GRADIENT DESCENT IS TAKING YOUR ANDROID DEVELOPER JOB AWAY

Photo by Adriana Harakalova, 2021

ALPHAGO

Photo by Adriana Harakalova, 2021

Andrej Karpathy
@karpathy

Gradient descent can write code better than you. I'm sorry.

10:56 PM · Aug 4, 2017 · Twitter Web Client

493 Retweets **92** Quote Tweets **2,611** Likes

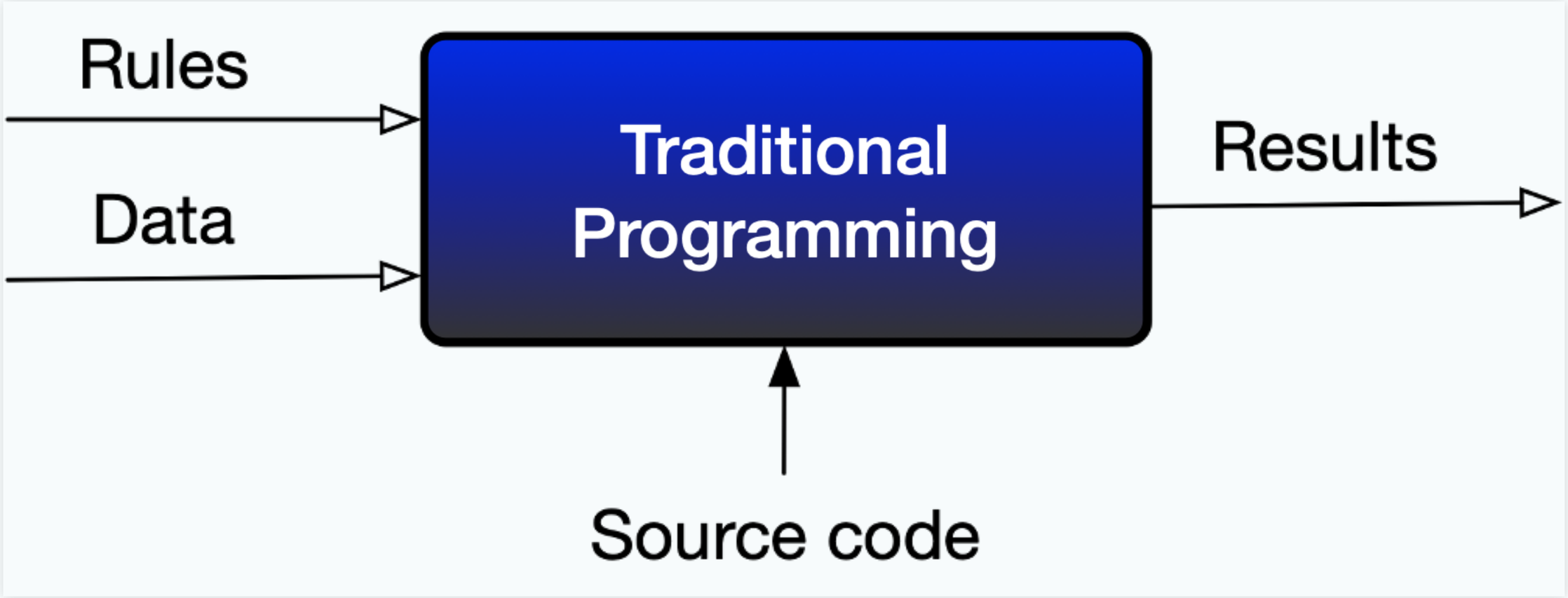
Photo by Adriana Harakalova, 2021

SOFTWARE 2.0


Mul
B ⟨64⟩
Add
B ⟨64⟩
Relu
Conv
W ⟨64×256×1×1⟩
BatchNormalization
scale
Conv
W ⟨96×64×3×3⟩
Mul
B ⟨96⟩
Add
B ⟨96⟩
Relu
Conv
W

SOFTWARE DEVELOPMENT

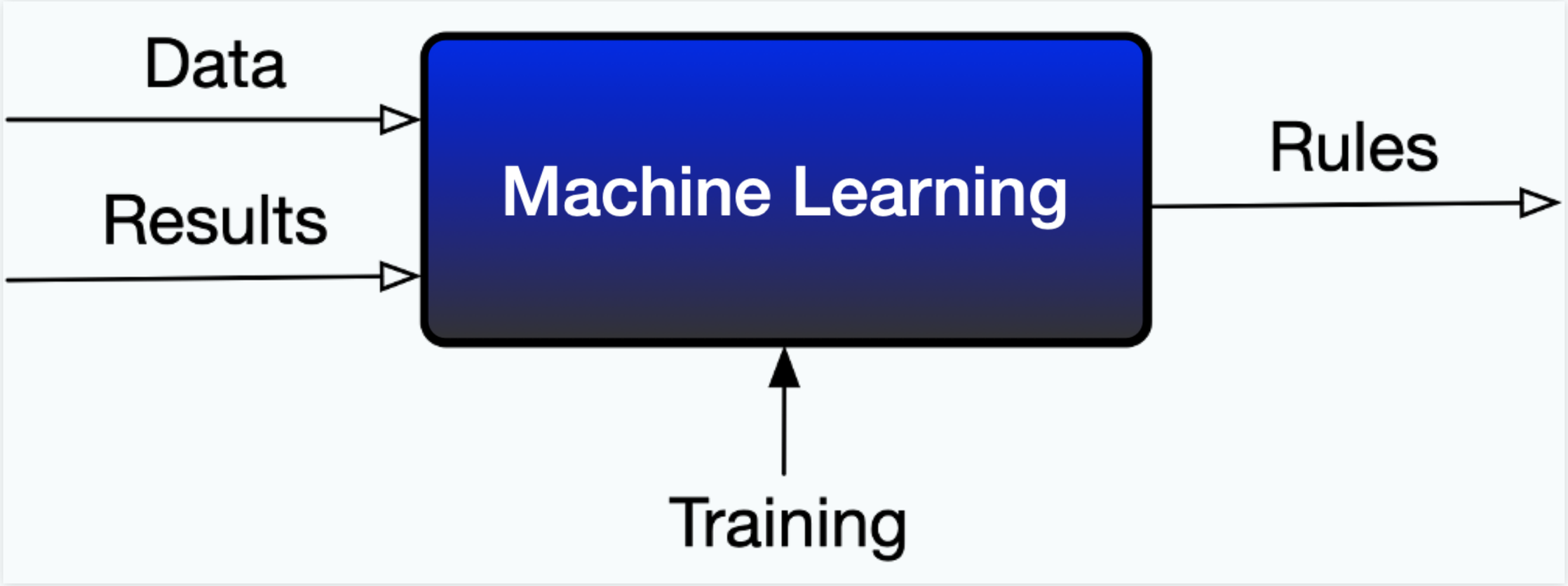
SOFTWARE DEVELOPMENT



MACHINE LEARNING

Photo by Adriana Harakalova, 2021

TRAINING



EXAMPLES

Photo by Adriana Harakalova, 2021

APPLICATIONS IN REAL LIFE

Photo by Adriana Harakalova, 2021

APPLICATIONS IN REAL LIFE

- visual recognition

APPLICATIONS IN REAL LIFE

- visual recognition
- speach recognition & syntesis

APPLICATIONS IN REAL LIFE

- visual recognition
- speach recognition & syntesis
- machine translation

APPLICATIONS IN REAL LIFE

- visual recognition
- speach recognition & syntesis
- machine translation
- recommendation engines

APPLICATIONS IN REAL LIFE

- visual recognition
- speach recognition & syntesis
- machine translation
- recommendation engines
- natural language processing

APPLICATIONS IN REAL LIFE

- visual recognition
- speach recognition & syntesis
- machine translation
- recommendation engines
- natural language processing
- financial sector

