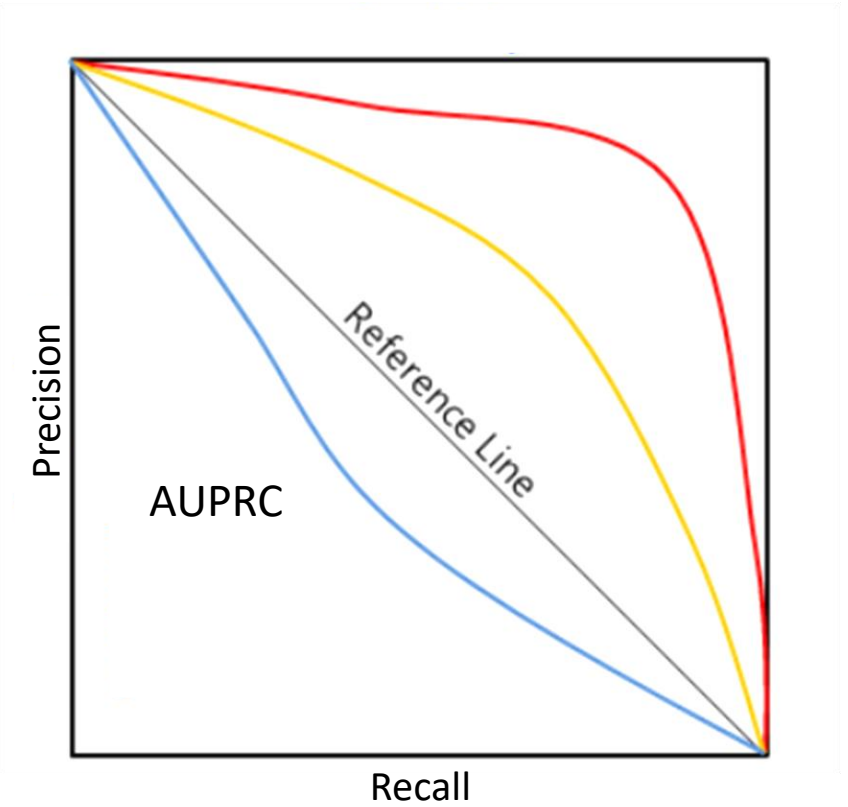
$$\text{Glioblastoma prediction Accuracy} = \frac{100{,}000}{100{,}003} = 99.99\%$$

