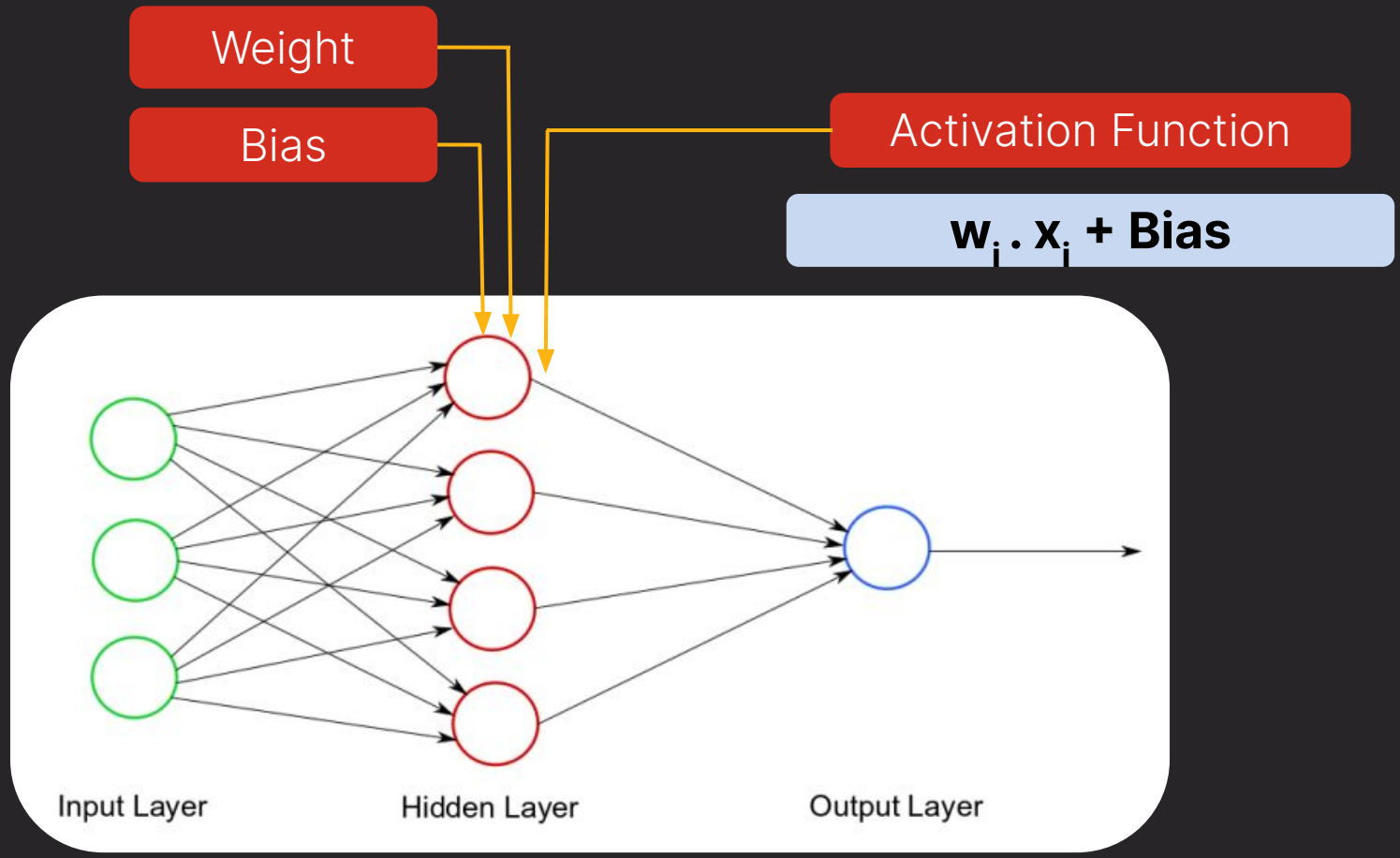
Understanding workings of Neural Networks