ROCKET SCIENCE

APOLLO 11 GUIDANCE COMPUTER

```
# SINGLE PRECISION SINE AND COSINE

                COUNT*  $$/INTER
SPCOS           AD      HALF            # ARGUMENTS SCALED AT PI
SPSIN           TS      TEMK
                TCF     SPT
                CS      TEMK
SPT             DOUBLE
                TS      TEMK
                TCF     POLLEY
                XCH     TEMK
                INDEX   TEMK
                AD      LIMITS
                COM
                AD      TEMK
                TS      TEMK
                TCF     POLLEY
                TCF     ARG90
POLLEY          EXTEND
                MP      TEMK
                TS      SQ
                EXTEND
                MP      C5/2
                AD      C3/2
                EXTEND
                MP      SQ
                AD      C1/2
                EXTEND
                MP      TEMK
                DDOUBL
                TS      TEMK
                TC      Q
ARG90           INDEX   A
                CS      LIMITS
                TC      Q               # RESULT SCALED AT 1.
```

Photo by Adriana Harakalova, 2021

MARGARET HAMILTON

Photo by Adriana Harakalova, 2021

SINE FUNCTION

$$y(x) = sin(x)$$

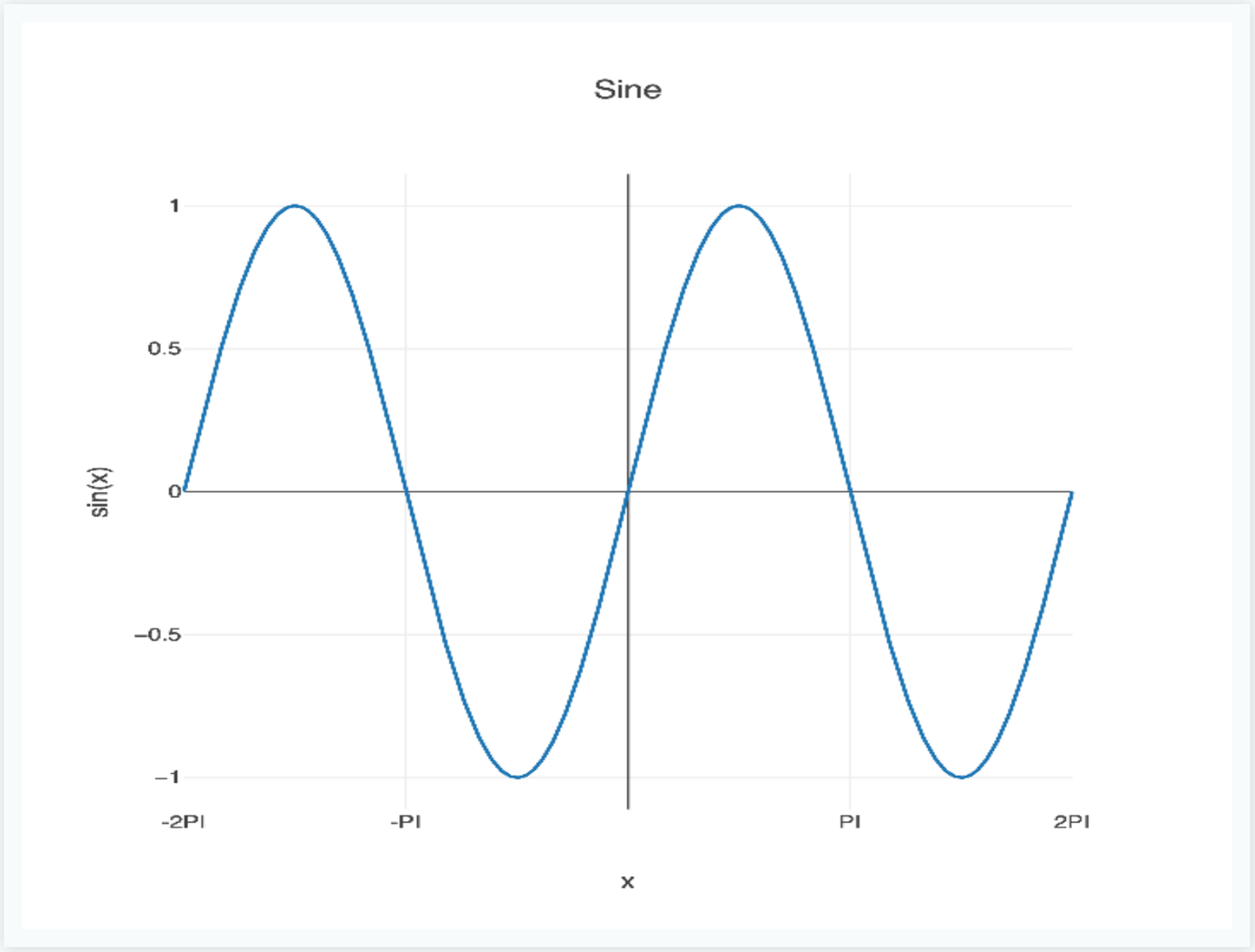


Photo by Adriana Harakalova, 2021

APOLLO 11 APPROXIMATION

$$y(x) = 0.7853134x - 0.3216147x^3 + 0.036551x^5$$

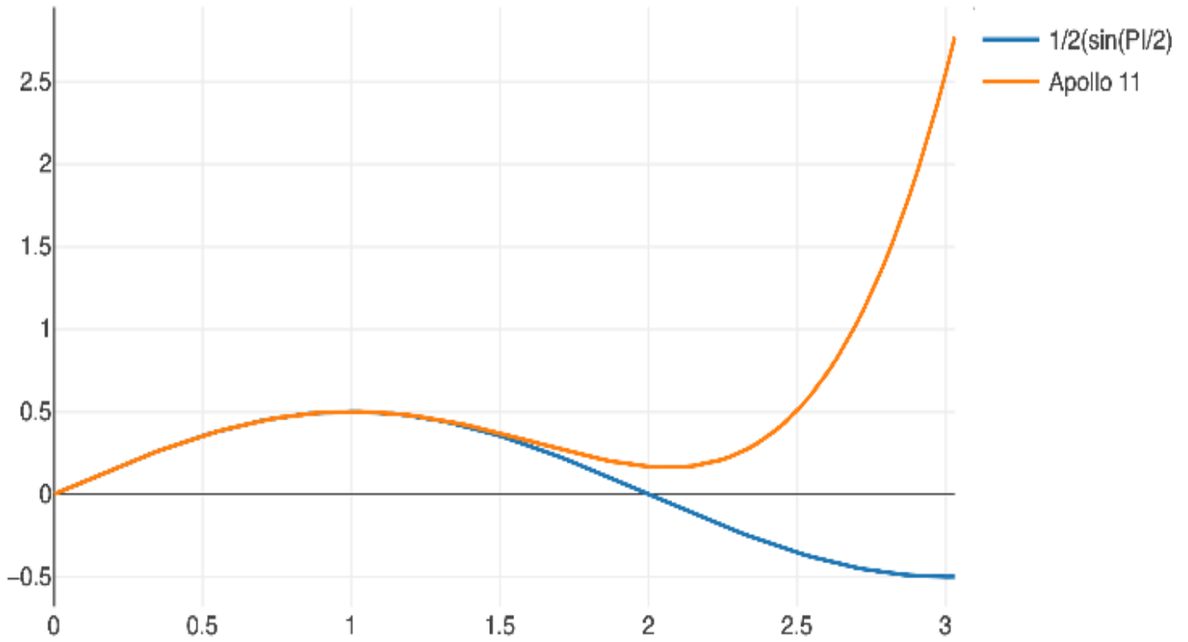
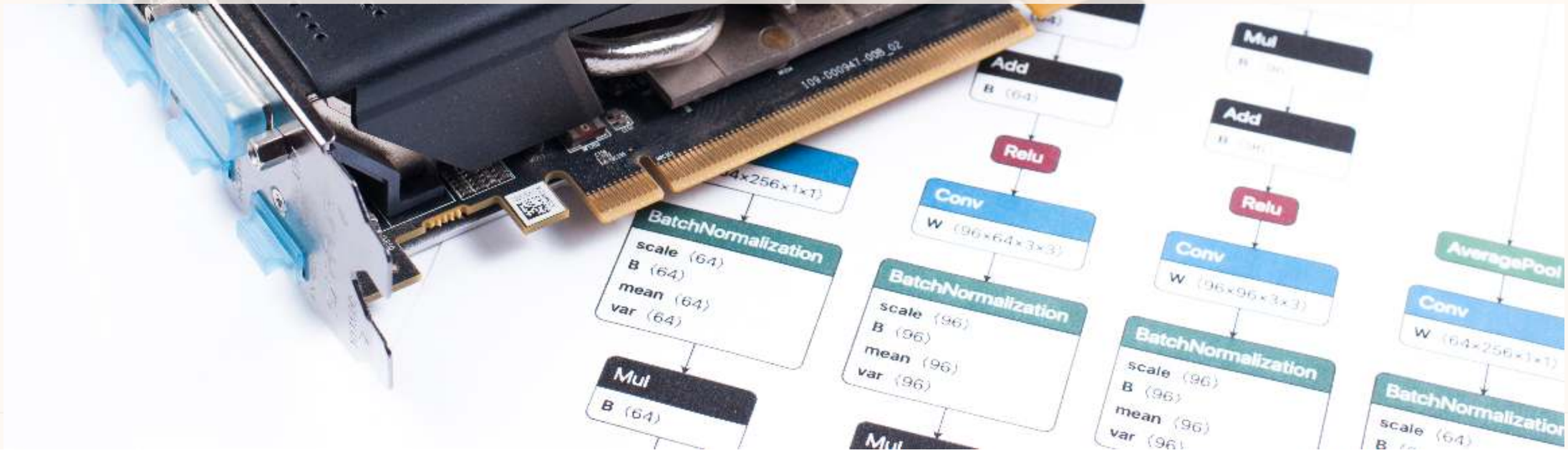
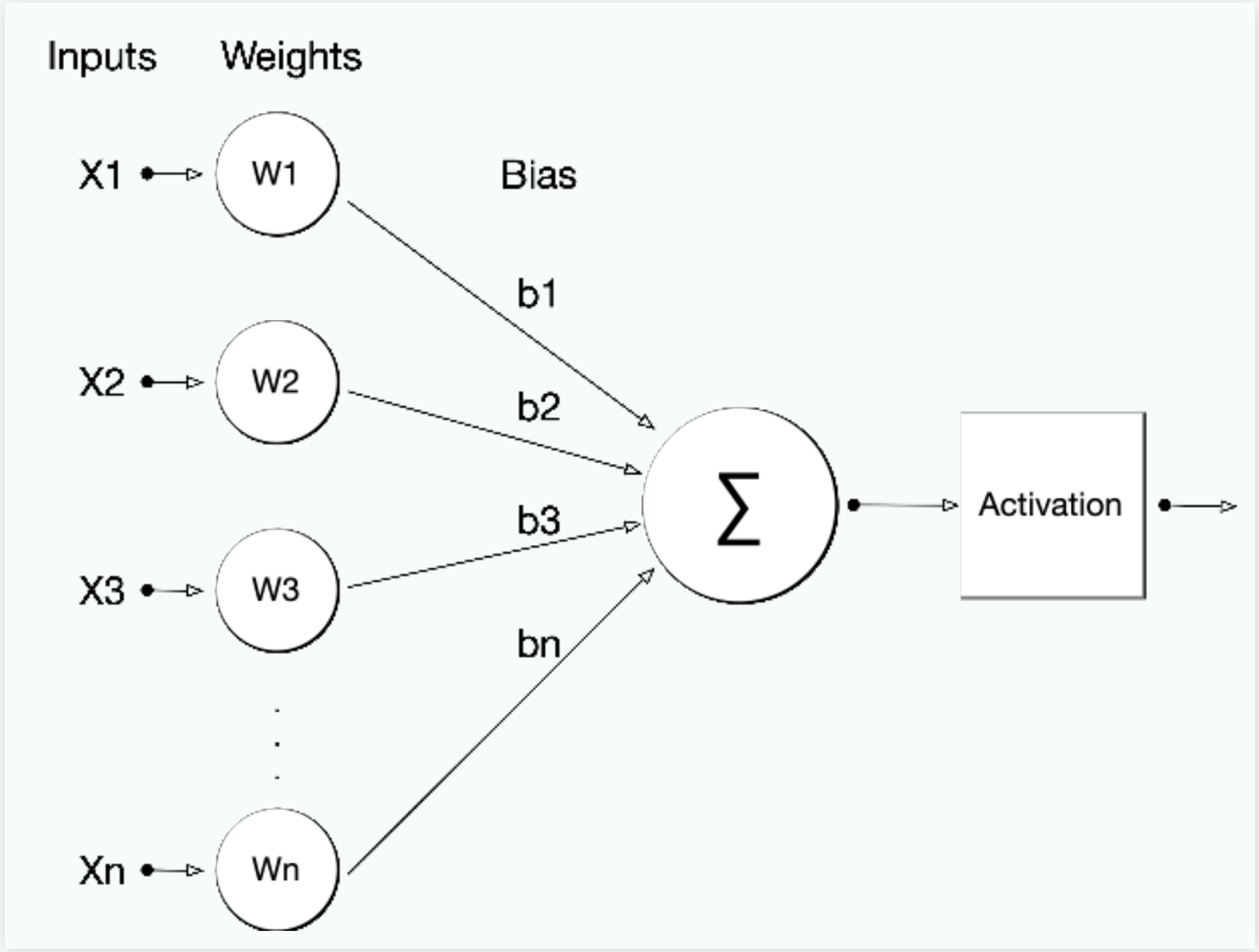