$$\text{Glioblastoma F1 score} = \frac{0}{3} = 0\%$$

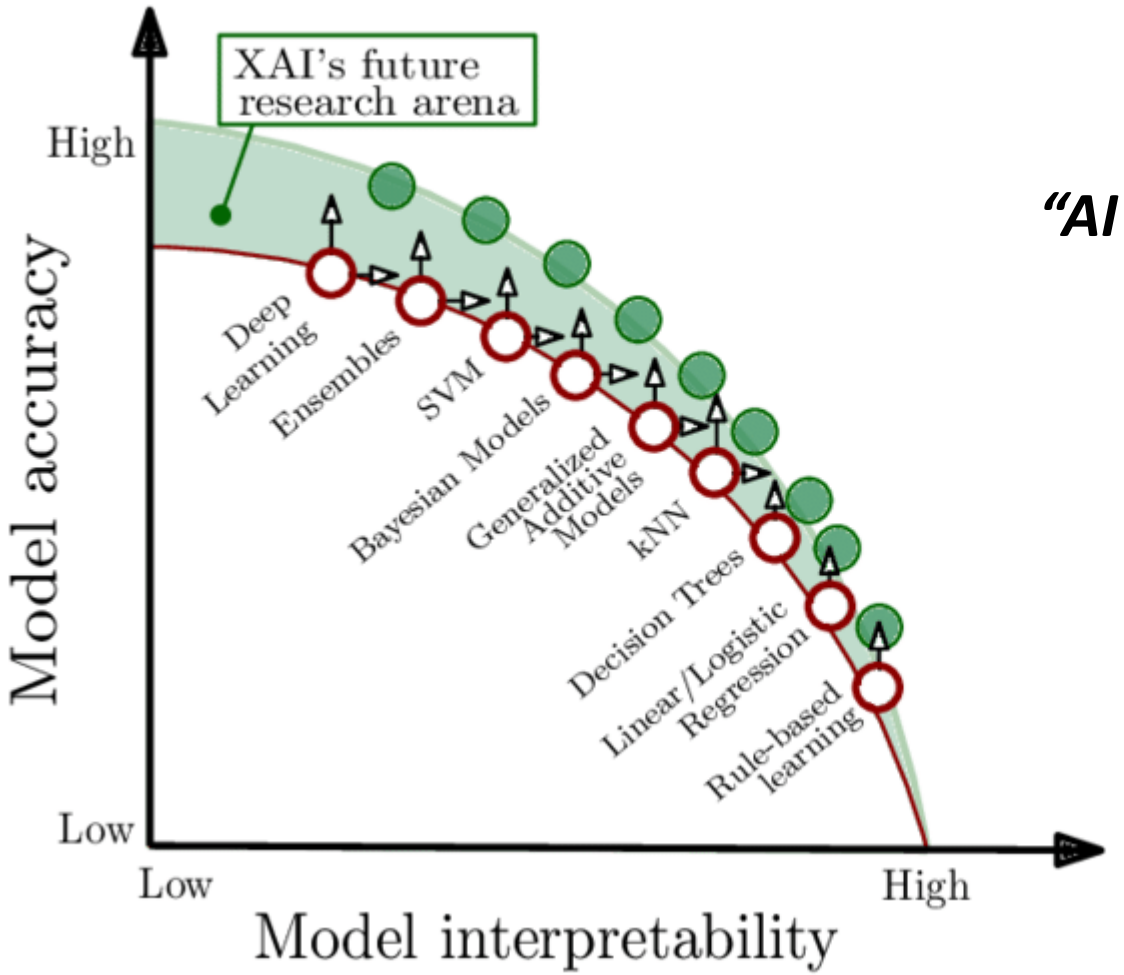
Recall (Sensitivity) vs Precision (PPV)

Trade-off exist !
ROC is not sufficient !