Video 1: How are Neural Networks trained: Forward Propagation

Analytics Vidhya

Introduction

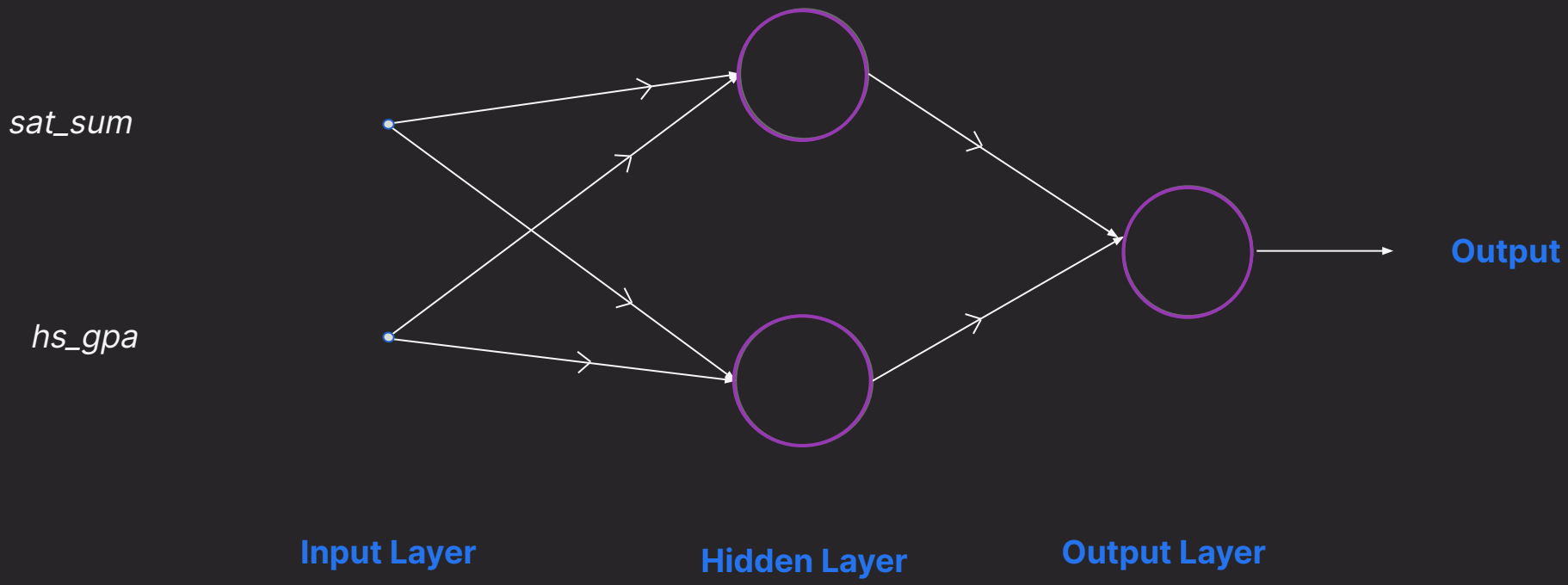