Photo by Adriana Harakalova, 2021

NEURAL NETWORKS

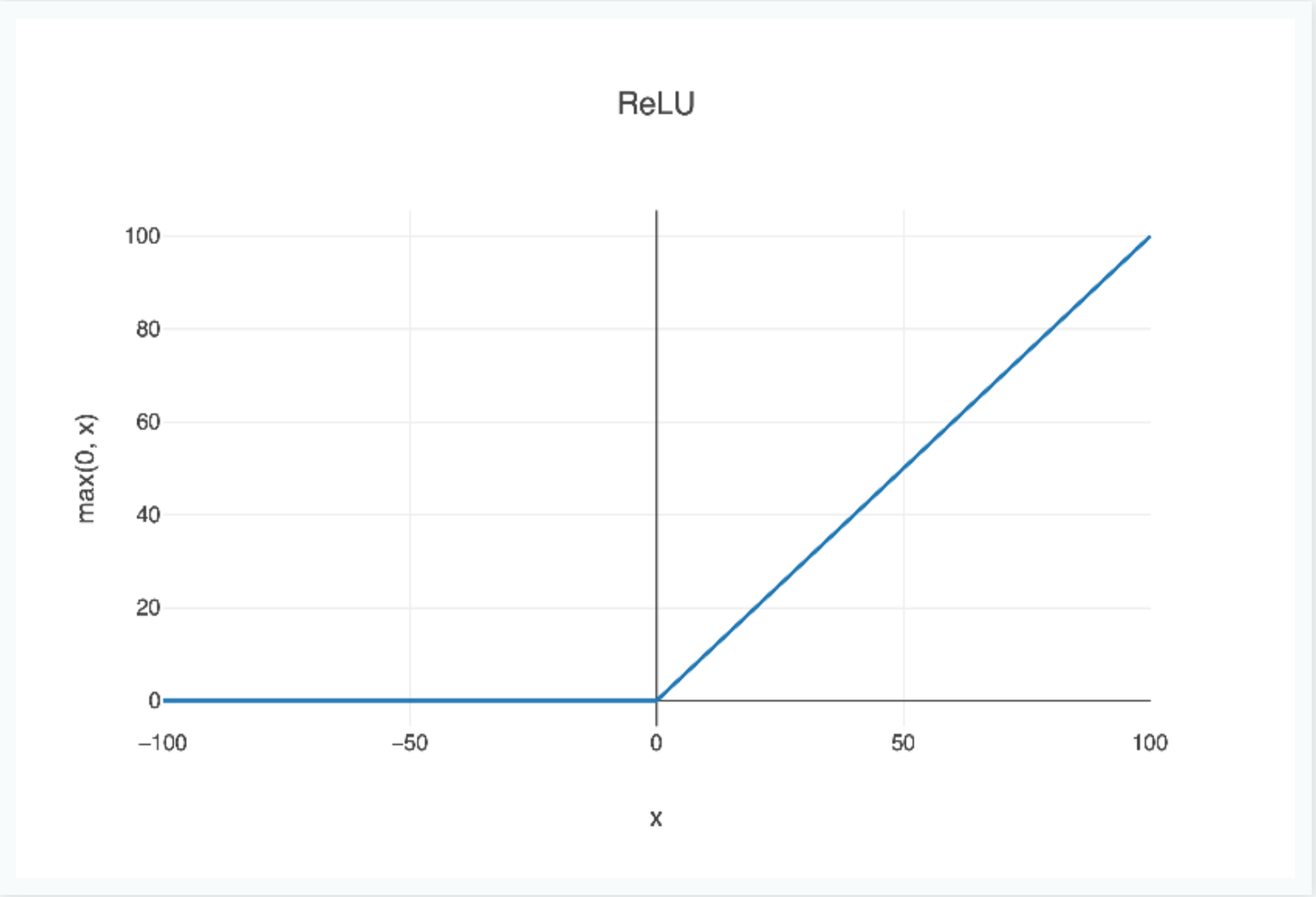


NEURON



$$\sum_{i=0}^{n} x_i w_i + b_i$$

ACTIVATION FUNCTION RELU



ReLU

$$f(x) = max(0, x)$$

NEURAL NETWORK ARCHITECTURE

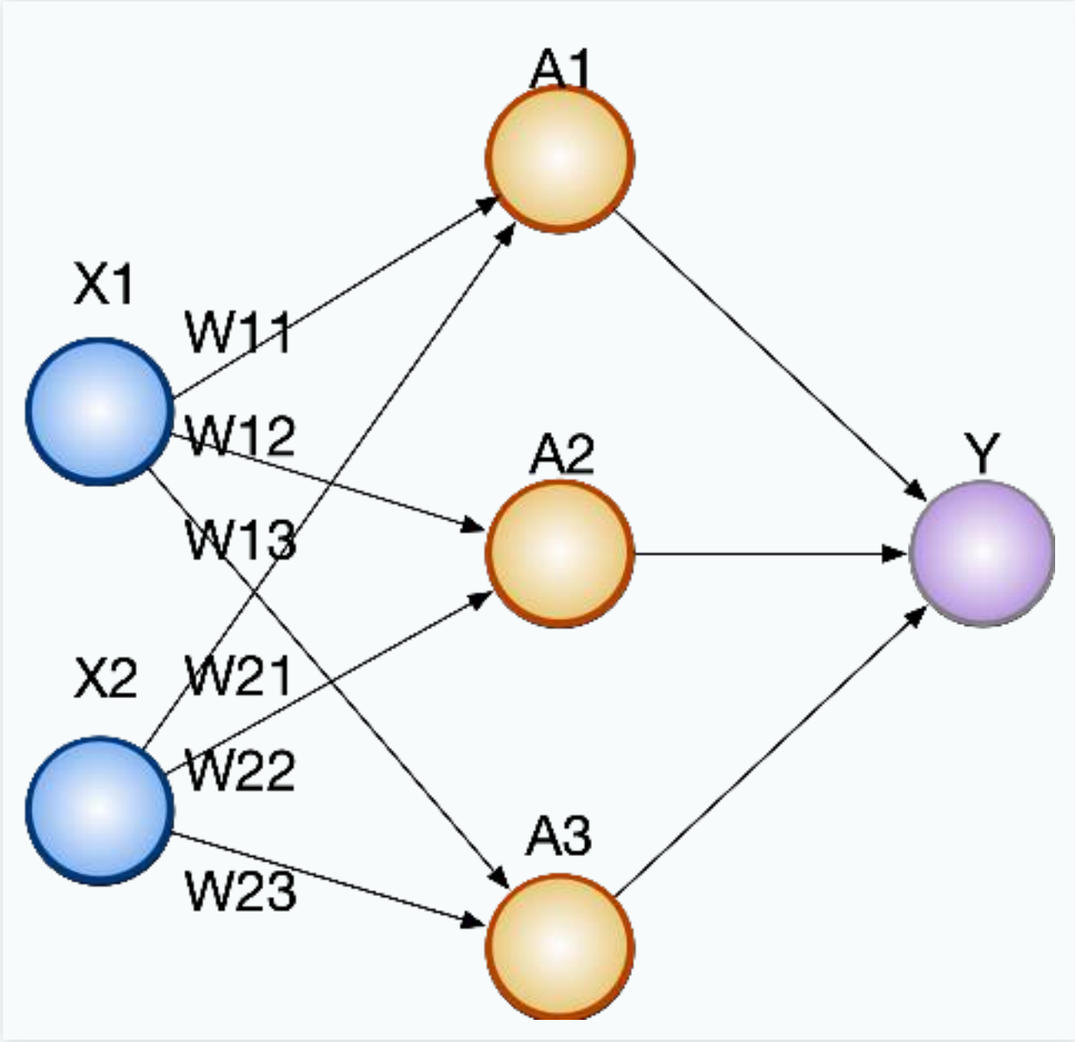
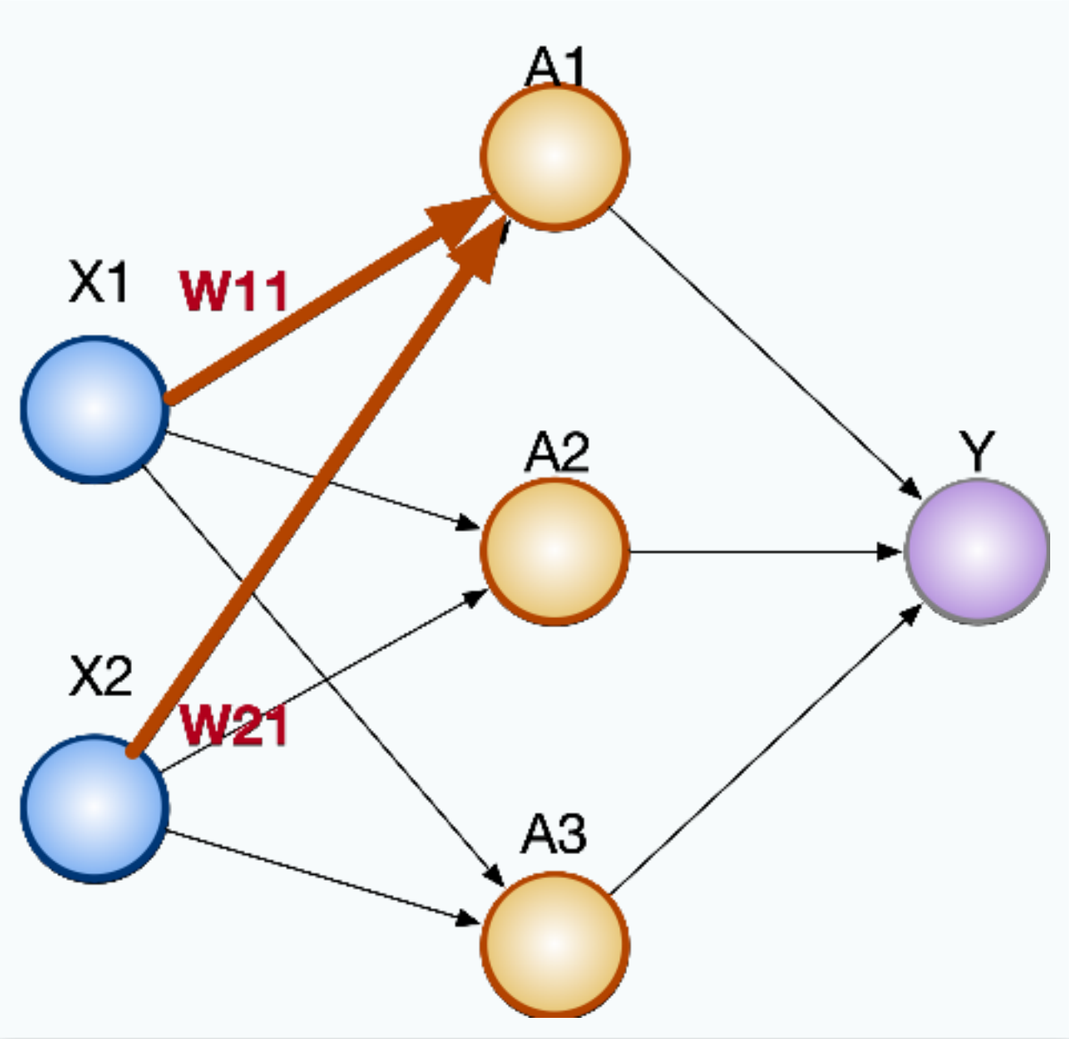