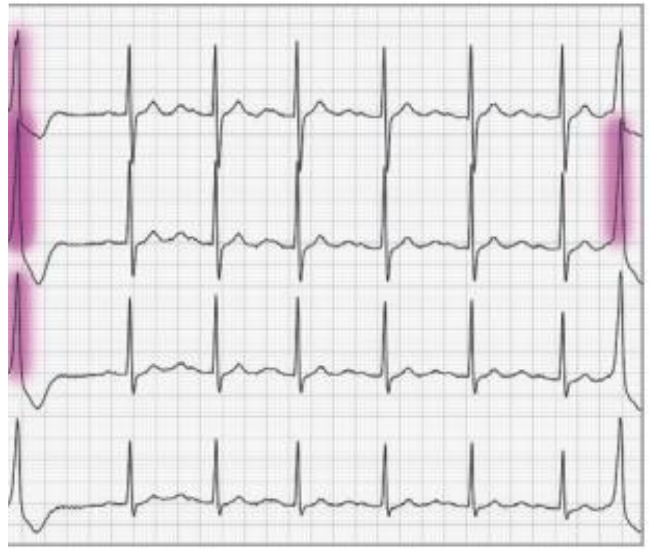




Trade-off between Accuracy & Interpretability

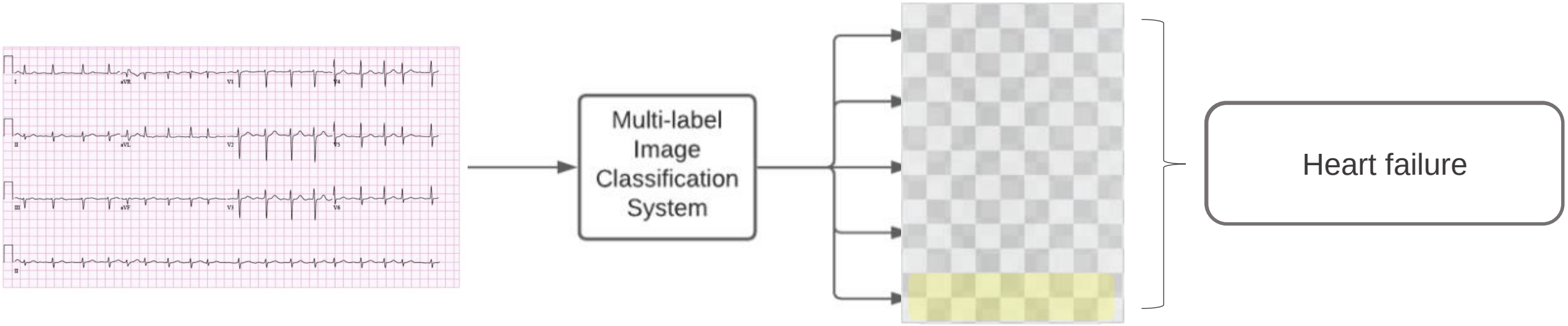


"AI is Black Box"



Explainable AI, is it possible?

What you see is, what you know !