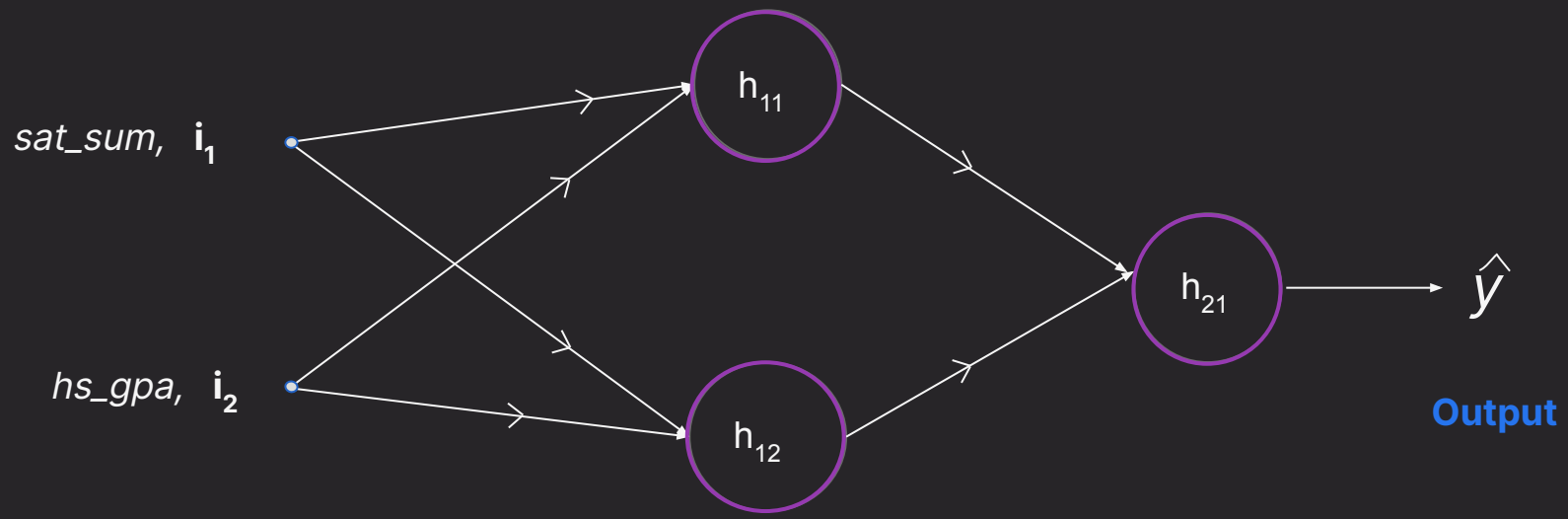
Analytics Vidhya

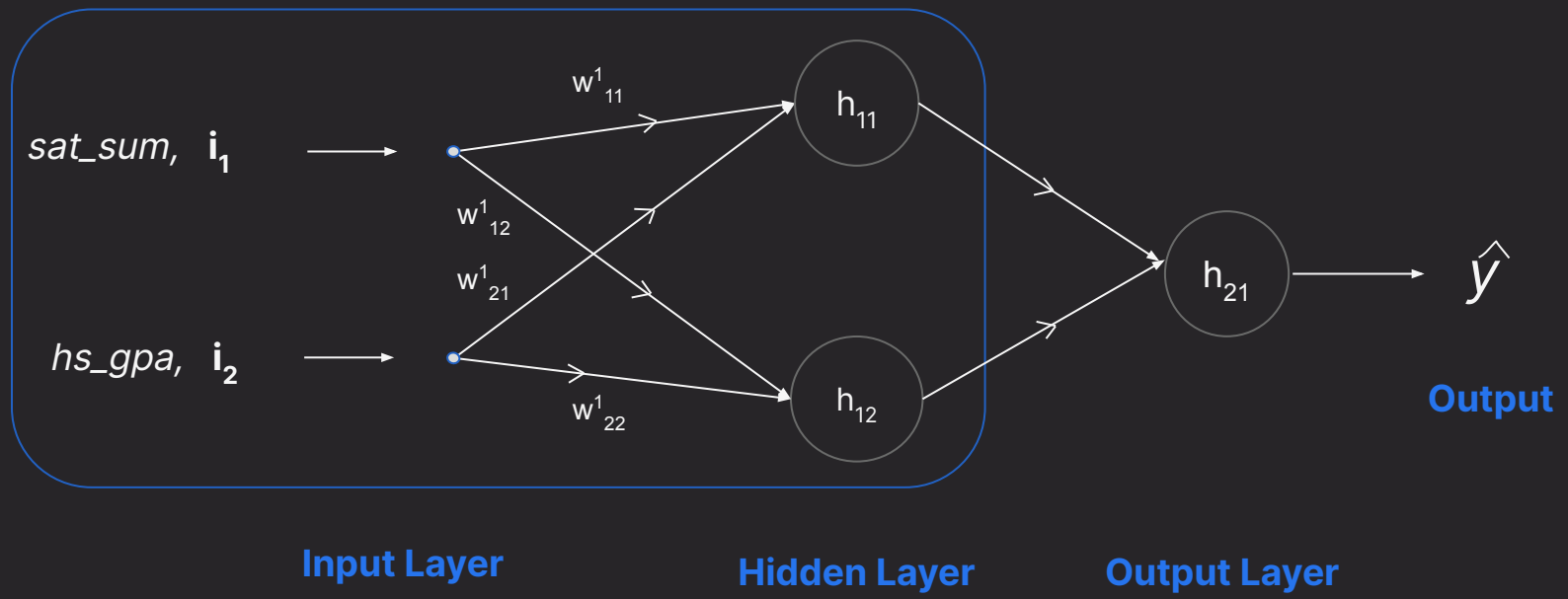
Predict 1st year college grades from high school SAT and GPA scores.