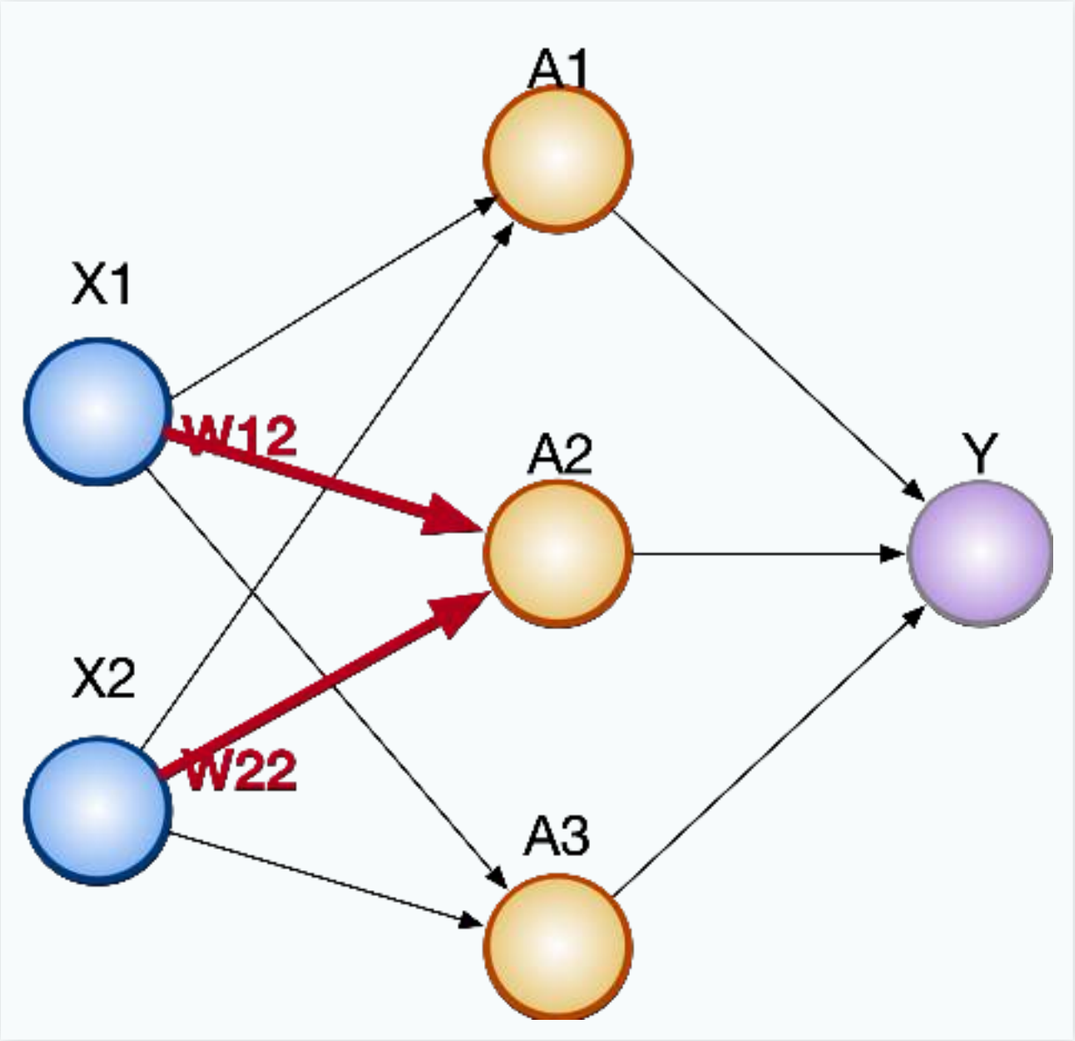
Photo by Adriana Harakalova, 2021

NEURAL NETWORK ARCHITECTURE



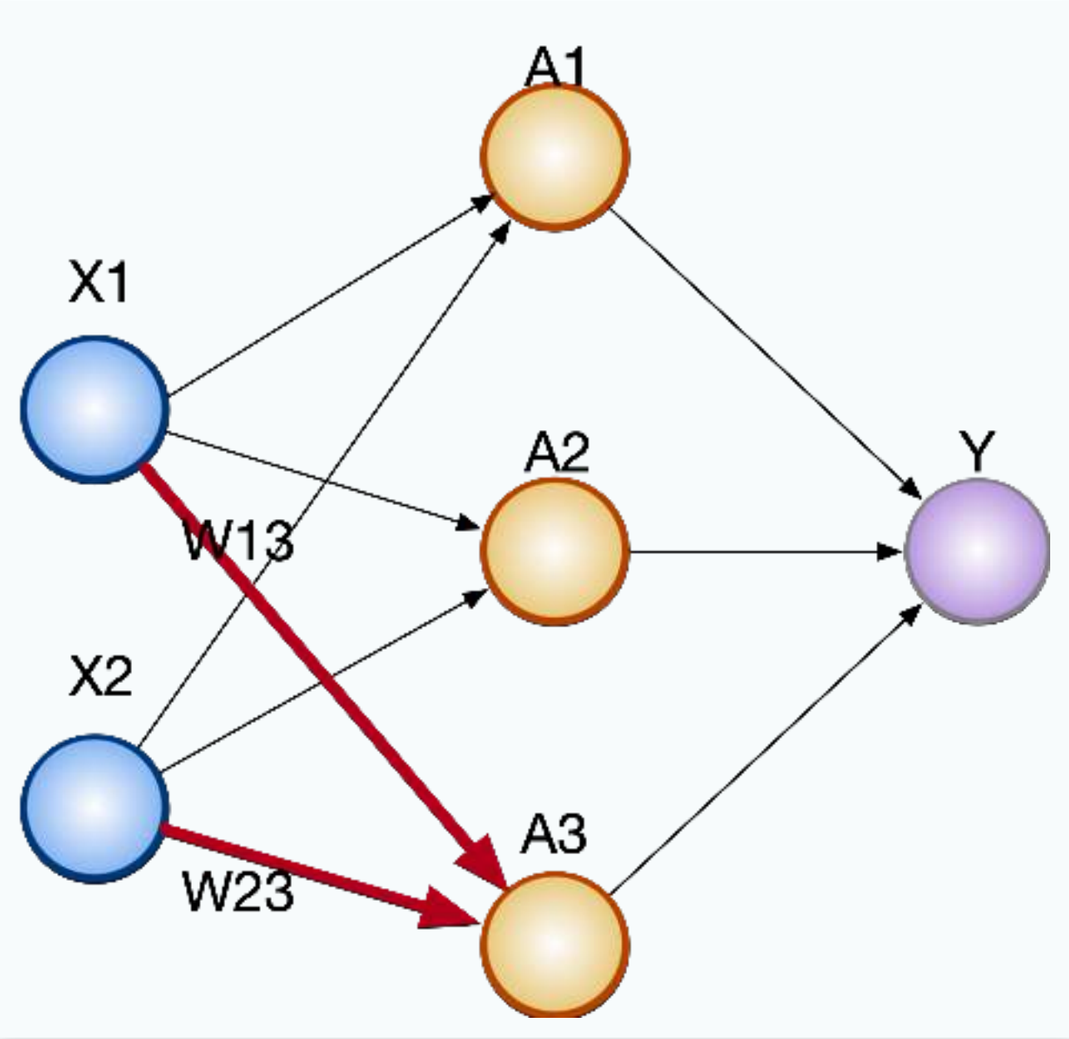
- $A_1 = X_1 W_{11} + X_2 W_{21}$

NEURAL NETWORK ARCHITECTURE



- $A_2 = X_2 W_{21} + X_1 W_{21}$

NEURAL NETWORK ARCHITECTURE



- $A_3 = X_1 W_{13} + X_2 W_{23}$

MATRICES

$$A = X.W$$

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{21} & W_{23} \end{bmatrix}$$

