

# A Framework for Selecting **DEEP LEARNING** Hyper-Parameters

British International Conference on Databases

Jim O' Donoghue

7<sup>th</sup> July 2015

# TABLE OF CONTENTS

Background + Motivation

Algorithms + The CDN

Experiments + Results

Future Work

Conclusions



NEED TO FIX NUMBERS



INTRODUCTION

# BACKGROUND



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



INTRODUCTION

# BACKGROUND **IN-MINDD**



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



INTRODUCTION

# BACKGROUND **IN-MINDD**

## Dementia Awareness + Prevention



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



INTRODUCTION

# BACKGROUND **IN-MINDD**

Dementia Awareness + Prevention

## Online Environment



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



INTRODUCTION

# BACKGROUND **IN-MINDD**

Dementia Awareness + Prevention

Online Environment

## Risk Prediction Algorithm



INTRODUCTION

# BACKGROUND **IN-MINDD**

Dementia Awareness + Prevention

Online Environment

## Risk Prediction Algorithm - Validation



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979





# INTRODUCTION **MOTIVATION**



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# INTRODUCTION MOTIVATION

# DATA



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# INTRODUCTION MOTIVATION

# DATA

# High-Dimensional



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# INTRODUCTION **MOTIVATION**

# DATA

High-Dimensional

**Variable  
Interactions**



# INTRODUCTION **MOTIVATION**

**DATA**

High-Dimensional

Variable  
Interactions

# DEEP LEARNING



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# INTRODUCTION **MOTIVATION**