sat_sum	hs_gpa	fy_gpa
727	3.40	3.18
722	4.00	3.33
716	3.75	3.25

← Target


sat_sum
hs_gpa
Output
Input Layer
Hidden Layer
Output Layer


sat_sum, i_1
hs_gpa, i_2
h_{11}
h_{12}
h_{21}
$\hat{y}$
Output
Input Layer
Hidden Layer
Output Layer

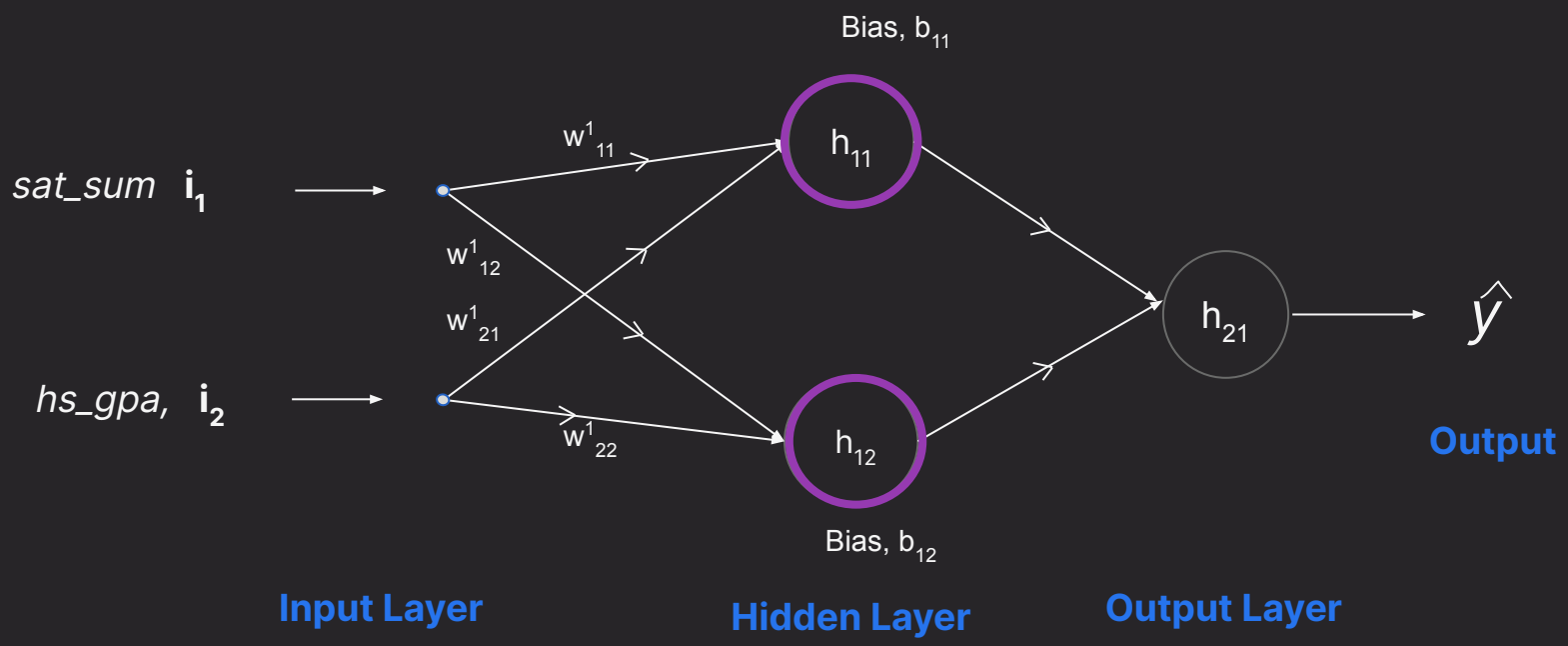


Nodes in Input and Hidden layer = 2 Each

Connections = 2*2 = 4

4 Weights