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{12} \end{bmatrix}$$

$$A_{11} = X_1 W_{11} + X_2 W_{21}$$

MATRICES

$$A = X . W$$

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{21} & W_{23} \end{bmatrix}$$

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{12} \end{bmatrix}$$

$$A_{11} = X_1 W_{21} + X_2 W_{22}$$

MATRICES

$$A = X . W$$

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{21} & W_{23} \end{bmatrix}$$

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{12} \end{bmatrix}$$

$$A_{11} = X_1 W_{13} + X_2 W_{23}$$

FORWARD PROPAGATION & INFERENCE

```
fun predict(inputMatrix: Matrix): Matrix {
        var data = inputMatrix
        layers.forEach { layer ->
            val weighted = data.multiply(layer.weights)
            val biased = weighted.add(layer.bias)
            data = layer.activation(biased)
        }
        return data
    }
```

TRAINING

Mean square error

$$\frac{1}{N} \sum_{i=0}^{N} y_i + \hat{y}_i$$

BACK PROPAGATION & GRADIENT DESCENT

SOLUTION IN KOTLIN

```
val model1 = Sequential.of(
    Input(1, name = "input_1"),
    Dense(
        20,
        Activations.Relu,
        kernelInitializer = HeNormal(SEED),
        biasInitializer = HeUniform(SEED),
        name = "dense_1"
    ),
    Dense(
        20,
        Activations.Relu,
        kernelInitializer = HeNormal(SEED),
        biasInitializer = HeUniform(SEED),
        name = "dense_2"
    ),
    Dense(1, Activations.Linear, name = "dense_3")
)
```

DEMO TIME

Photo by Adriana Harakalova, 2021

LEARNING TO LEARN

LEARNING TO LEARN

- pimp up your brain

LEARNING TO LEARN

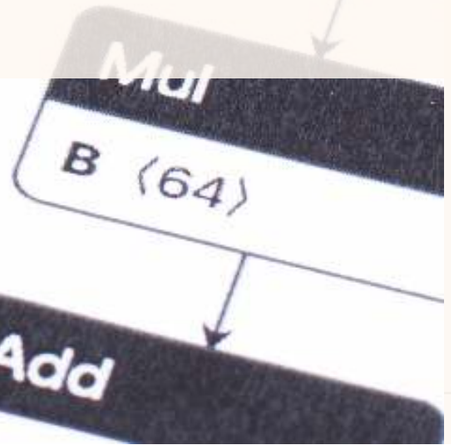
- pimp up your brain
- healthy lifestyle

LEARNING TO LEARN

- pimp up your brain
- healthy lifestyle
- massive open online courses(MOOC)

LEARNING TO LEARN

- pimp up your brain
- healthy lifestyle
- massive open online courses(MOOC)
- learning by doing



DATA SCIENCE FOR ANDROID DEVELOPER

Photo by Adriana Harakalova, 2021

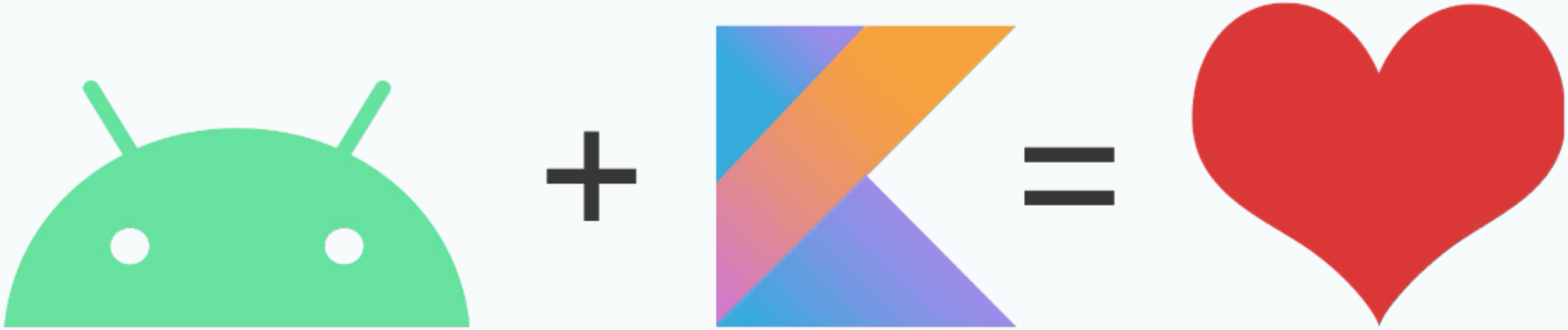


Photo by Adriana Harakalova, 2021

DATA SCIENCE AND KOTLIN

Photo by Adriana Harakalova, 2021

DATA SCIENCE AND KOTLIN

- Jetbrains and data science

DATA SCIENCE AND KOTLIN

- Jetbrains and data science
- kscript

DATA SCIENCE AND KOTLIN

- Jetbrains and data science
- kscript
- kmath, kotlingrad

DATA SCIENCE AND KOTLIN

- Jetbrains and data science
- kscript
- kmath, kotlingrad
- krangl, plotly.kt

Photo by Adriana Harakalova, 2021

DATA SCIENCE AND KOTLIN

- Jetbrains and data science
- kscript
- kmath, kotlingrad
- krangl, plotly.kt
- KotlinDL

