

GRADIENT DESCENT IS  
TAKING YOUR  
ANDROID DEVELOPER  
JOB AWAY

The image features a Go board with a grid of intersections. Several black and white stones are placed on the board. In the background, a neural network diagram is visible, showing nodes and connections. The word "ALPHAGO" is overlaid in the center.

# ALPHAGO

Photo by Adriana Harakalova, 2021



**Andrej Karpathy** ✓

@karpathy



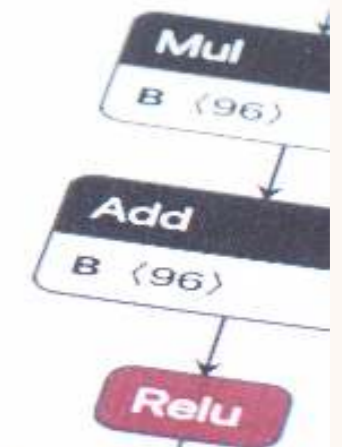
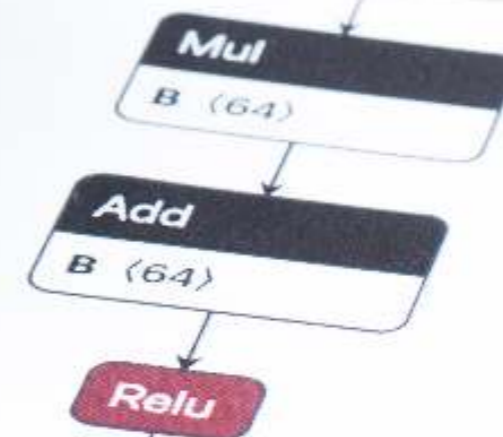
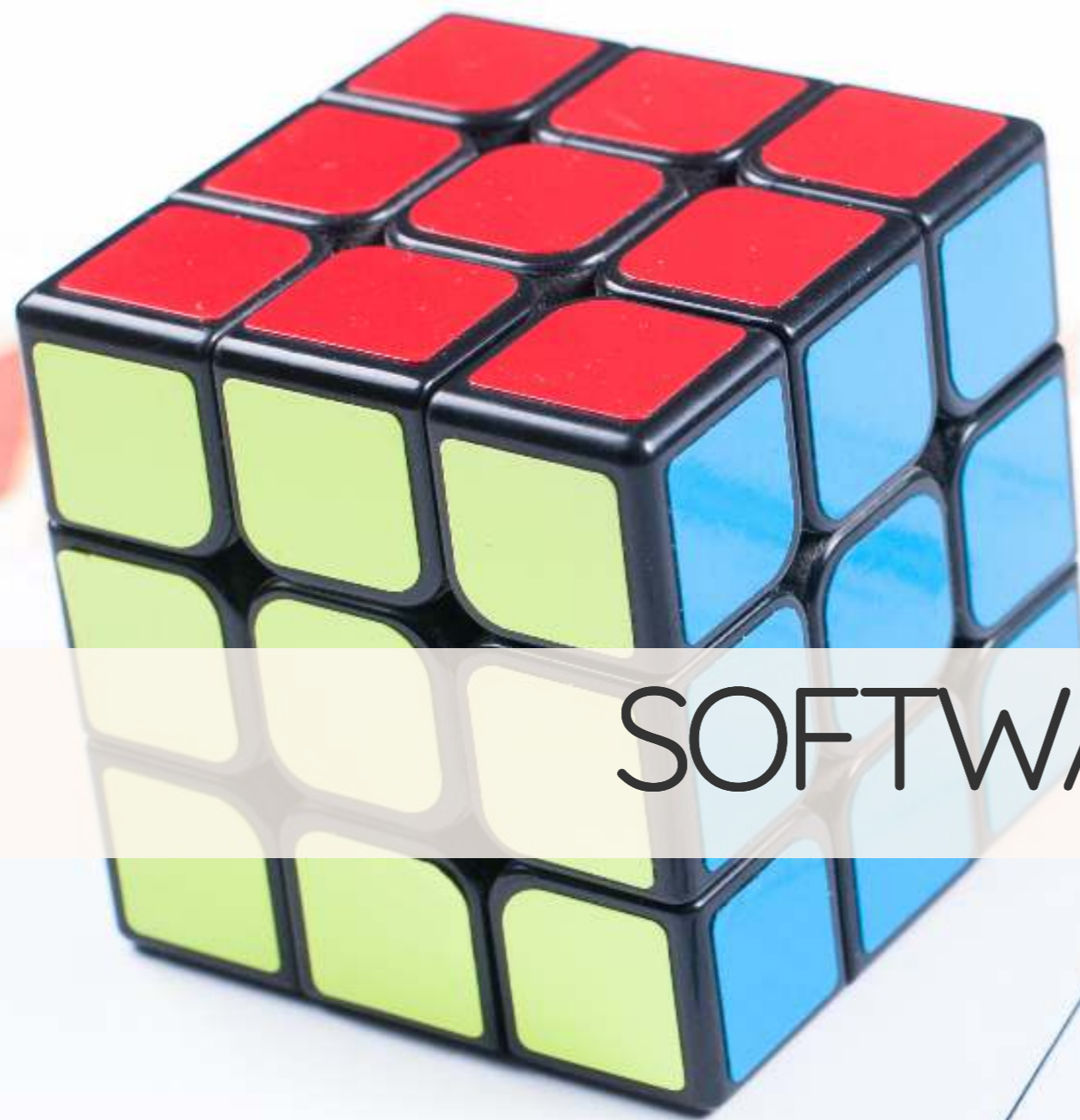
Gradient descent can write code better than you. I'm sorry.