w^1_{11} , w^1_{12}, w^1_{21}, w^1_{22}



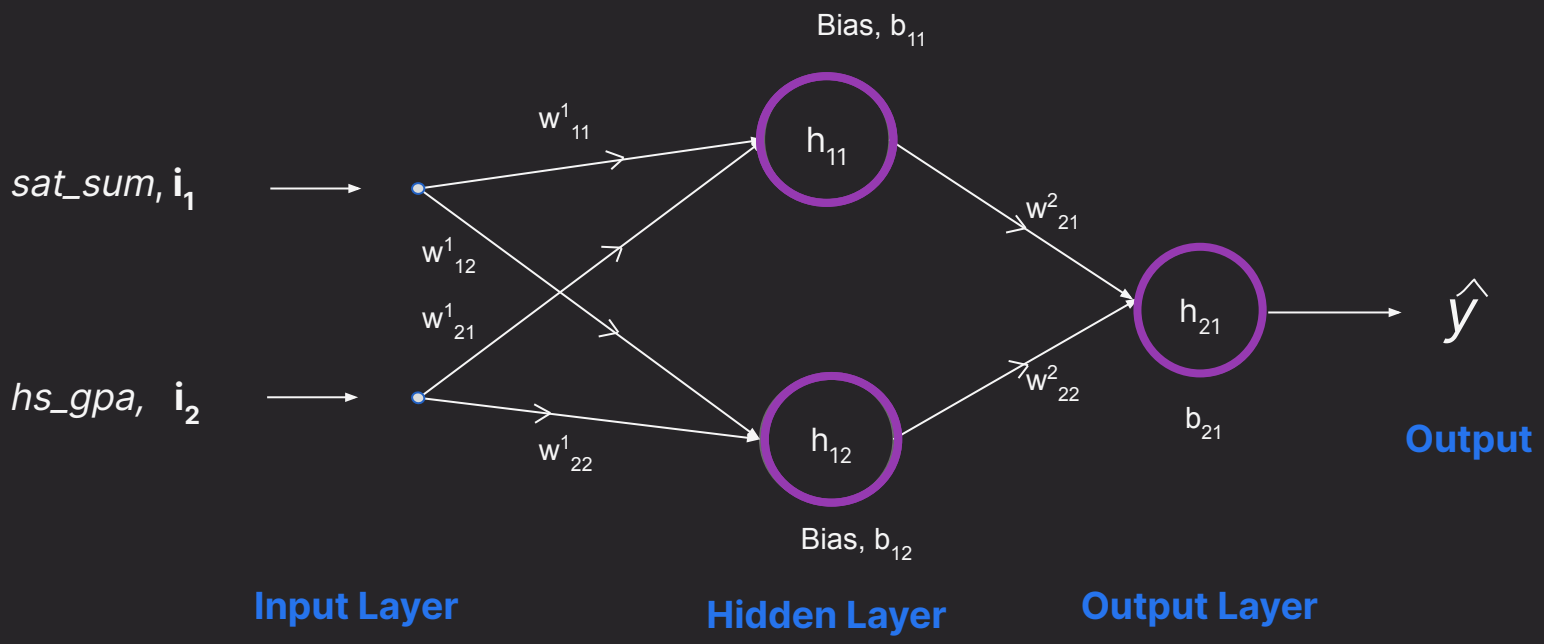
Connections = 2*2 = 4

Nodes = 2

Biases= 4 + 2

6 Dependent Parameters

$w^1_{11}, w^1_{12}, w^1_{21}, w^1_{22}, b_{11}, b_{12}$

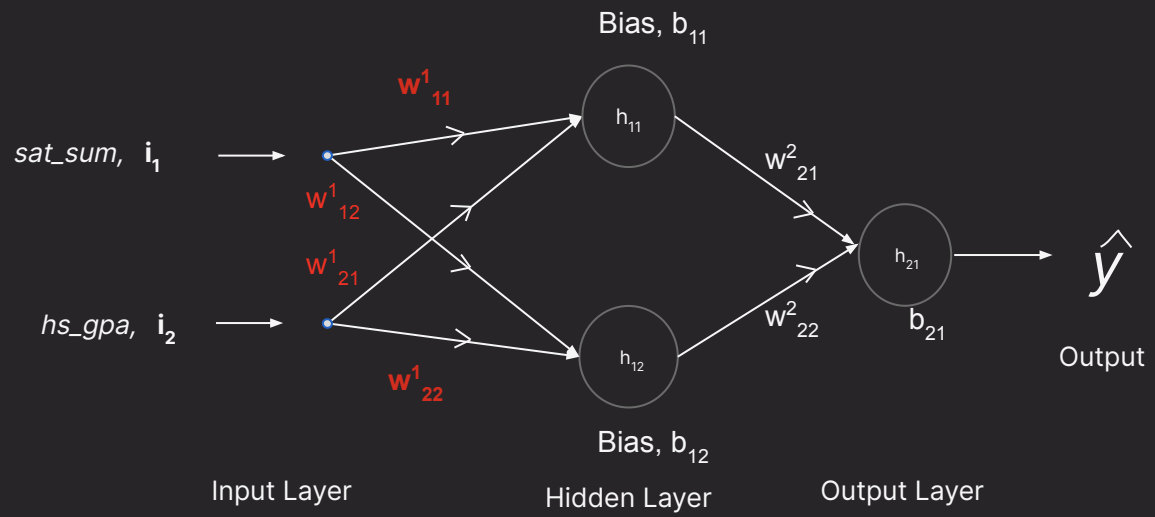


6 Dependent Parameters
+
3 New Parameters
=
9 Dependent Parameters

migration

Analytics Vidhya