**DATA**

High-Dimensional

Variable  
Interactions

**DEEP LEARNING**

**Hyper-Parameter  
Selection**



# INTRODUCTION MOTIVATION

DATA

High-Dimensional

Variable

# FRAMEWORK

Hyper-Parameter Selection

DEEP LEARNING



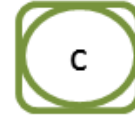




## ALGORITHM OVERVIEW

## DEEP LEARNING

Visible  
Output Layer



C

— Class

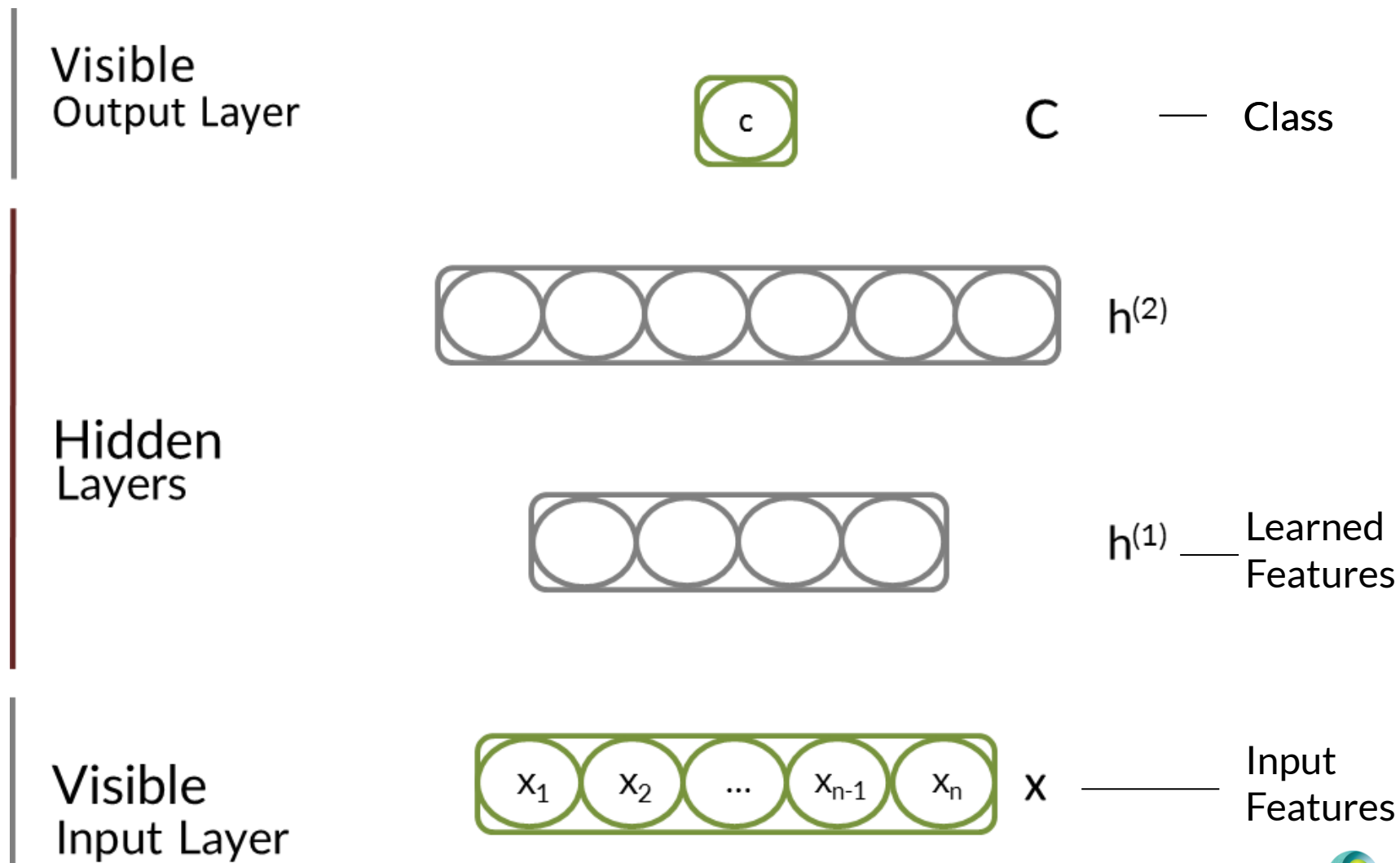
Visible  
Input Layer



X

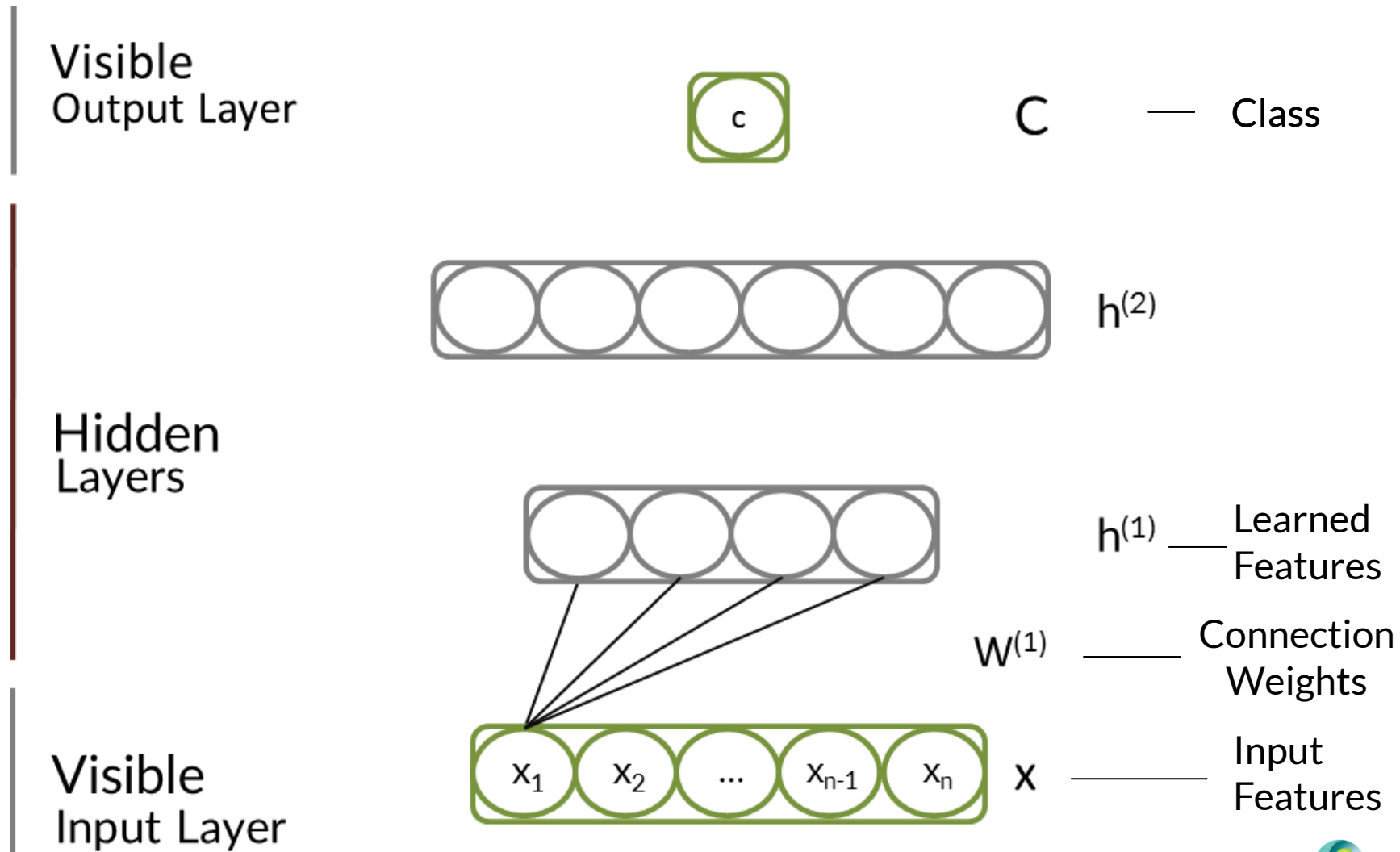