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

---

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

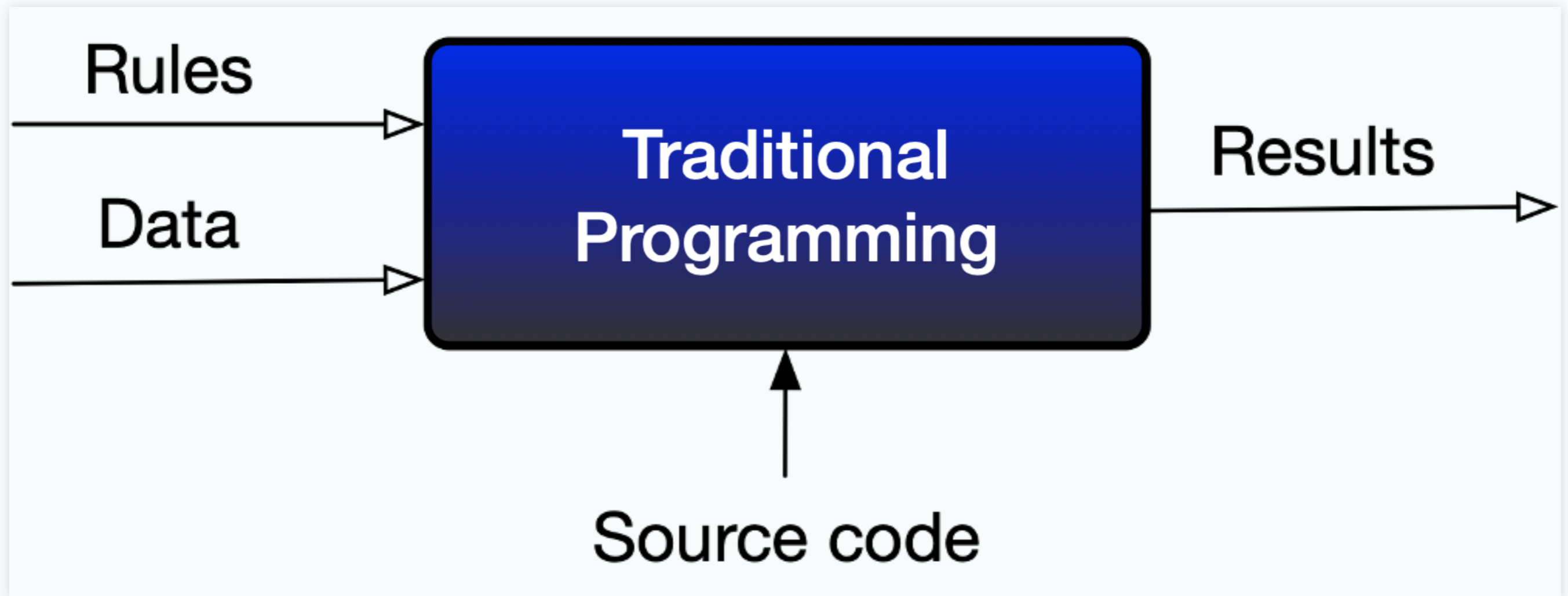
# SOFTWARE 2.0



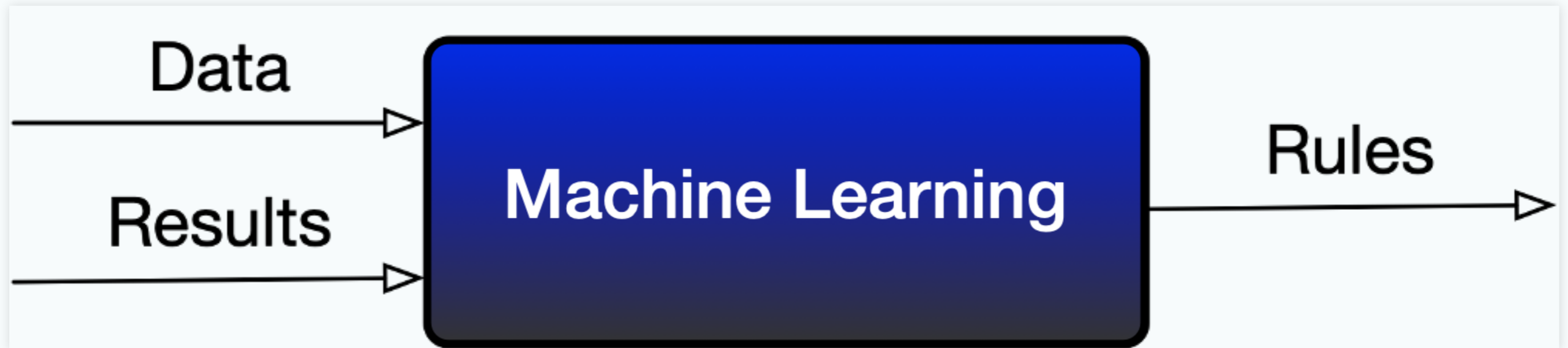
# SOFTWARE DEVELOPMENT



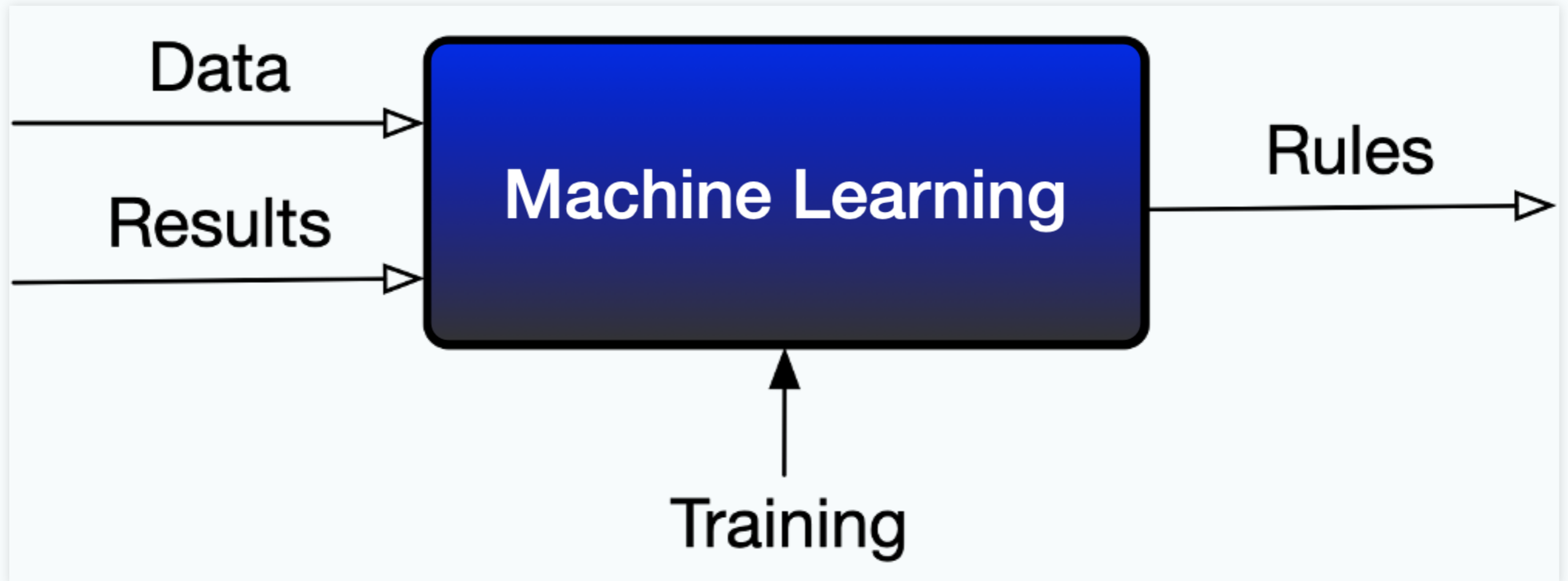
# SOFTWARE DEVELOPMENT



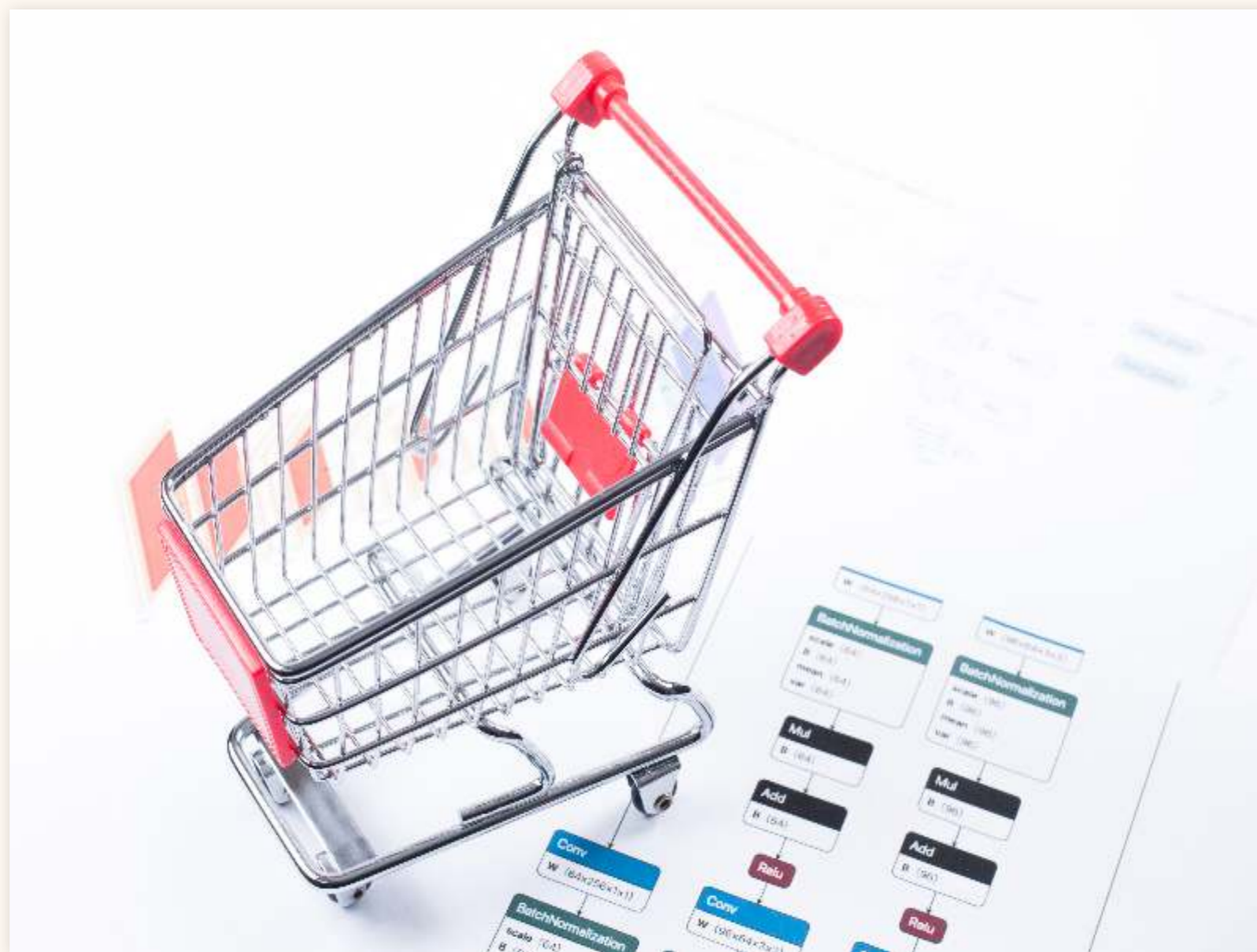
# MACHINE LEARNING



# TRAINING



# EXAMPLES



# APPLICATIONS IN REAL LIFE

# APPLICATIONS IN REAL LIFE

- visual recognition

# APPLICATIONS IN REAL LIFE

- visual recognition
- speech recognition & synthesis

# APPLICATIONS IN REAL LIFE

- visual recognition
- speech recognition & synthesis
- machine translation

# APPLICATIONS IN REAL LIFE

- visual recognition
- speech recognition & synthesis
- machine translation
- recommendation engines

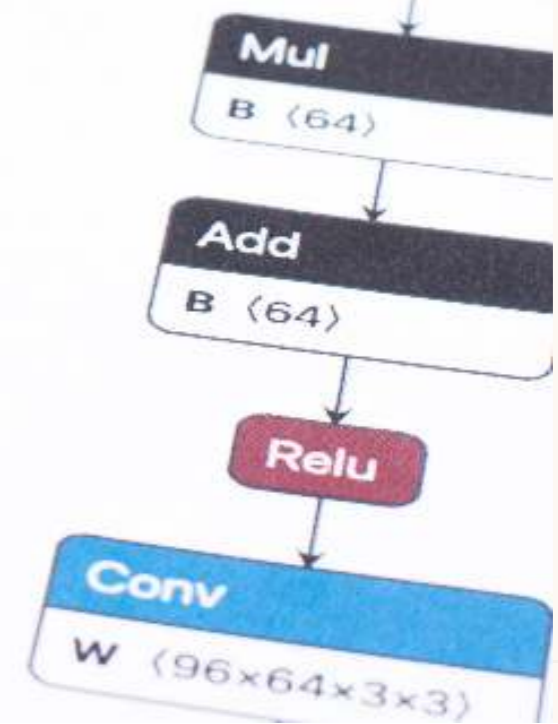
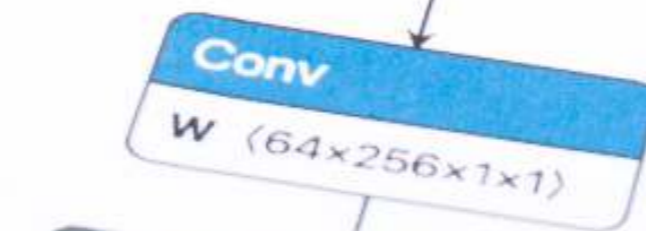
# APPLICATIONS IN REAL LIFE

- visual recognition
- speech recognition & synthesis
- machine translation
- recommendation engines
- natural language processing