1. Calculate the weighted sum of the inputs

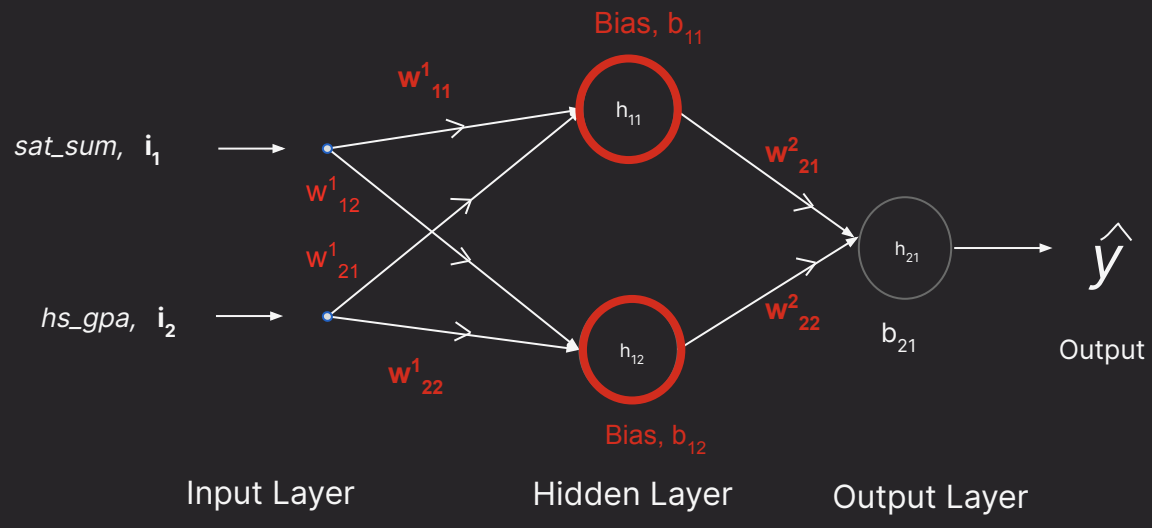


$$\mathbf{I} = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix}$$

$$\mathbf{W_1} = \begin{bmatrix} w^1_{11} \\ w^1_{12} \end{bmatrix} \qquad \mathbf{W_2} = \begin{bmatrix} w^1_{21} \\ w^1_{22} \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix}$$