Input  
Features

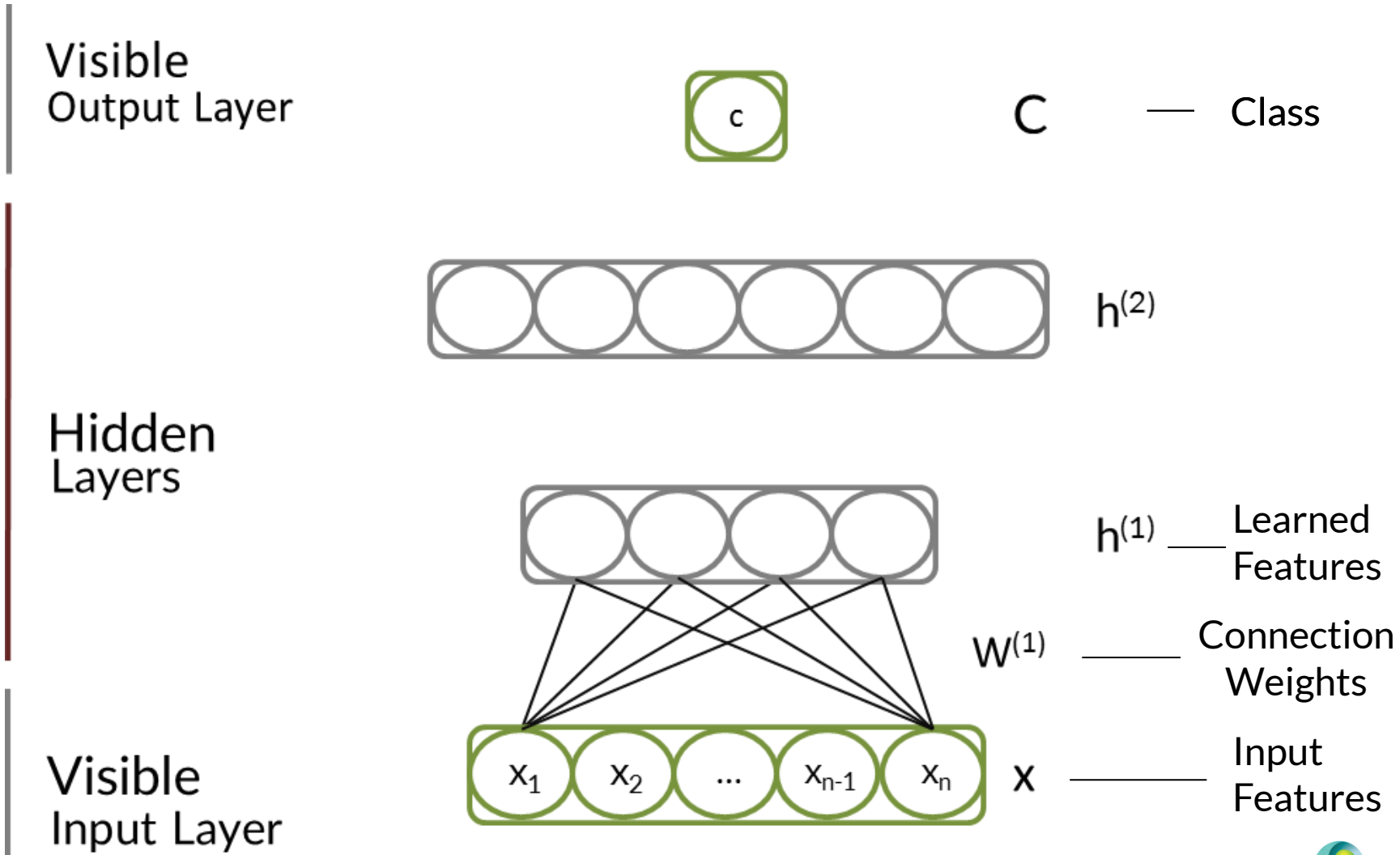


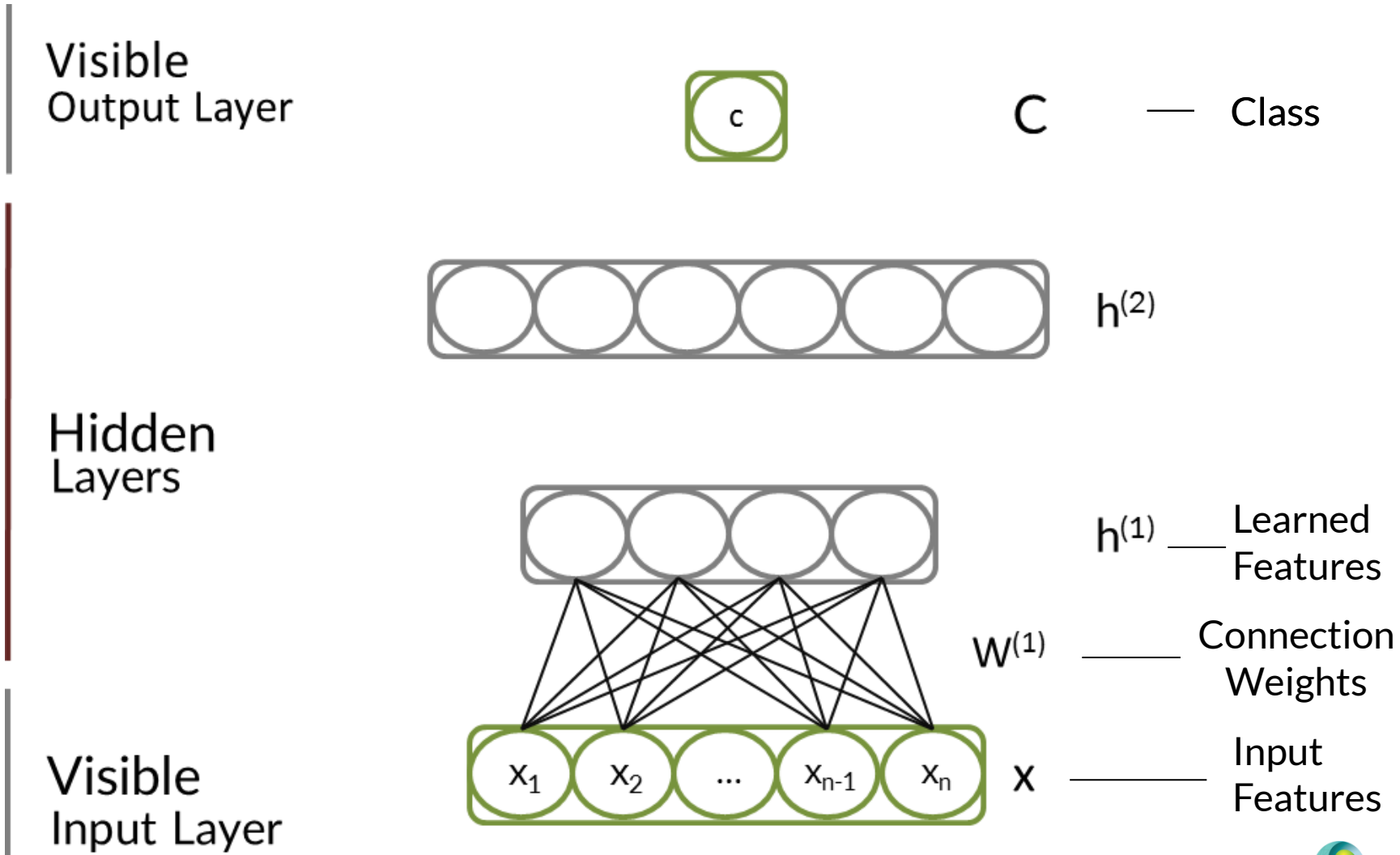