# APPLICATIONS IN REAL LIFE

- visual recognition
- speech recognition & synthesis
- machine translation
- recommendation engines
- natural language processing
- financial sector

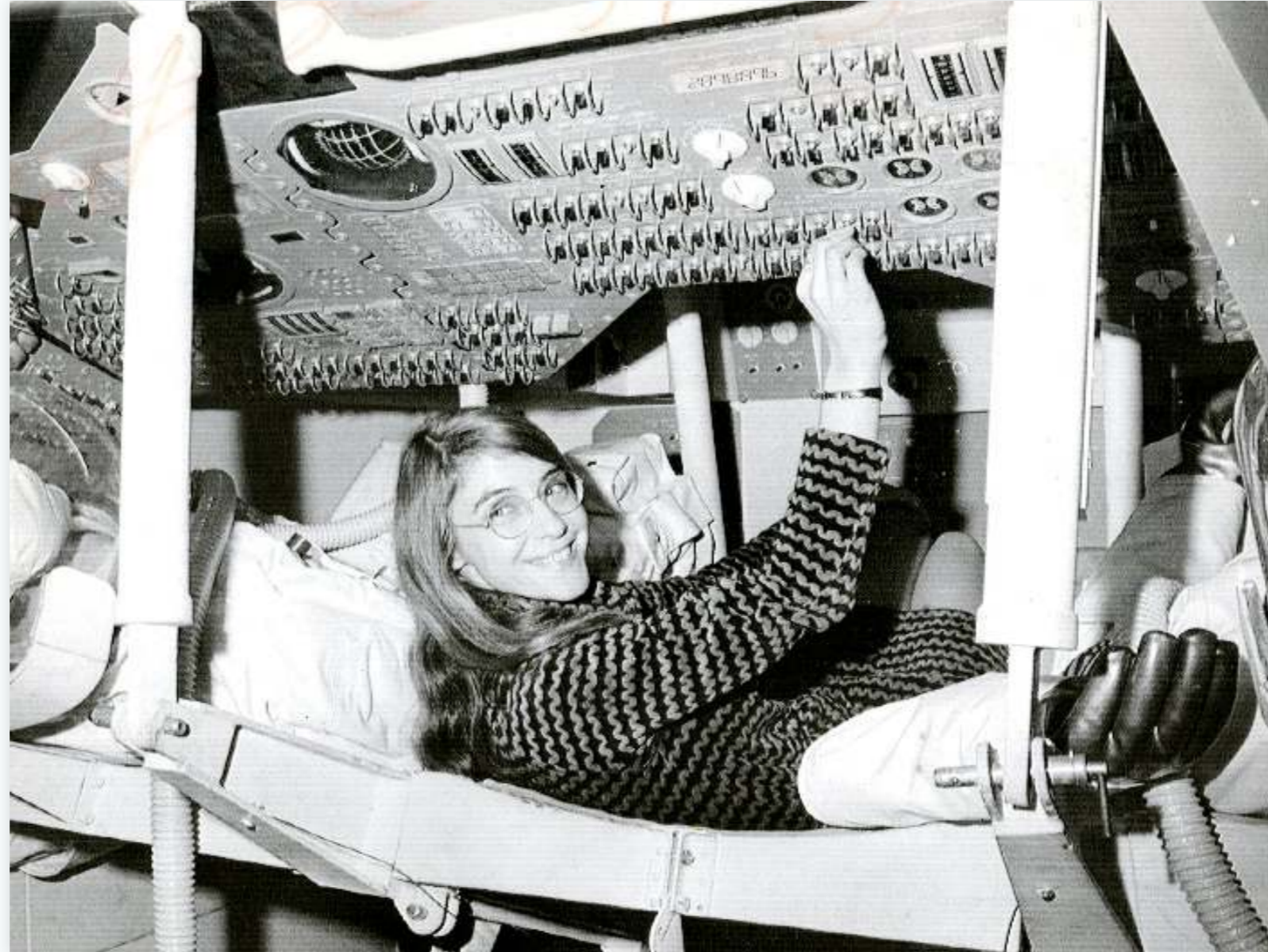
# ROCKET SCIENCE



# APOLLO 11 GUIDANCE COMPUTER

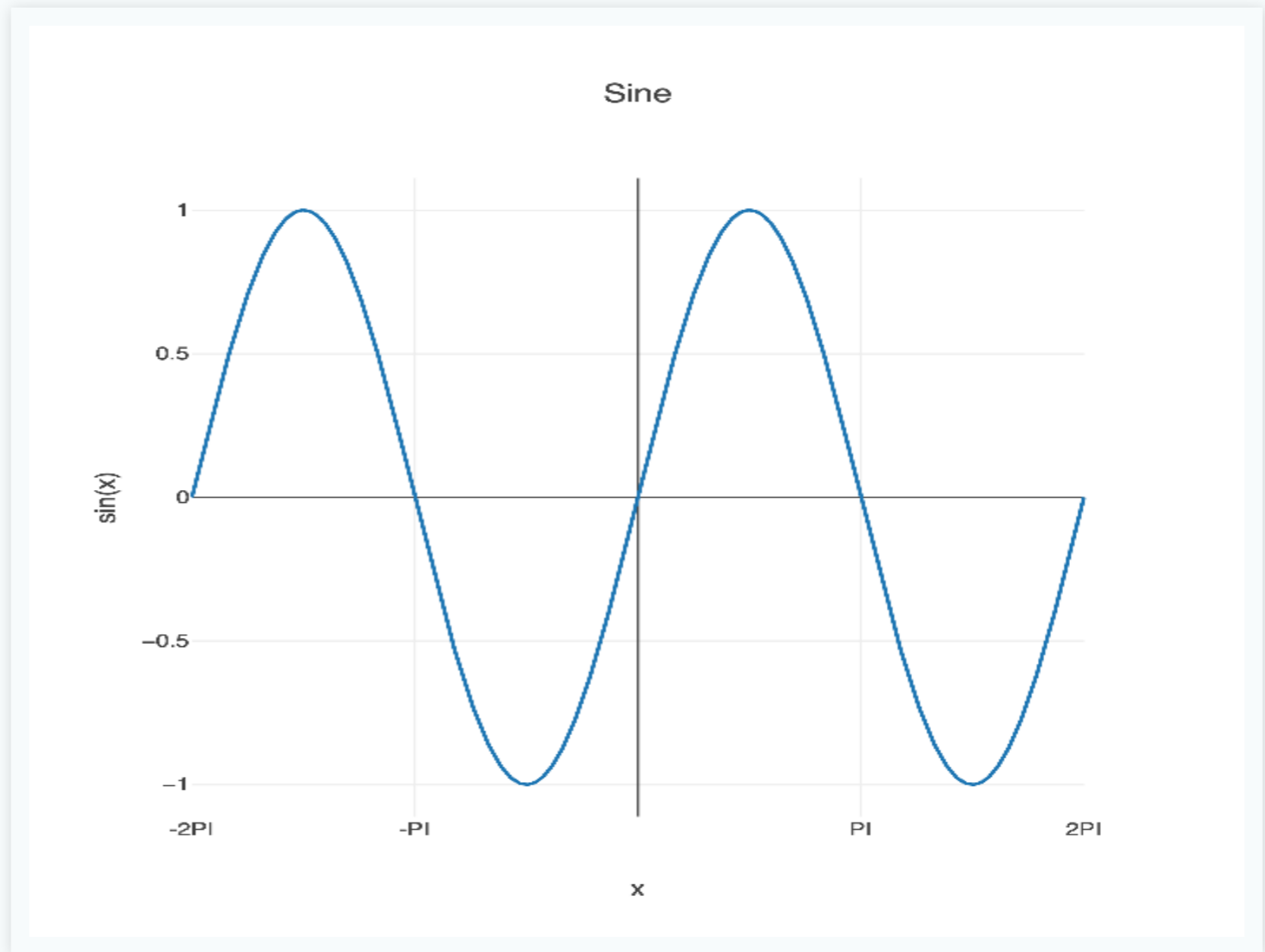
# SINGLE PRECISION SINE AND COSINE			
	COUNT*	\$\$/INTER	
SPCOS	AD	HALF	# ARGUMENTS SCALED AT PI
SPSIN	TS	TEMK	
	TCF	SPT	
	CS	TEMK	
SPT	DOUBLE		
	TS	TEMK	
	TCF	POLLEY	
	XCH	TEMK	
	INDEX	TEMK	
	AD	LIMITS	
	COM		
	AD	TEMK	
	TS	TEMK	
	TCF	POLLEY	
	TCF	ARG90	
POLLEY	EXTEND		
	MP	TEMK	
	TS	SQ	
	EXTEND		
	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.