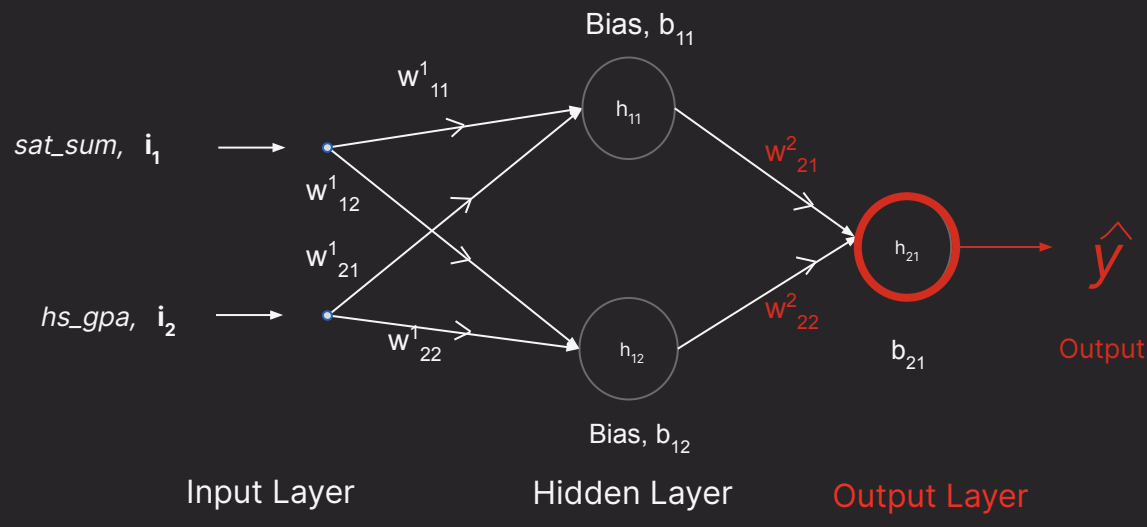
2. Apply activation function to get outputs



$$O_1 = \text{Sigmoid}\,(w^1_{11} * i_1 + w^1_{21} * i_2 + b_{11})$$

$$O_2 = \text{Sigmoid}\,(w^1_{12} * i_1 + w^1_{22} * i_2 + b_{12})$$

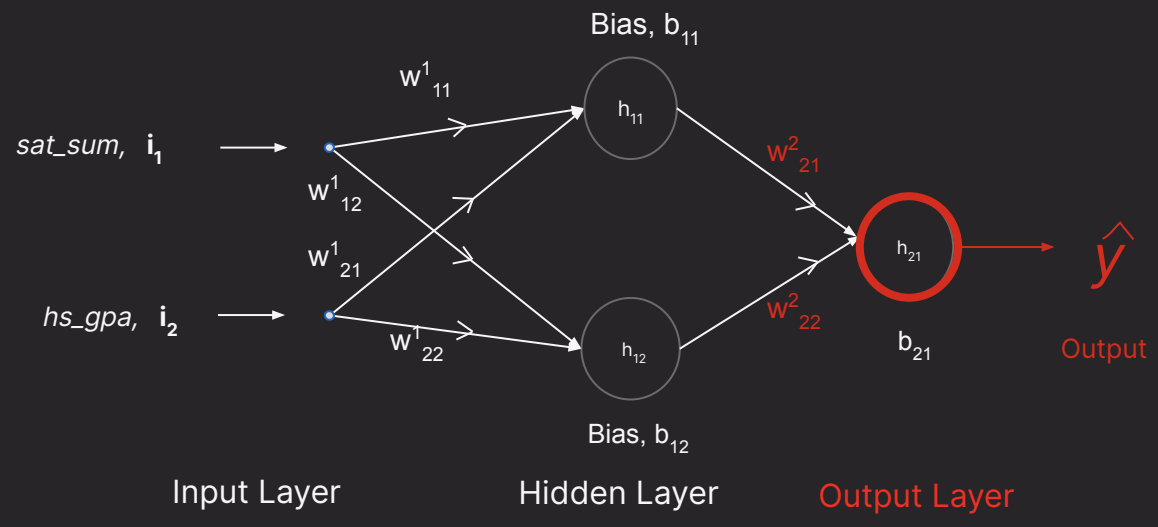
3. Outputs act as input for the next layer