MACHINE LEARNING ON ANDROID

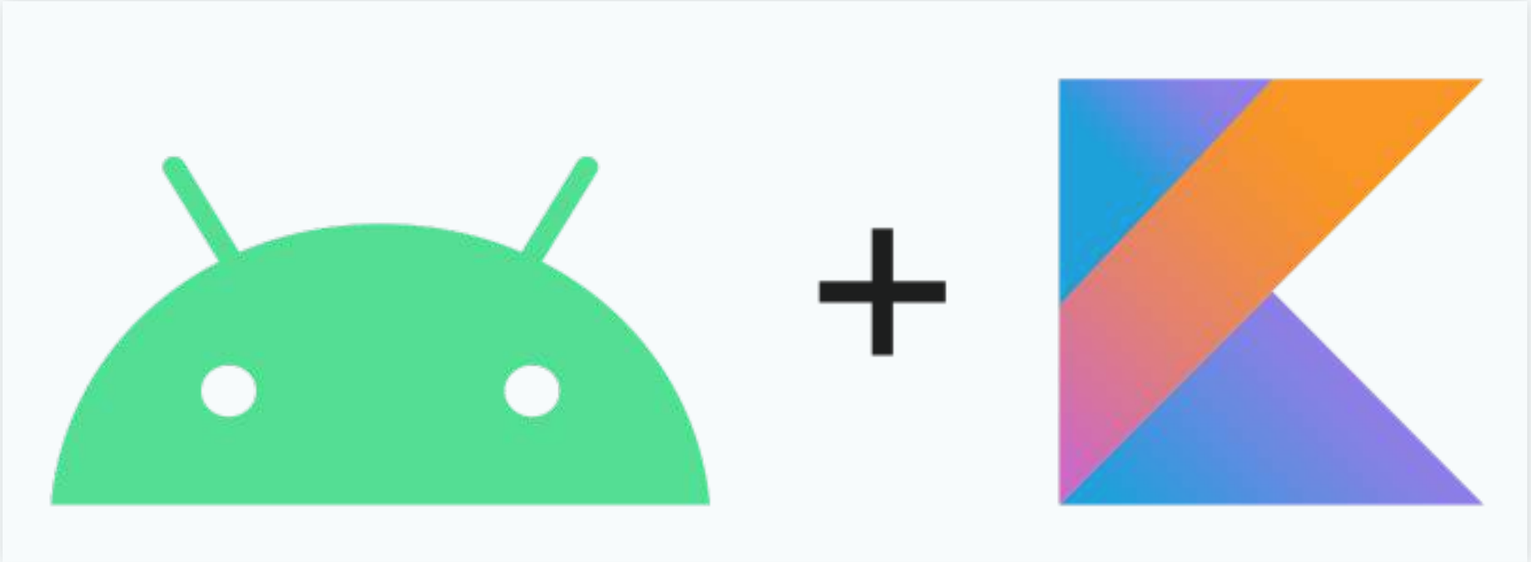
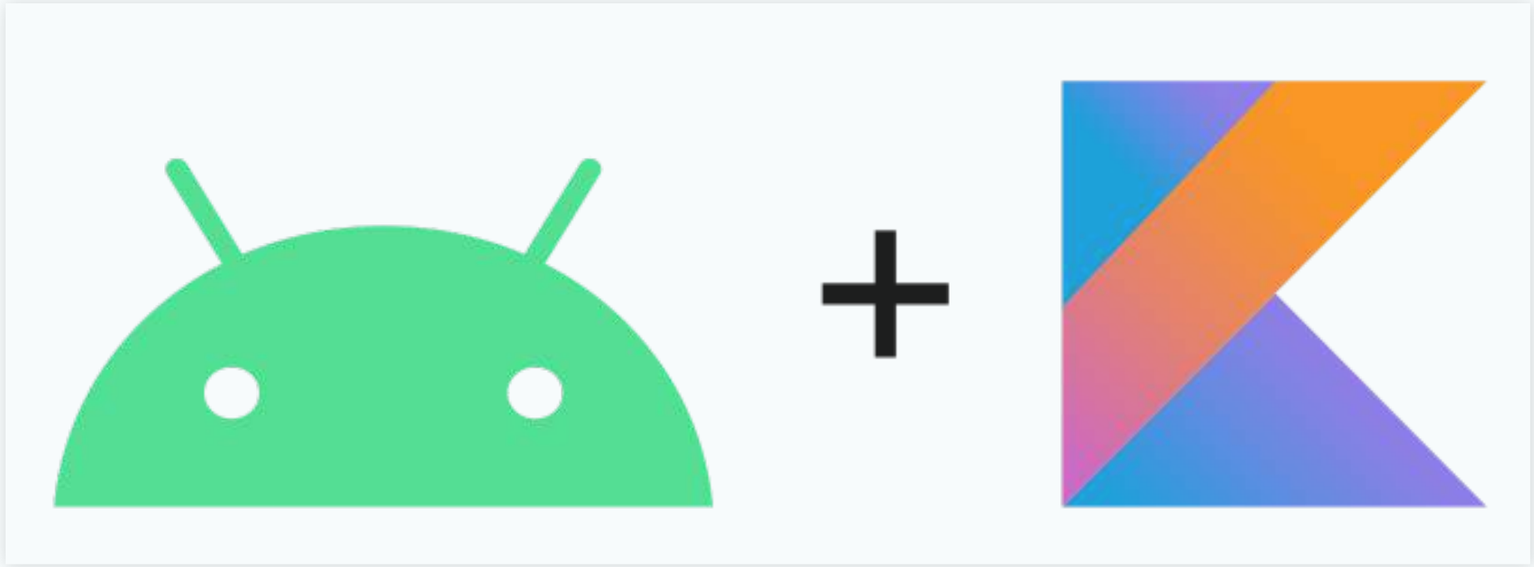