# MARGARET HAMILTON



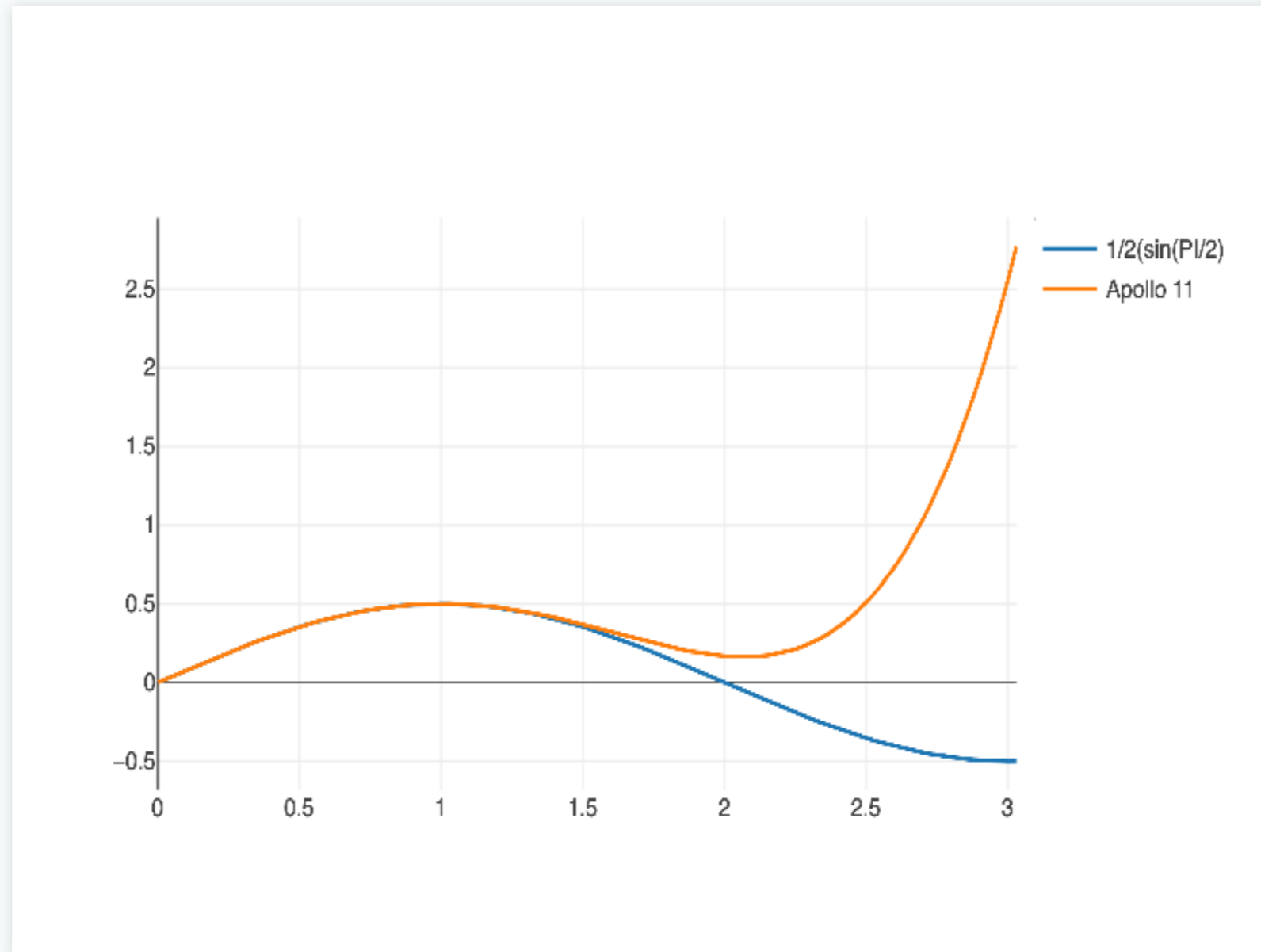
# SINE FUNCTION

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



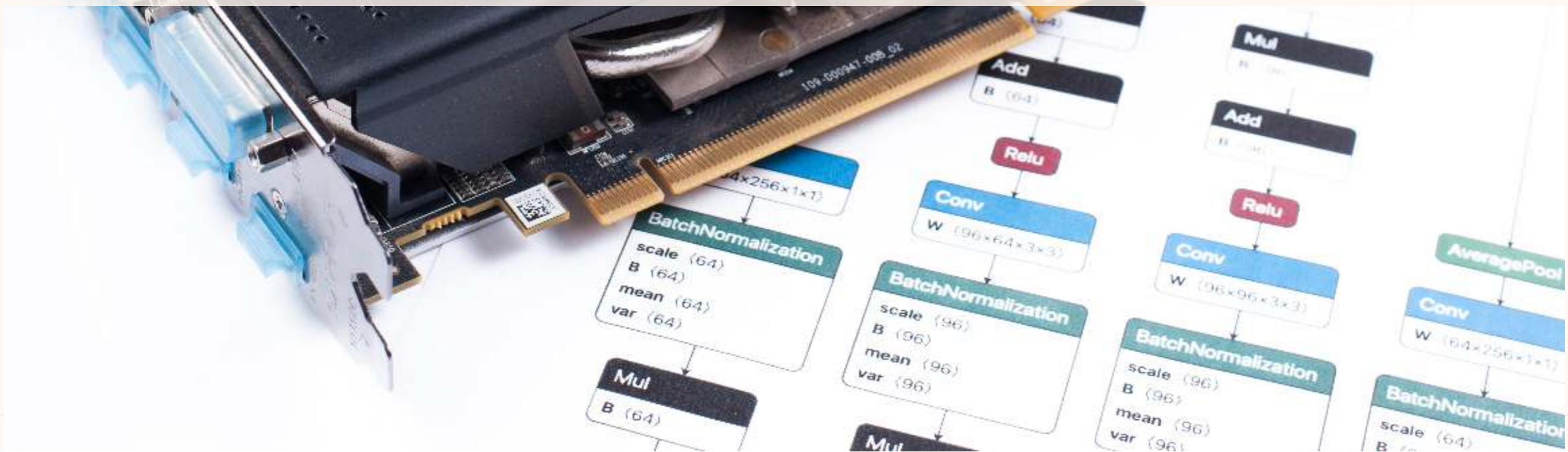
# APOLLO 11 APPROXIMATION

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

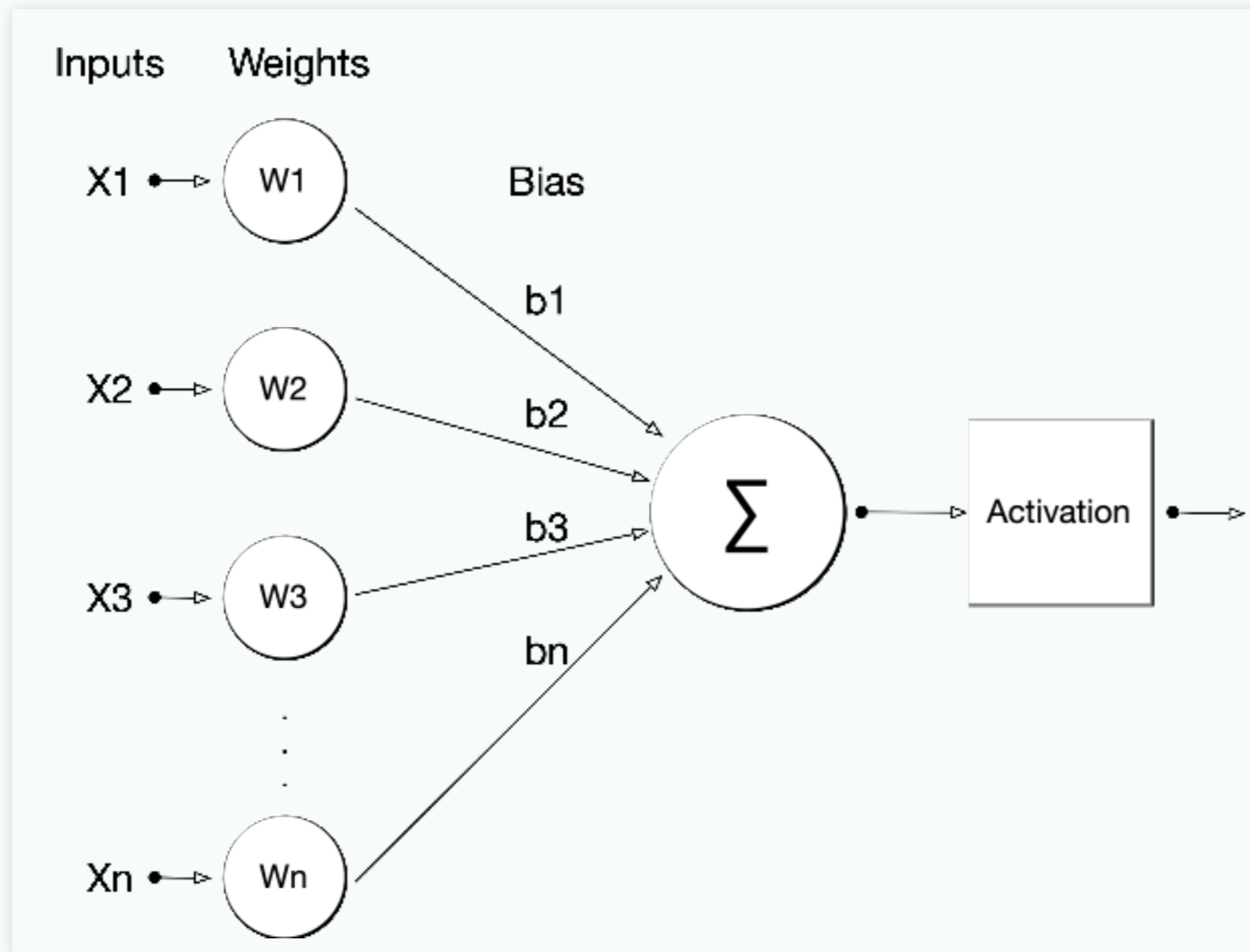




# NEURAL NETWORKS

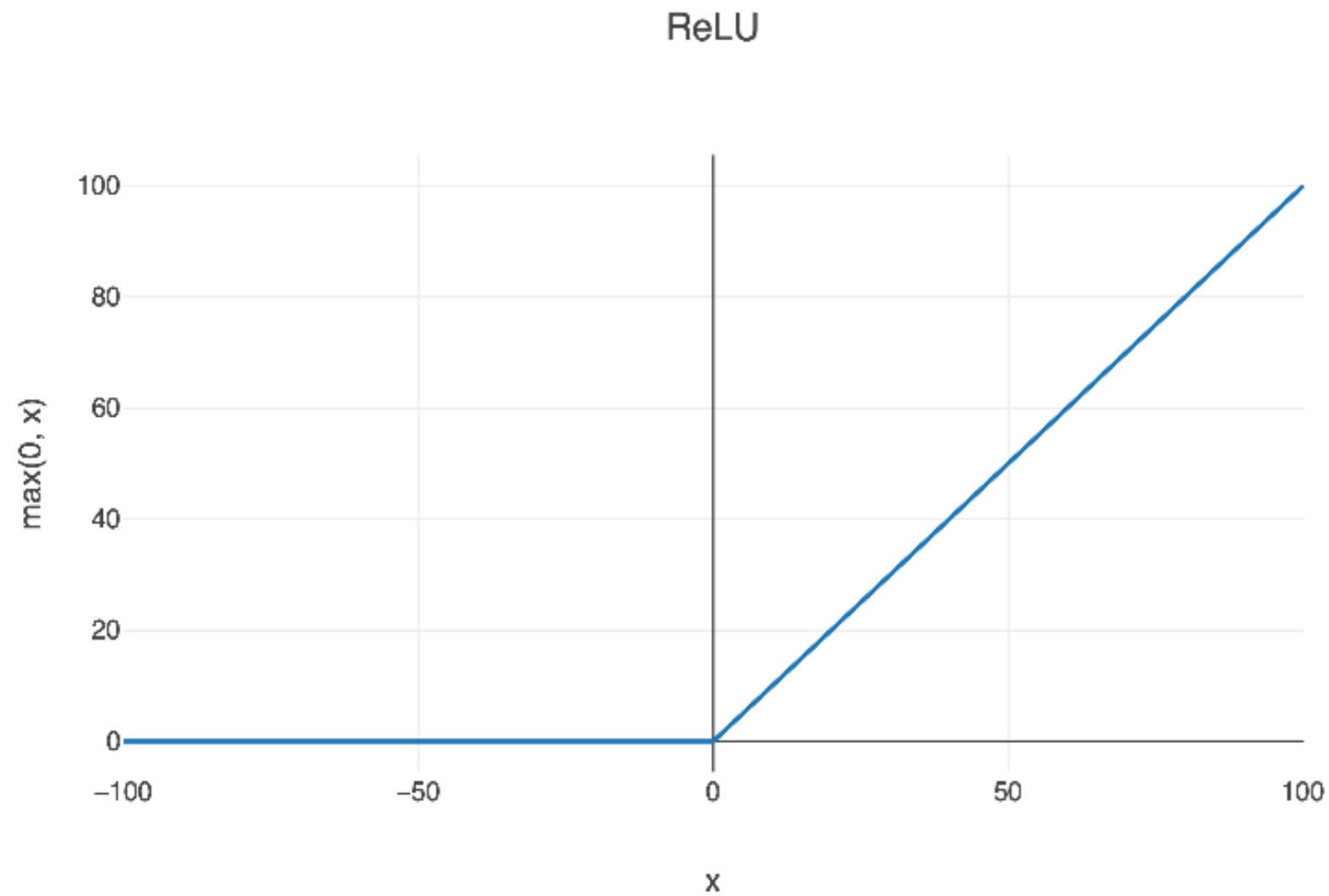


# NEURON



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

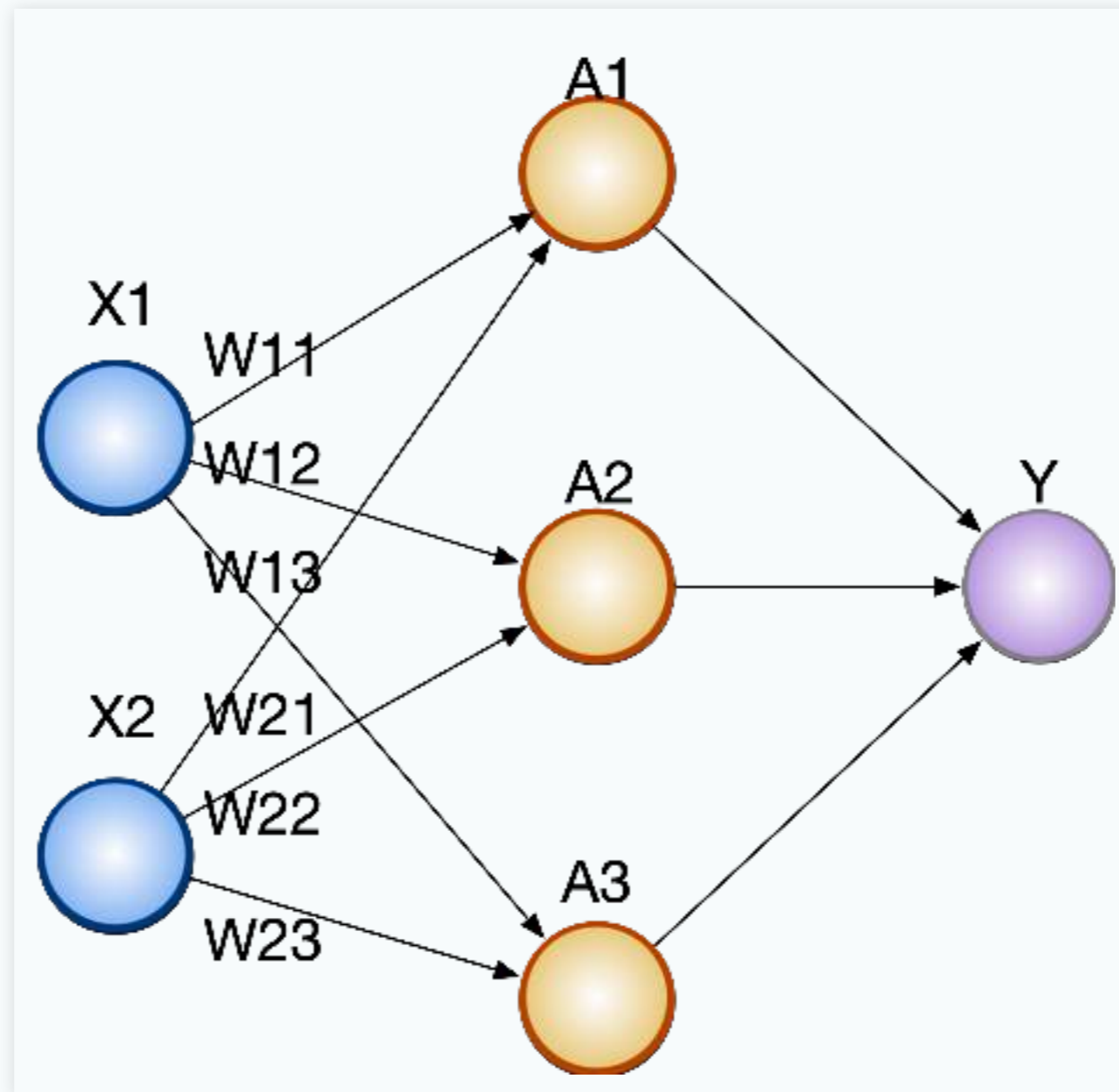
# ACTIVATION FUNCTION RELU



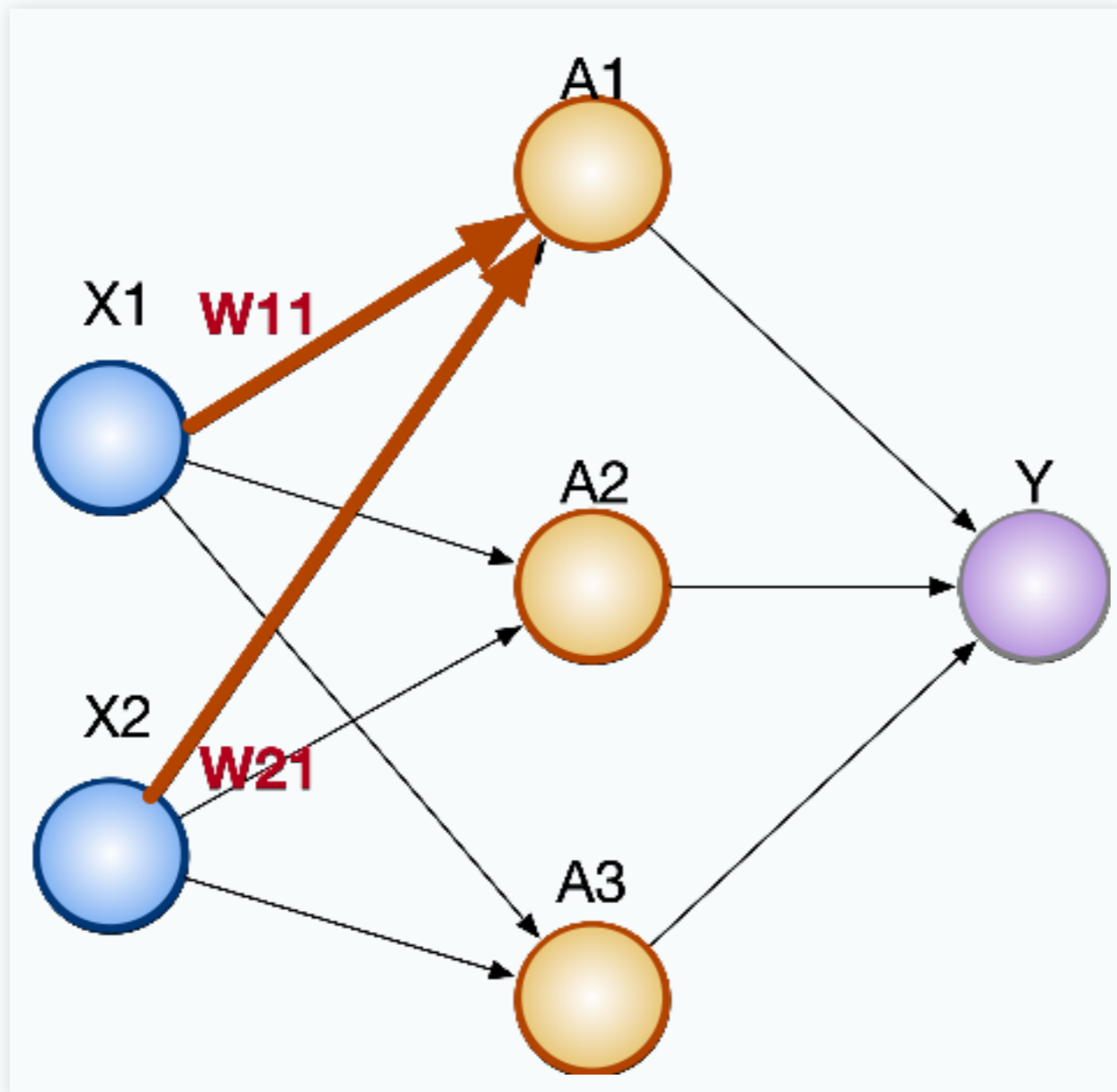
ReLU

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

# NEURAL NETWORK ARCHITECTURE

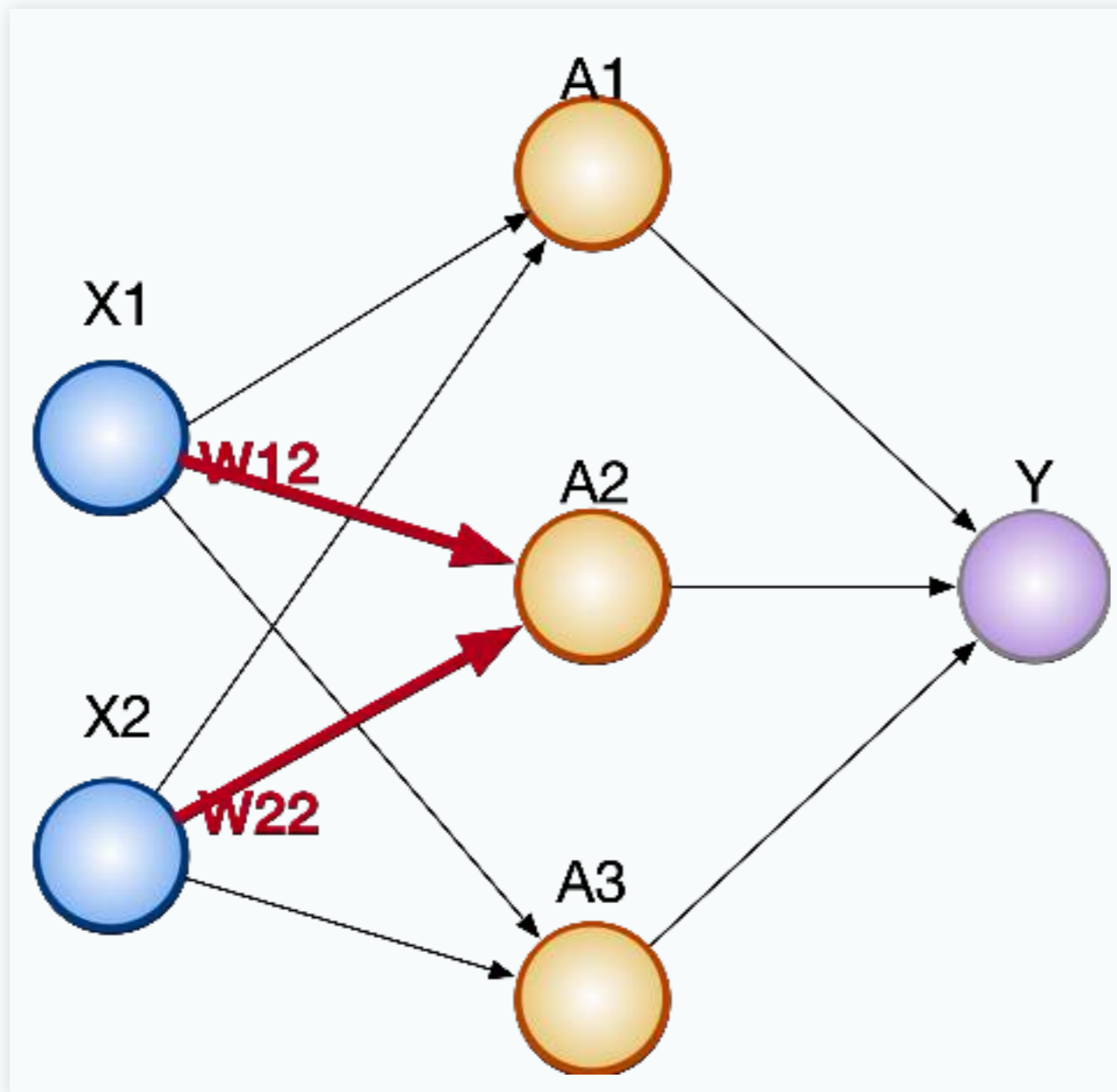


# NEURAL NETWORK ARCHITECTURE



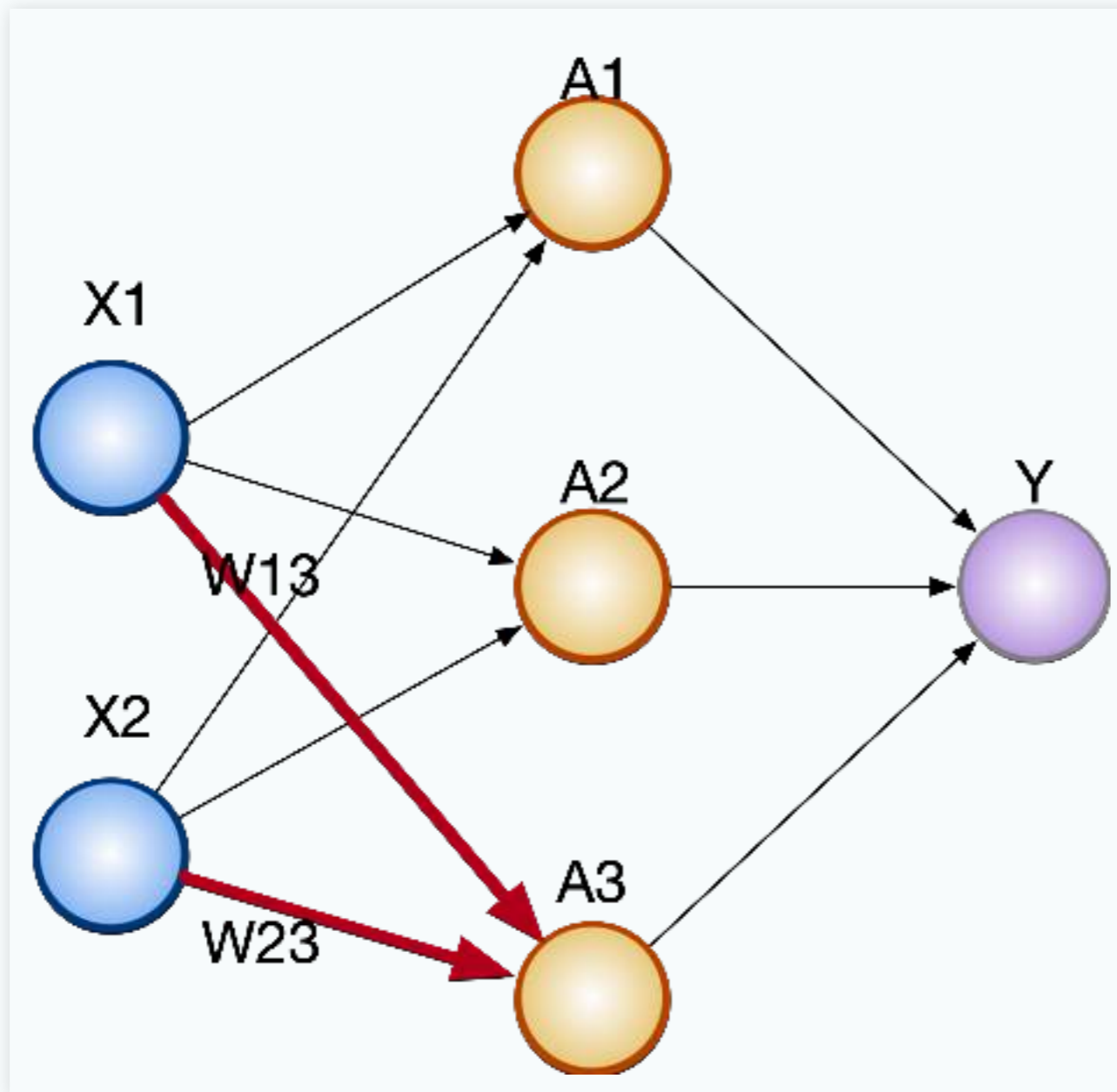
- $A_1 = X_1 W_{11} + X_2 W_{21}$

# 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

$$A = X \cdot W$$

$$X = \begin{bmatrix} \textcolor{blue}{X}_1 & \textcolor{blue}{X}_2 \end{bmatrix}$$