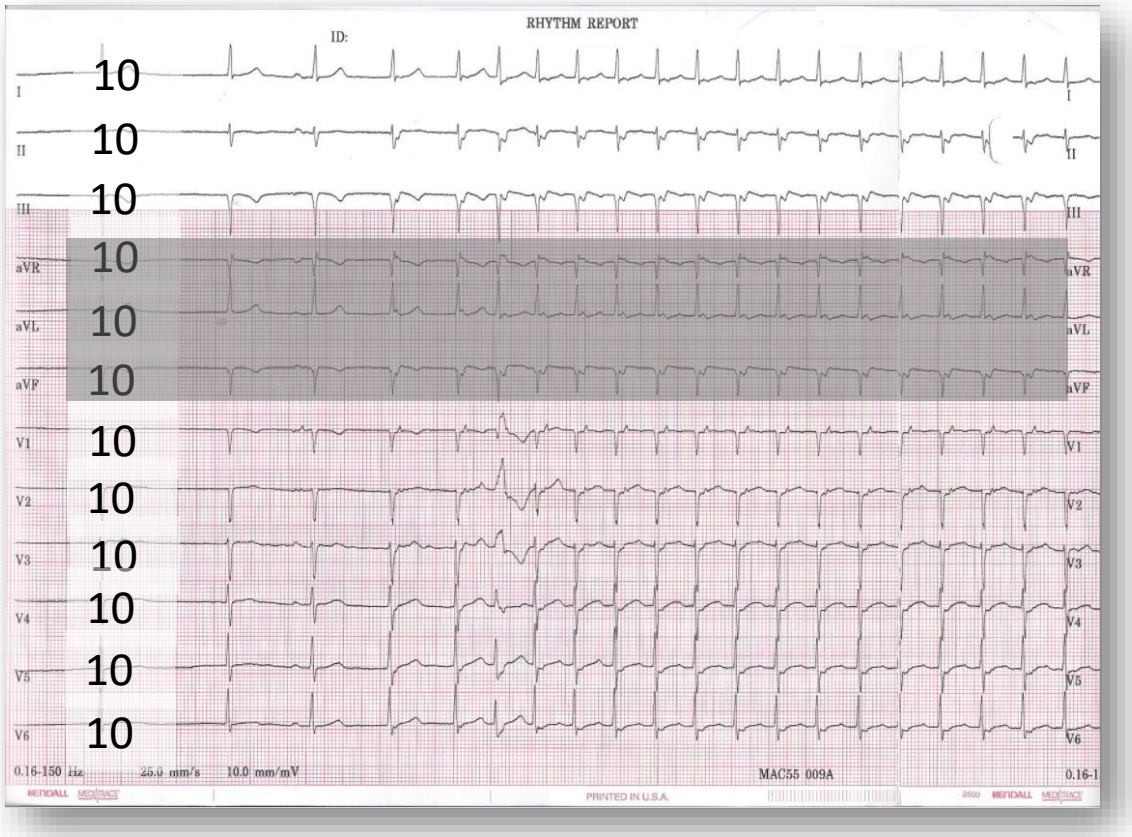


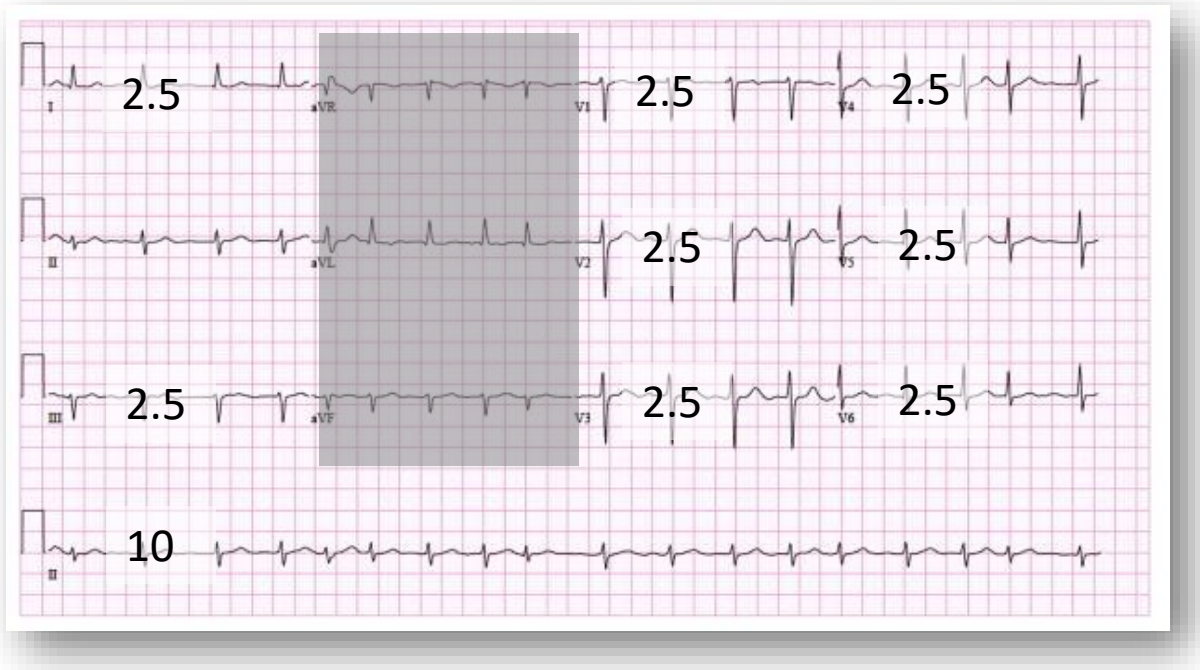
Contents

How ECG AI works?

Amount of Information

Computer vs. Human



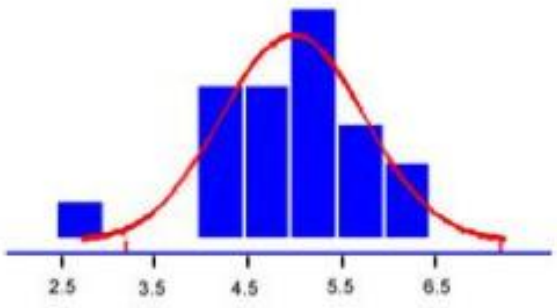


Amount of information is 12L x 10sec = 120
Considering augmented lead, total information would be **90**