Photo by Adriana Harakalova, 2021

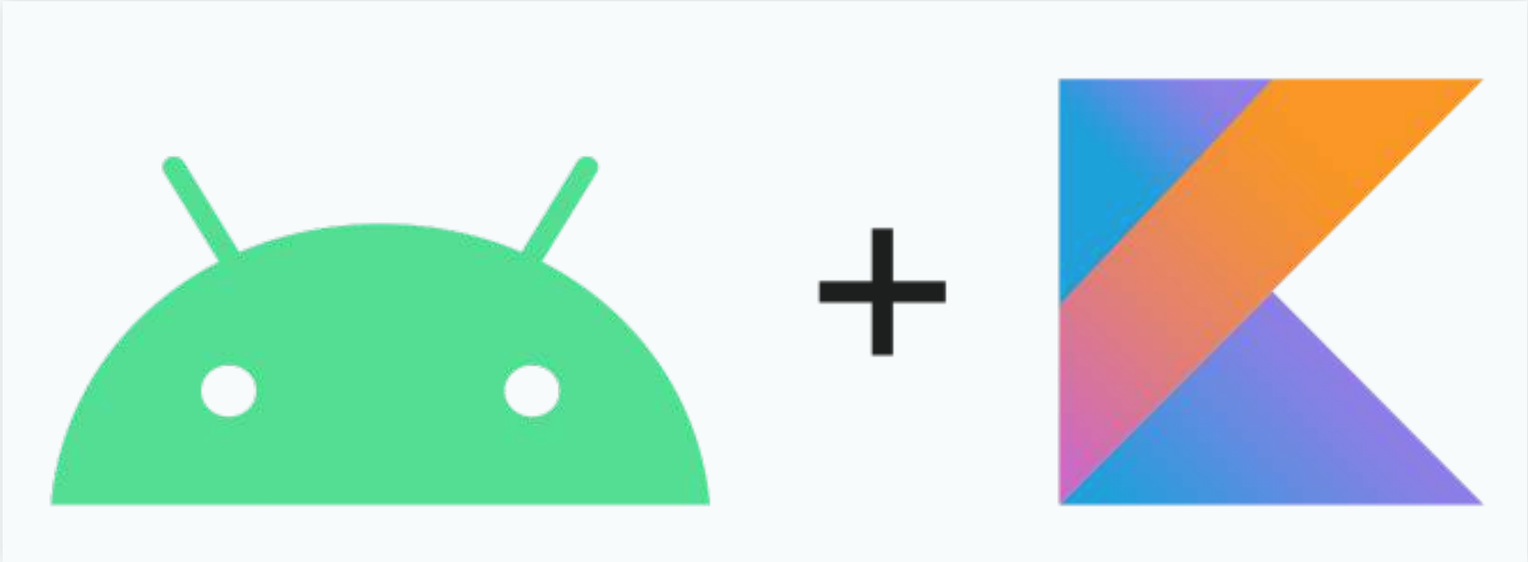
MACHINE LEARNING ON ANDROID

- Kotlin & JVM



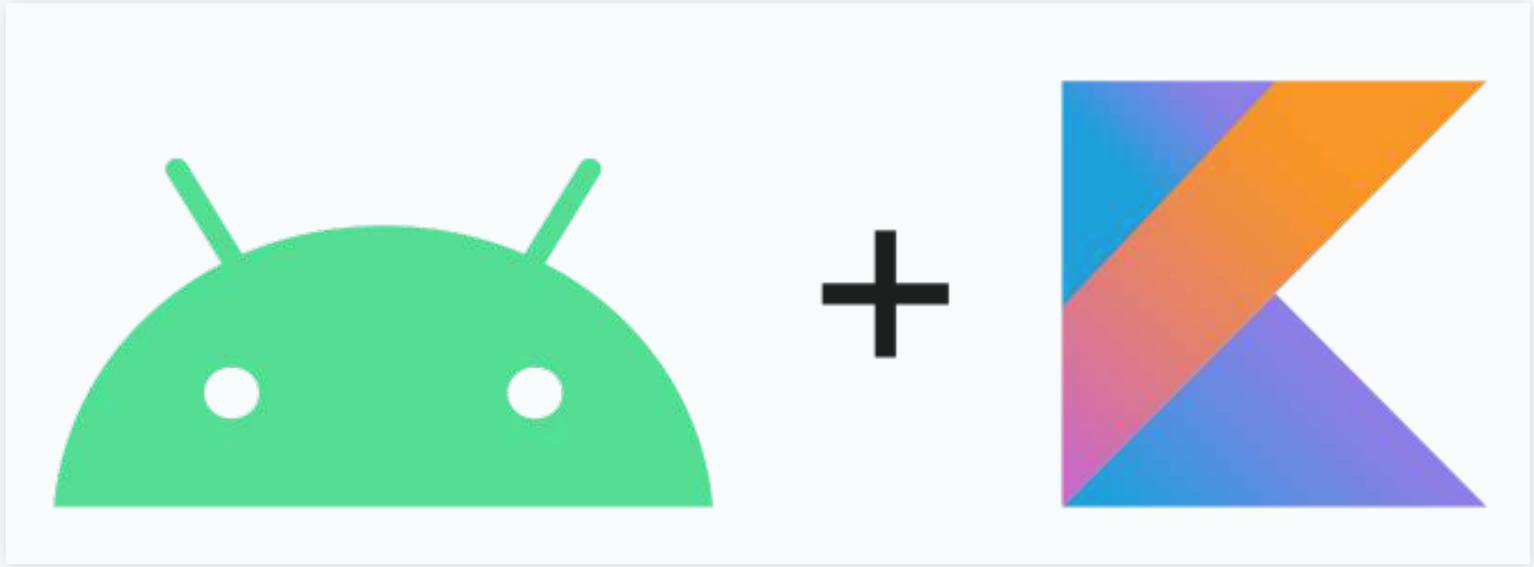
MACHINE LEARNING ON ANDROID

- Kotlin & JVM
- machine learning on the edge



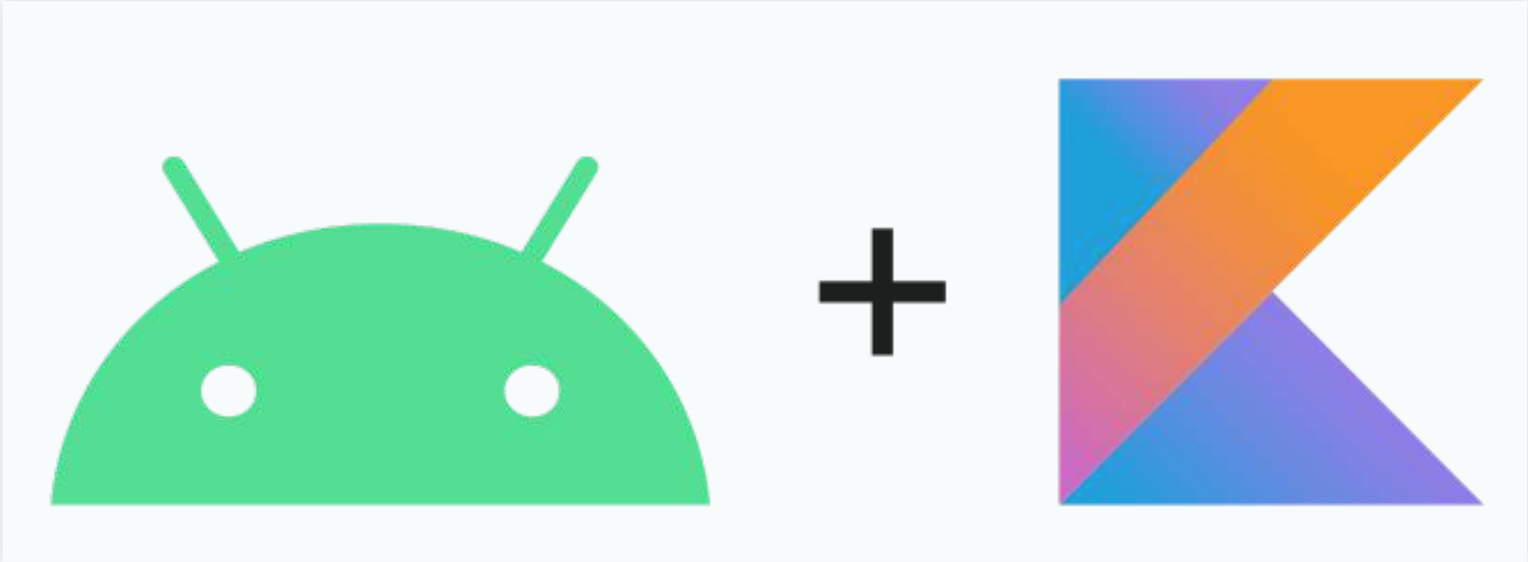
MACHINE LEARNING ON ANDROID

- Kotlin & JVM
- machine learning on the edge
- TinyML



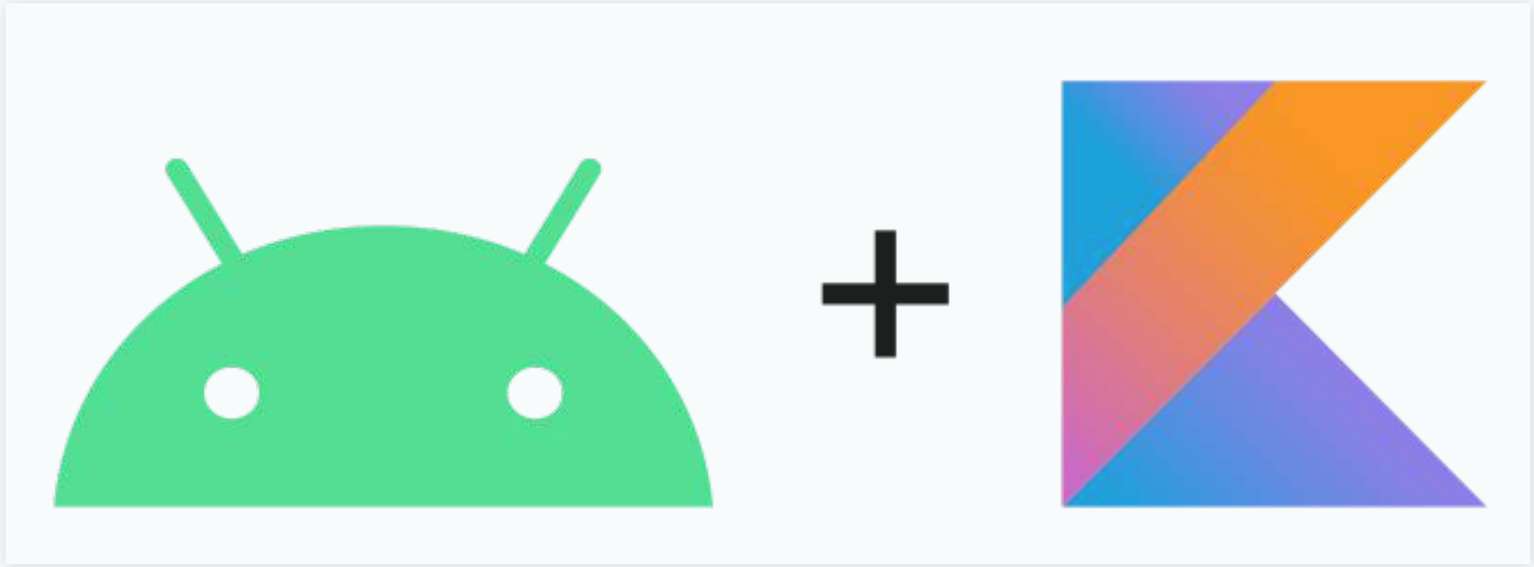
MACHINE LEARNING ON ANDROID

- Kotlin & JVM
- machine learning on the edge
- TinyML
- Tensorflow & Tensorflow Lite



MACHINE LEARNING ON ANDROID

- Kotlin & JVM
- machine learning on the edge
- TinyML
- Tensorflow & Tensorflow Lite
- Tensorflow Hub




DURACELL
PLUS
THANK YOU
Conv
W (64×256×1×1)
BatchNormalization
scale (64)
B (64)
mean (64)
var (64)
Mul
B (64)
Add
B (64)
W (64×256×1×1)
BatchNormalization
scale (96)
B (96)
mean (96)
var (96)
Mul
B (96)

SLIDES

- slides
- code
- notebook

Photo by Adriana Harakalova, 2021

LINKS

Slides and code, arXiv papers, papers with code, AlphaGo, Scale of the human brain, The total number of neurons in the human neocortex unbiasedly estimated using optical disectors

Apollo 11 Program

Wikipedia about Apollo Guidance Computer, Simplified implementation of sin/cos, https://de.wikipedia.org/wiki/Margaret_Hamilton_(Wissenschaftlerin)#/media/Datei:Margaret_Hamilton_in_action.jpg, Virtual AGC project

Learning

Learning how to learn,

Android&Math

Flag Sam, Pose estimation Tensorflow lite sample, kotlingrad

Logos&pictures

Android, Kotlin, Adriana Harakalova ©2021, Margaret Hamilton, Wikipedia

Photo by Adriana Harakalova, 2021