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

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

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

# MATRICES

$$A = X \cdot W$$

$$X = \begin{bmatrix} \textcolor{red}{X}_1 & \textcolor{red}{X}_2 \end{bmatrix}$$

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

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

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

# MATRICES

$$A = X \cdot W$$

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

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

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

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

# 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

# LEARNING TO LEARN



The background of the slide features a composite image. On the left, a silver spiral notebook is partially visible. In the center, a clear fountain pen with a blue nib lies diagonally. On the right, a technical diagram is shown, featuring a box labeled 'BatchNorm' with parameters 'scale <64>', 'B <64>', 'mean <64>', and 'var <64>'. Below it is a 'Mul' box with 'B <64>', and further down is an 'Add' box. Arrows indicate a flow from the BatchNorm box to the Mul box, and then to the Add box.

# LEARNING TO LEARN

- pimp up your brain

The background of the slide features a close-up, slightly blurred image of a spiral-bound notebook with a silver metal spiral binding on the left. A clear, silver-colored fountain pen lies diagonally across the center. To the right, a technical drawing or flowchart is visible, showing a box labeled 'BatchNo' with fields for 'scale <64>', 'B <64>', 'mean <64>', and 'var <64>'. Below this, an arrow points to a box labeled 'Mul' containing 'B <64>', which in turn points to a box labeled 'Add'.