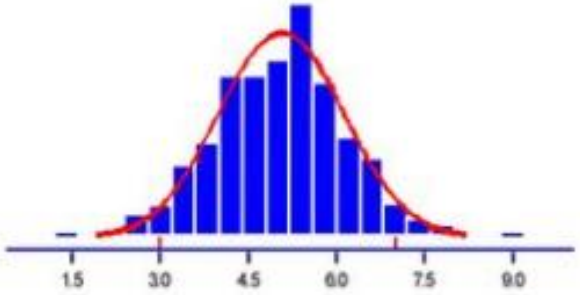
Amount of information is **30**

More Data = For a feature → More accurate mean value

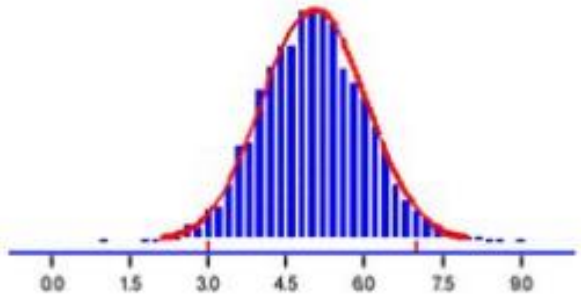
Normal Distribution



표본수:20개



표본수:200개



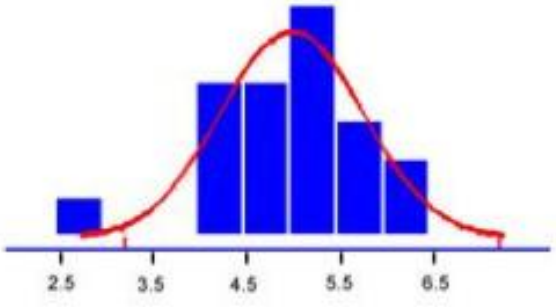
표본수:2,000개

Law of Large Numbers

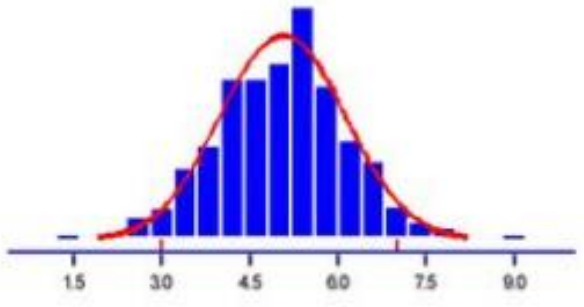
표본집단의 크기가 커지면, 그 표본평균이 모평균에 가까워 짐. 즉, 취합하는 표본의 수가 많을 수록, 통계적 정확도는 올라간다.

More Data = More features

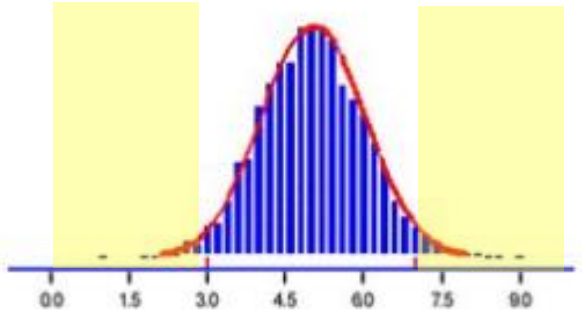
Rare feature could be retained.