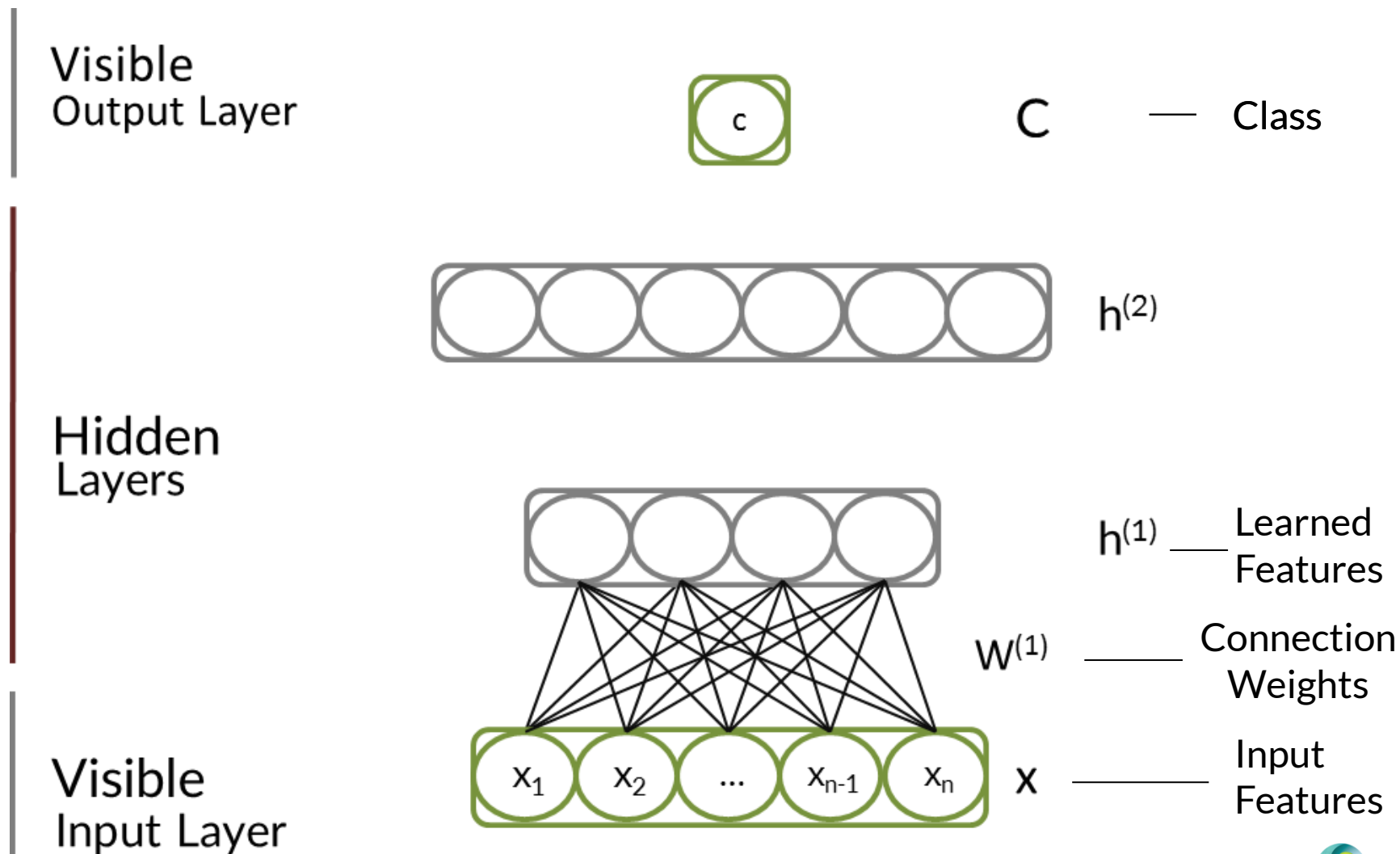
## ALGORITHM OVERVIEW

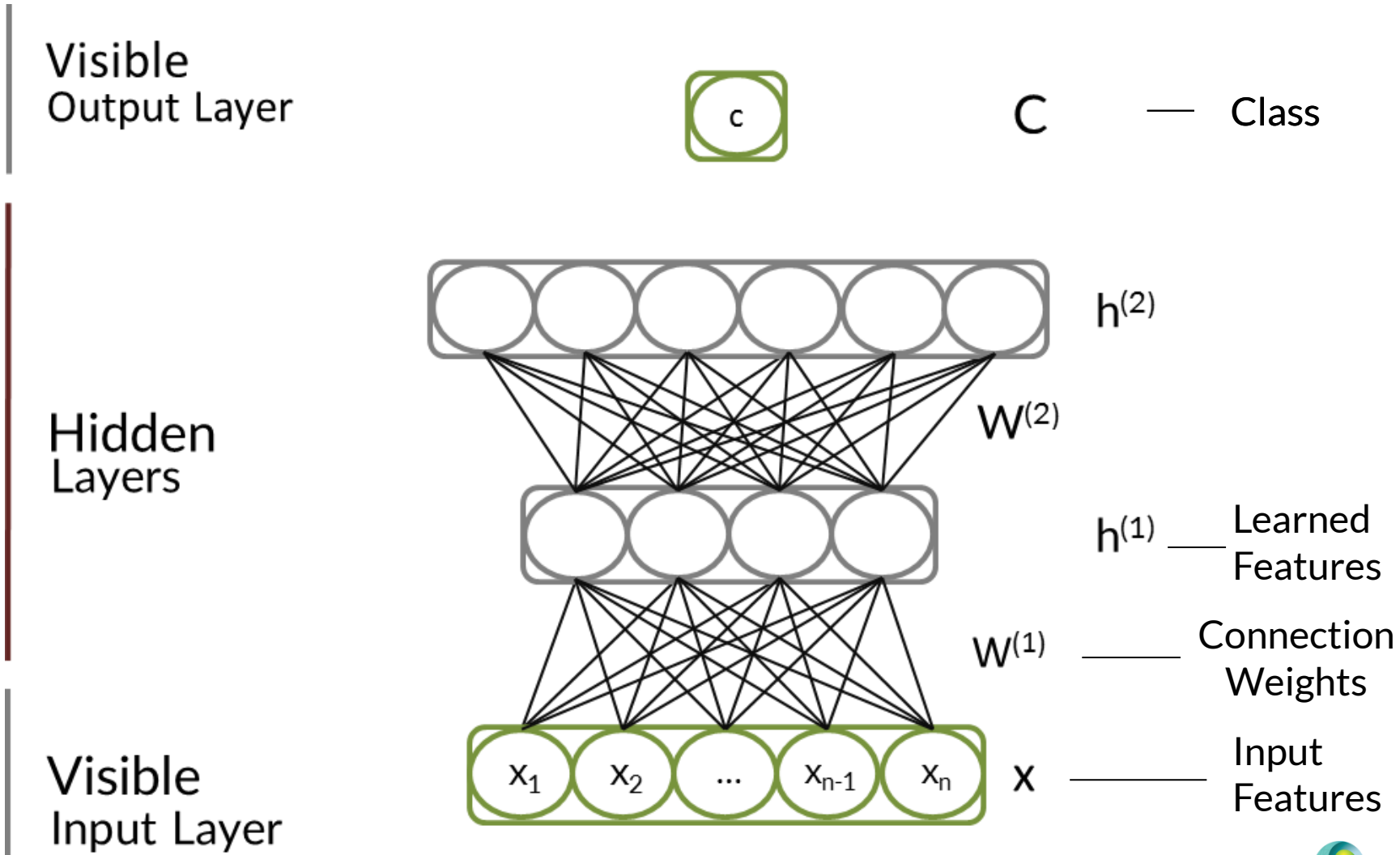