# LEARNING TO LEARN

- pimp up your brain
- healthy lifestyle

The background of the slide features a close-up, slightly blurred image of a spiral-bound notebook on the left. A silver pen lies diagonally across the center. To the right, a document is visible, showing a flowchart with boxes labeled 'BatchNo', 'Mul', and 'Add'. The 'BatchNo' box contains the text 'scale <64>', 'B <64>', 'mean <64>', and 'var <64>'. Arrows indicate a flow from 'BatchNo' to 'Mul', and from 'Mul' to 'Add'.

# LEARNING TO LEARN

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

The background of the slide features a close-up, slightly blurred image of a spiral-bound notebook with a silver metal spiral. A blue pen lies diagonally across the notebook. In the lower right corner, a portion of a document is visible, showing a flowchart with rectangular boxes. One box is labeled 'BatchNo' and contains the text 'scale <64>', 'B <64>', 'mean <64>', and 'var <64>'. Below it, an arrow points to a box labeled 'Mul' containing 'B <64>'. Another arrow points from 'Mul' to a box labeled 'Add'.

# LEARNING TO LEARN

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



# DATA SCIENCE FOR ANDROID DEVELOPER







+



=



# DATA SCIENCE AND KOTLIN

# 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

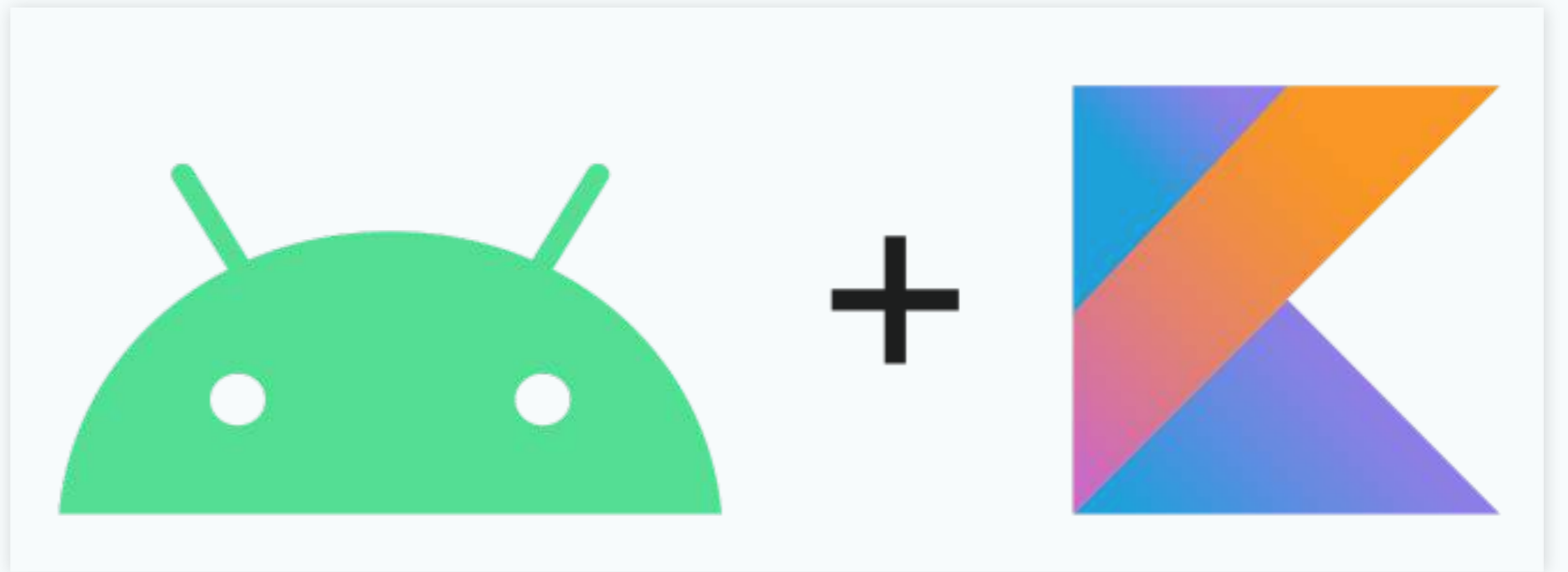
# DATA SCIENCE AND KOTLIN

- JetBrains and data science
- kscript
- kmath, kotlingrad
- kraml, plotly.kt

# DATA SCIENCE AND KOTLIN

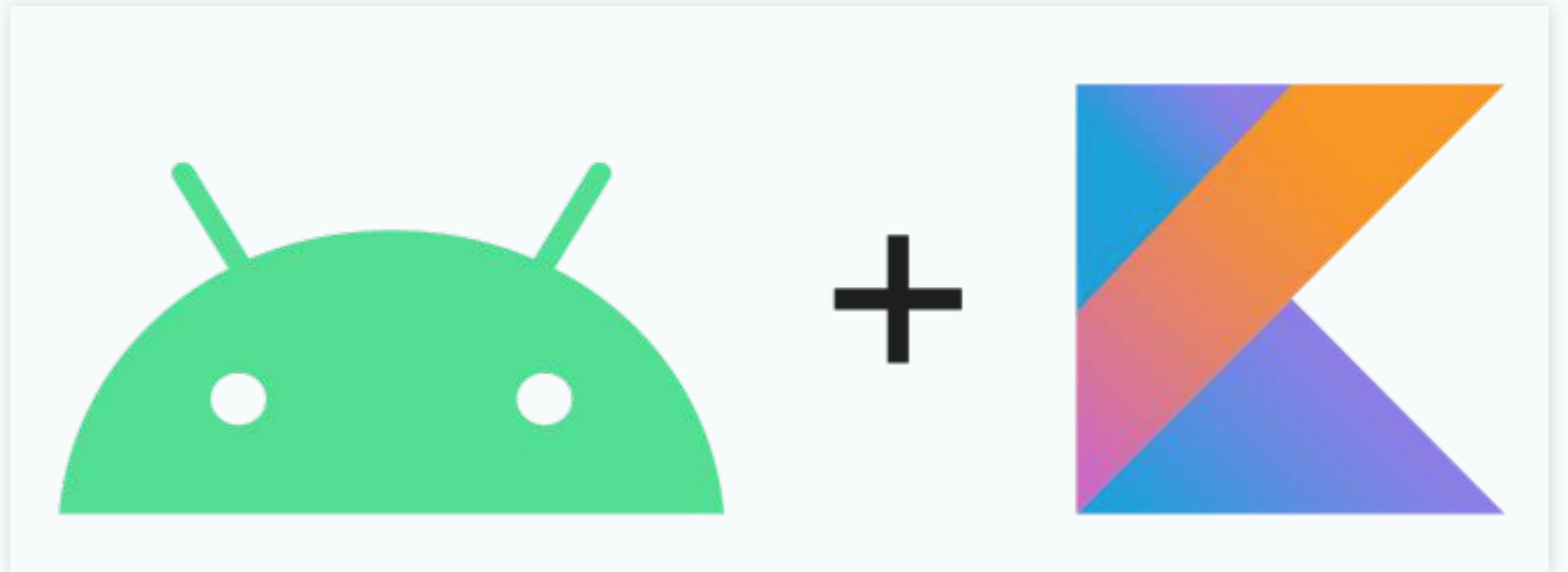
- JetBrains and data science
- kscript
- kmath, kotlingrad
- kraml, plotly.kt
- KotlinDL