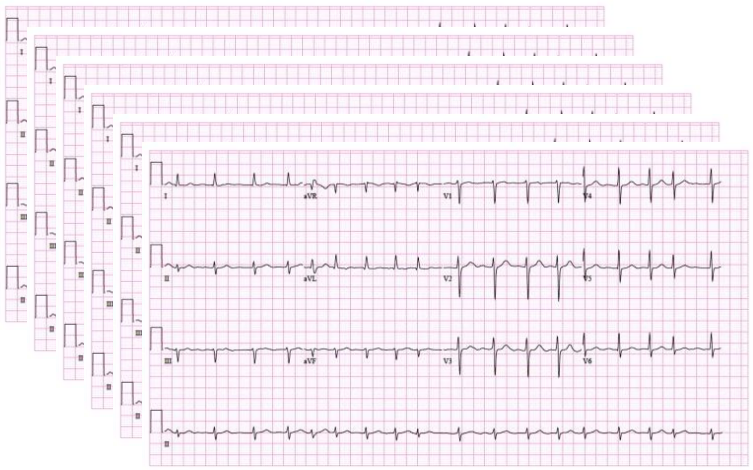
표본수:20개

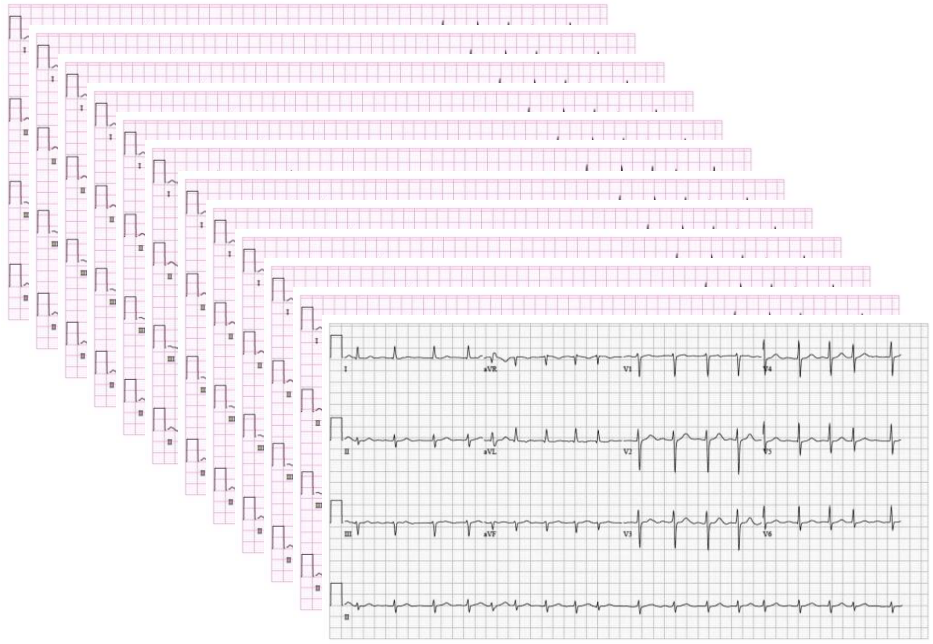


표본수:200개



표본수:2,000개





Empirical Risk Minimization : Goal of ML

$$Remp(h) = \frac{1}{n}\sum_{i=1}^{n} L(h(x_i), Yi)$$

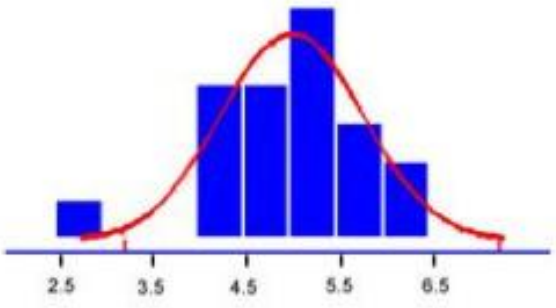
XML data

Ground Truth

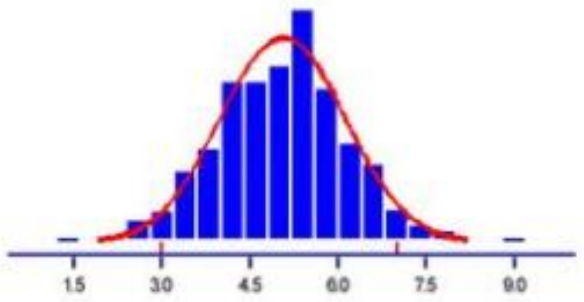