$$O_1 = \text{Sigmoid}\,(w^1_{11} * i_1 + w^1_{21} * i_2 + b_{11})$$

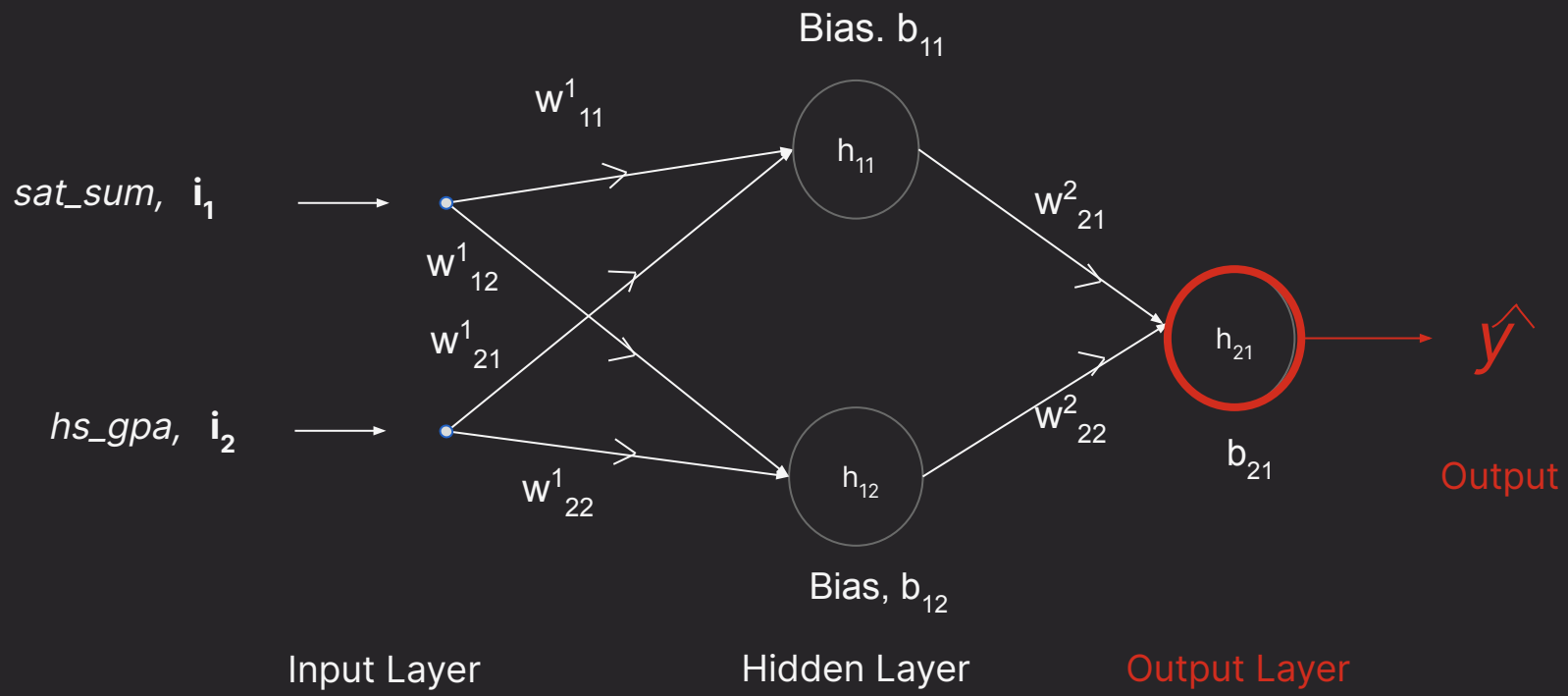
$$O_2 = \text{Sigmoid}\,(w^1_{12} * i_1 + w^1_{22} * i_2 + b_{12})$$

4. Calculate weighted sum of outputs



$$W^2_{21} * O_1 + W^2_{22} * O_2 + b_{21}$$

5. Apply Activation Function to get Output



$$Y = \text{Activation Function}\,(W^2_{21} * O_1 + W^2_{22} * O_2 + b_{21})$$

Analytics Vidhya

Activation Function of Last Layer : Regression Problems

$$Y = \textbf{Activation Function}\ (W^2_{21} * O_1 + W^2_{22} * O_2 + b_{21})$$

Linear Activation Function

✓ Regression Problems with negative output

Example: Temperature Prediction

ReLu Activation Function

✓ Regression Problems with positive output

Example: House price Prediction

Analytics Vidhya

Activation Function of Last Layer : Classification Problems

$$Y = \textbf{Activation Function}\ (W^2_{21} * O_1 + W^2_{22} * O_2 + b_{21})$$

Sigmoid Function

✓ Binary Classification Problems

Softmax Function

✓ Multi Classification Problems