# MACHINE LEARNING ON ANDROID



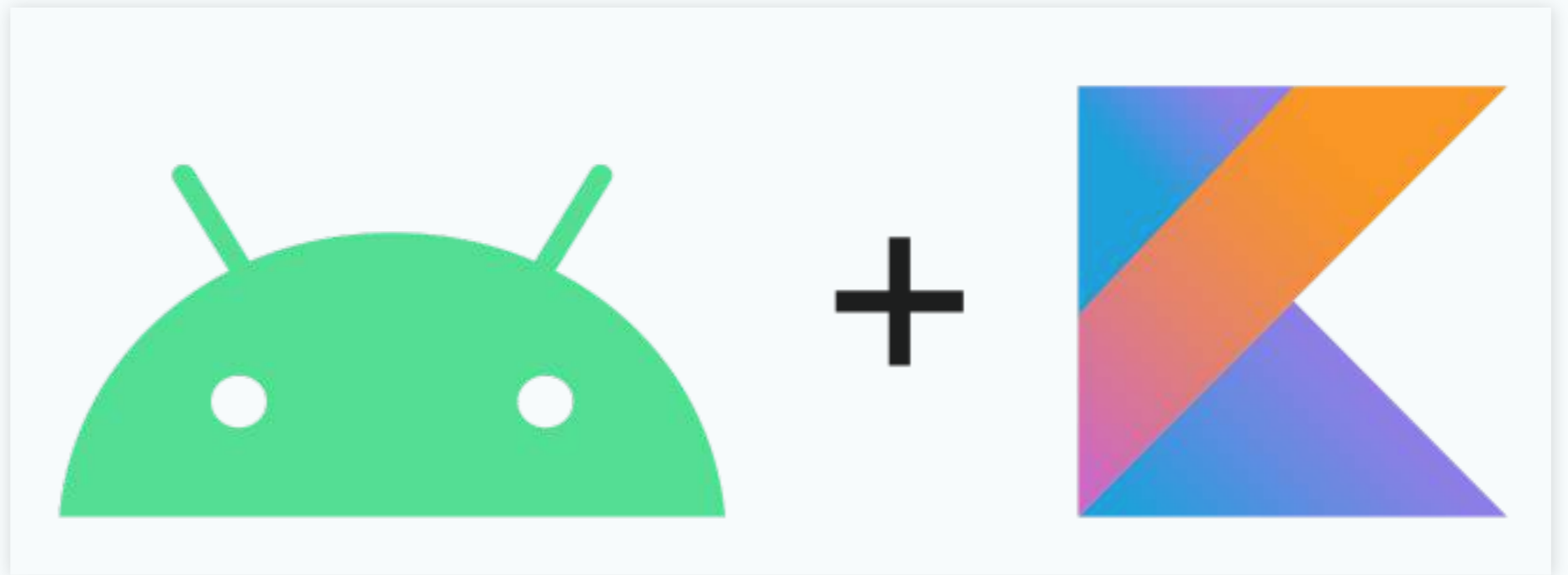
# MACHINE LEARNING ON ANDROID

- Kotlin & JVM



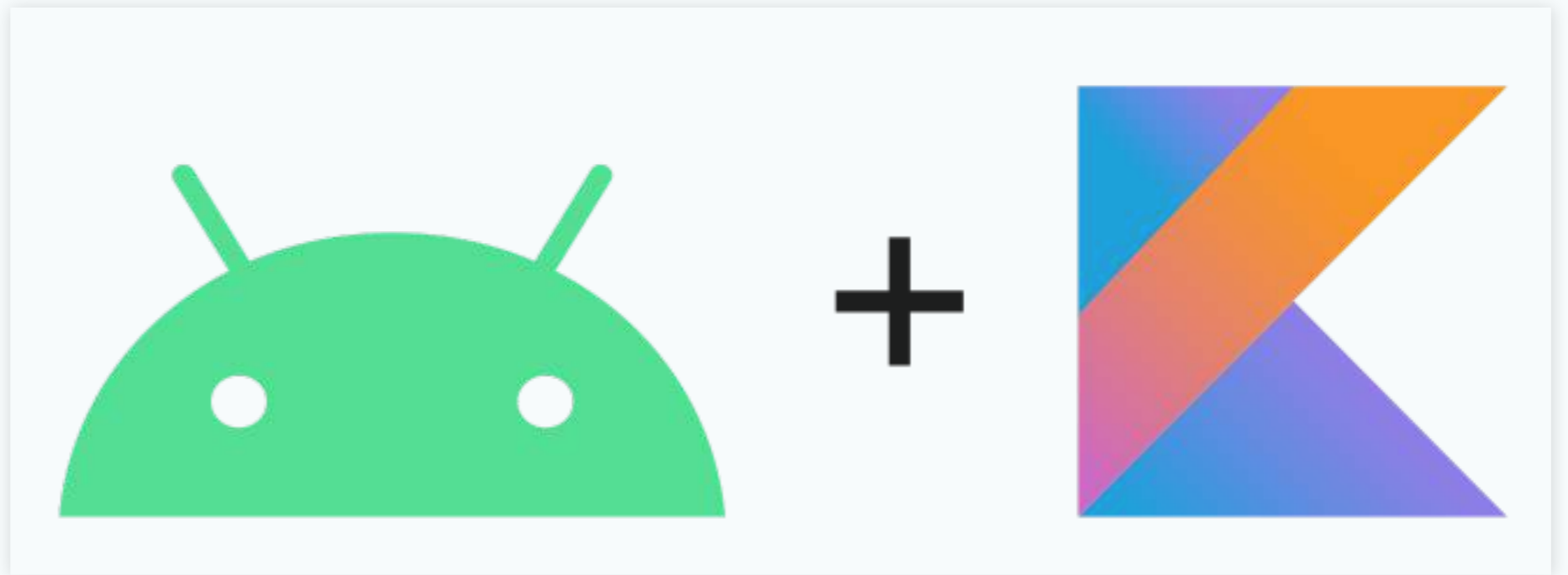
# MACHINE LEARNING ON ANDROID

- Kotlin & JVM
- machine learning on the edge



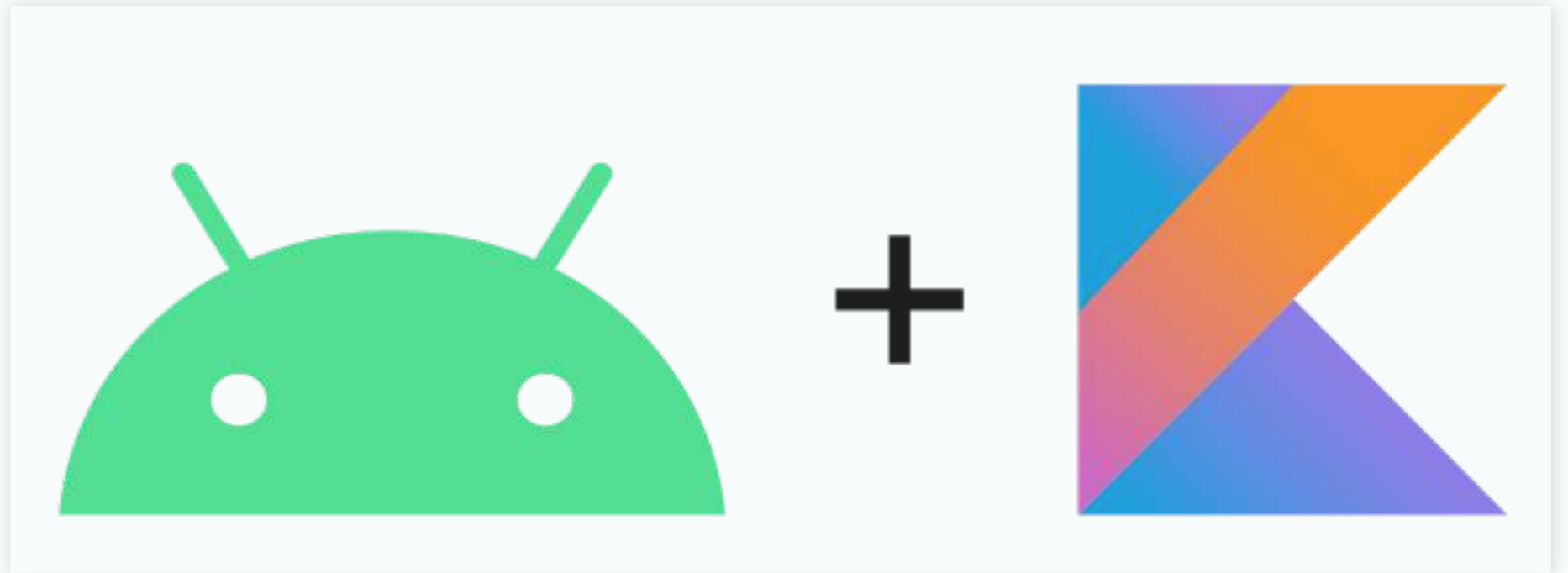
# MACHINE LEARNING ON ANDROID

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



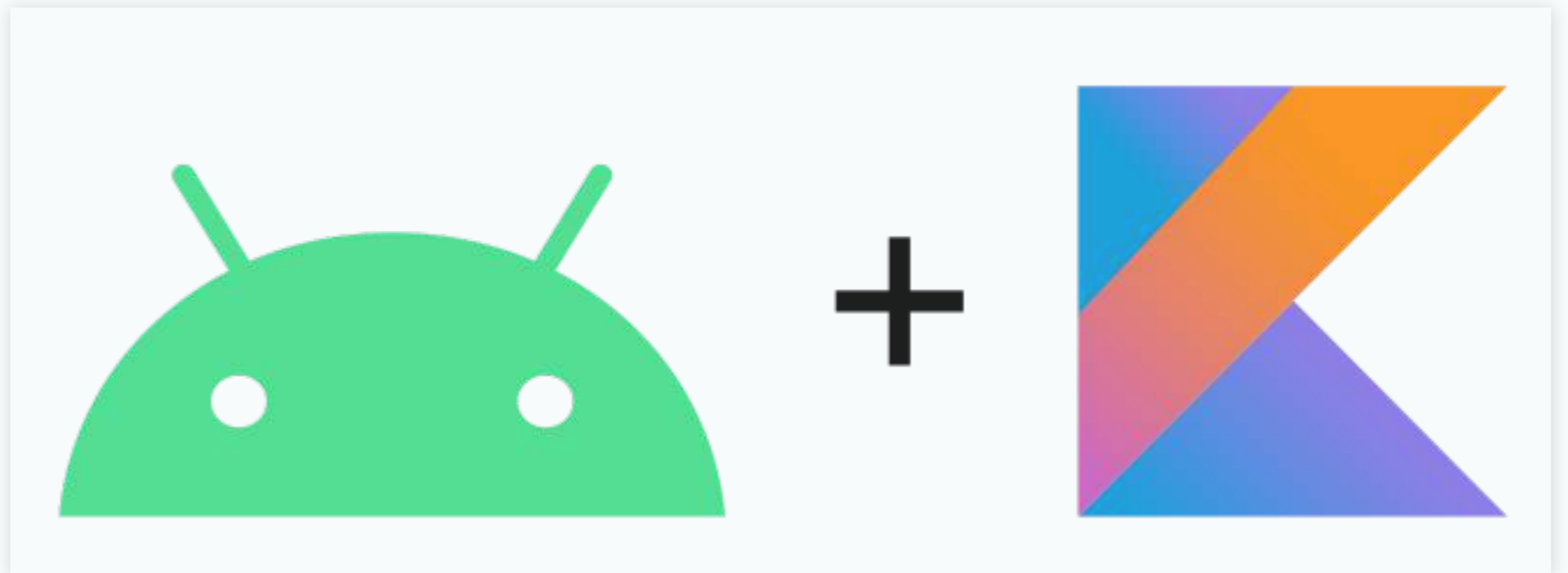
# MACHINE LEARNING ON ANDROID

- 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





THANK YOU



# SLIDES

- slides
- code
- notebook



# 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:Date:Margaret\\_Hamilton\\_in\\_action.jpg](https://de.wikipedia.org/wiki/Margaret_Hamilton_(Wissenschaftlerin)#/media:Date: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