모델함수

평균

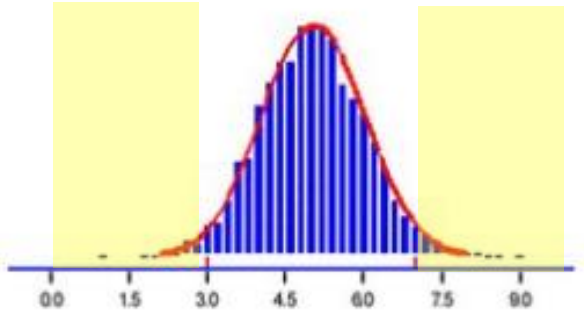
Loss Function : 모델과 Ground Truth와의 차이



표본수:20개

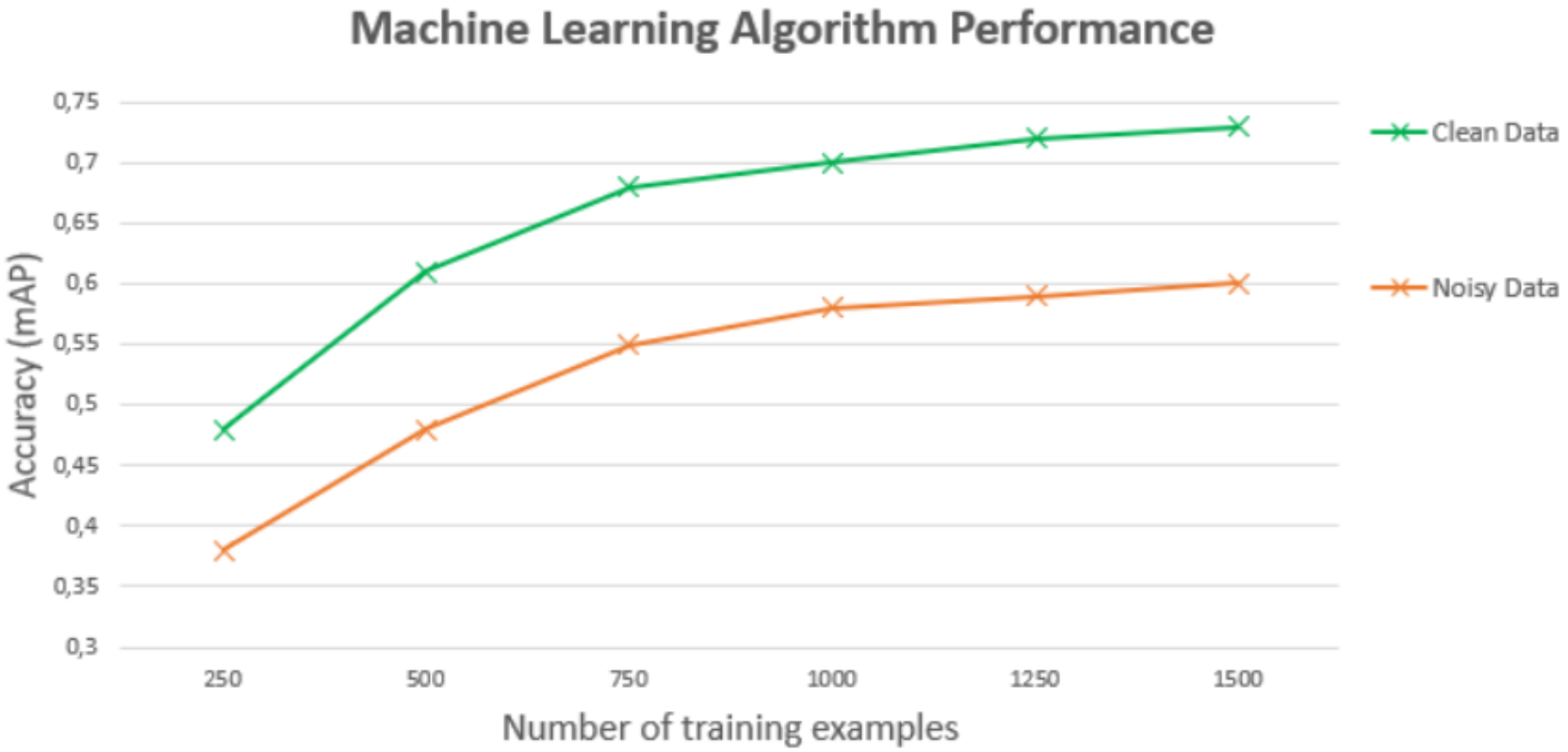


표본수:200개



표본수:2,000개

The more information, the better performance ?



Thank you

mineok.chang@vuno.co
oklizard81@gmail.com

T 02.515.6646 E hello@vuno.co H www.vuno.co