## DEEP LEARNING

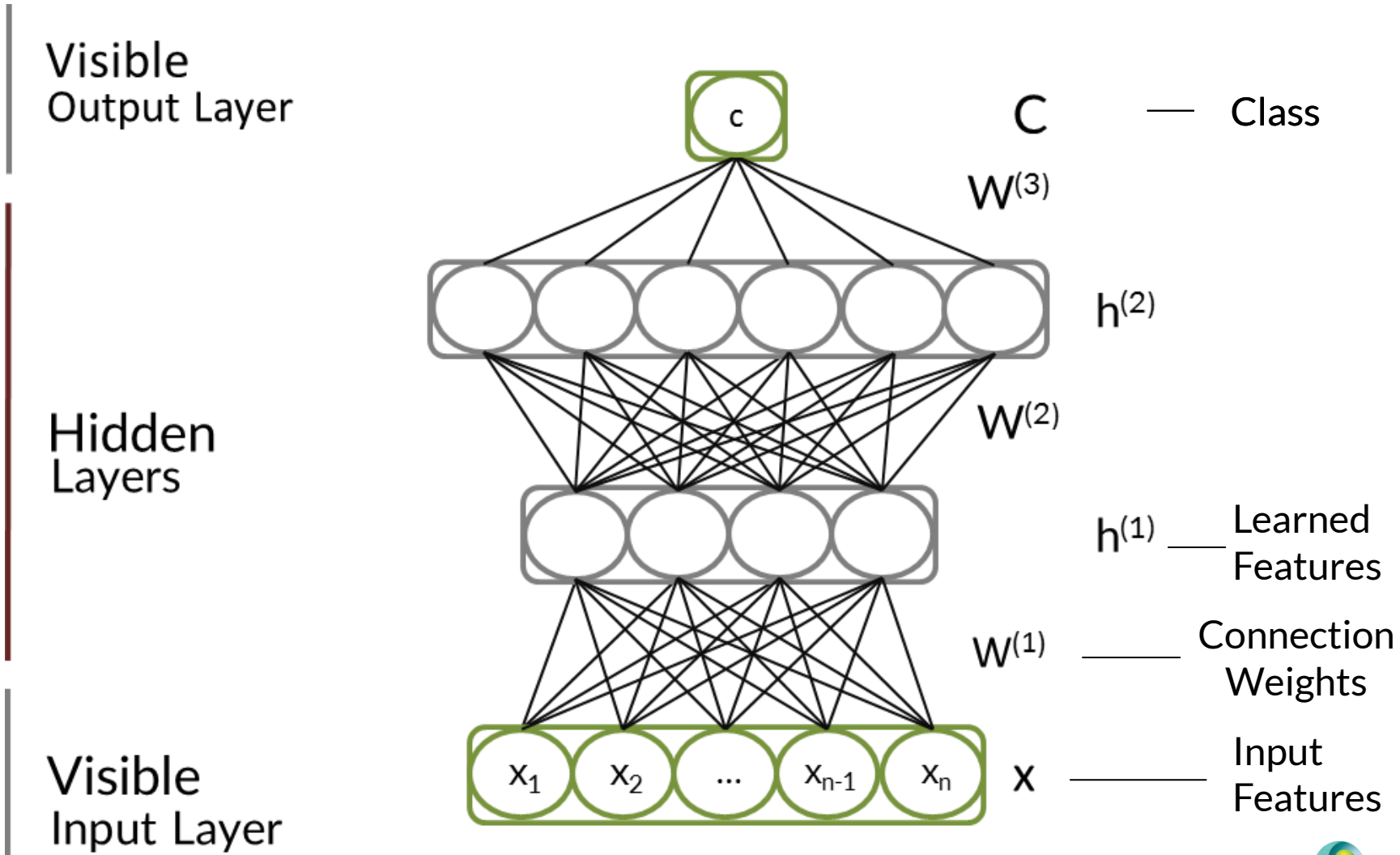




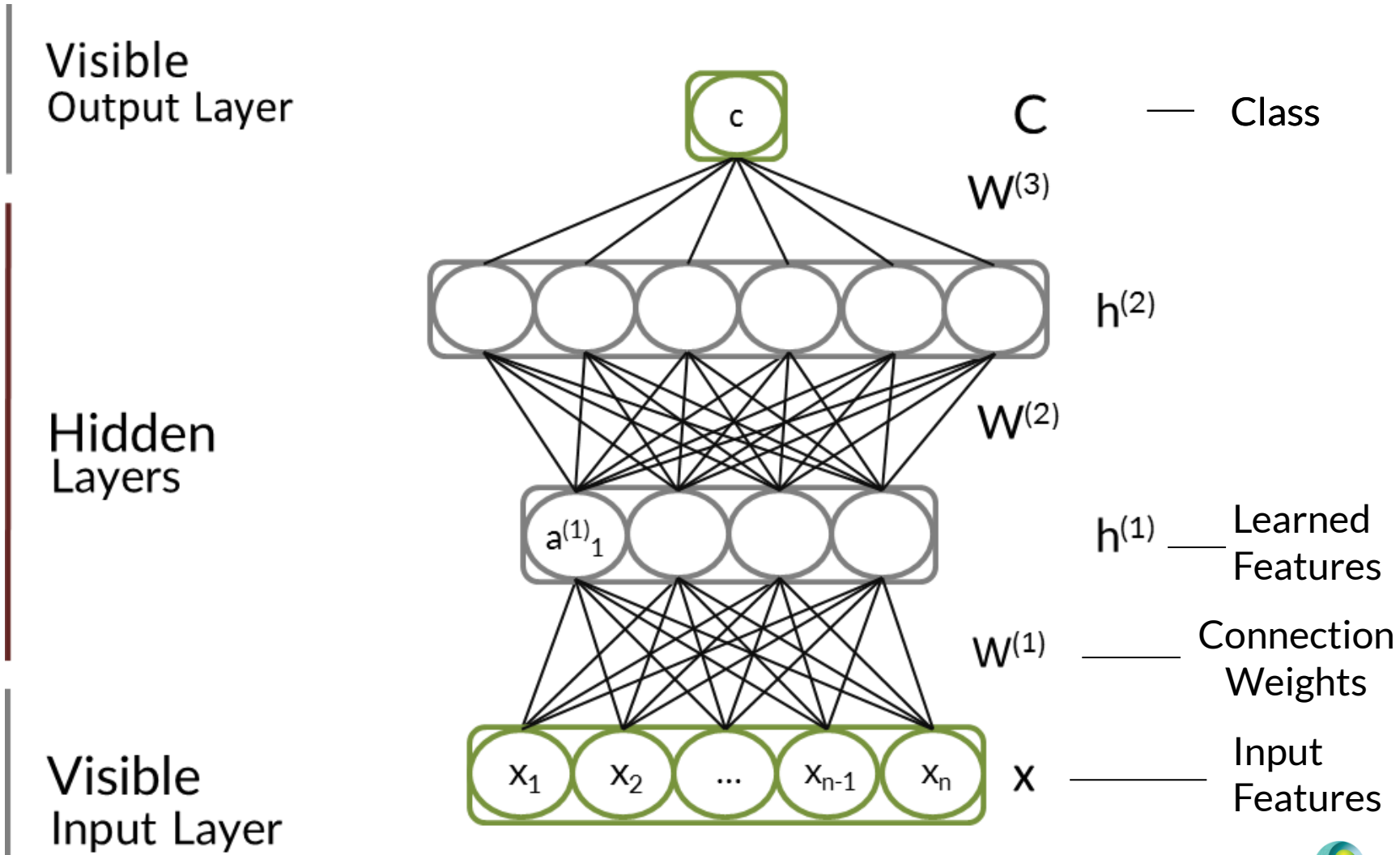


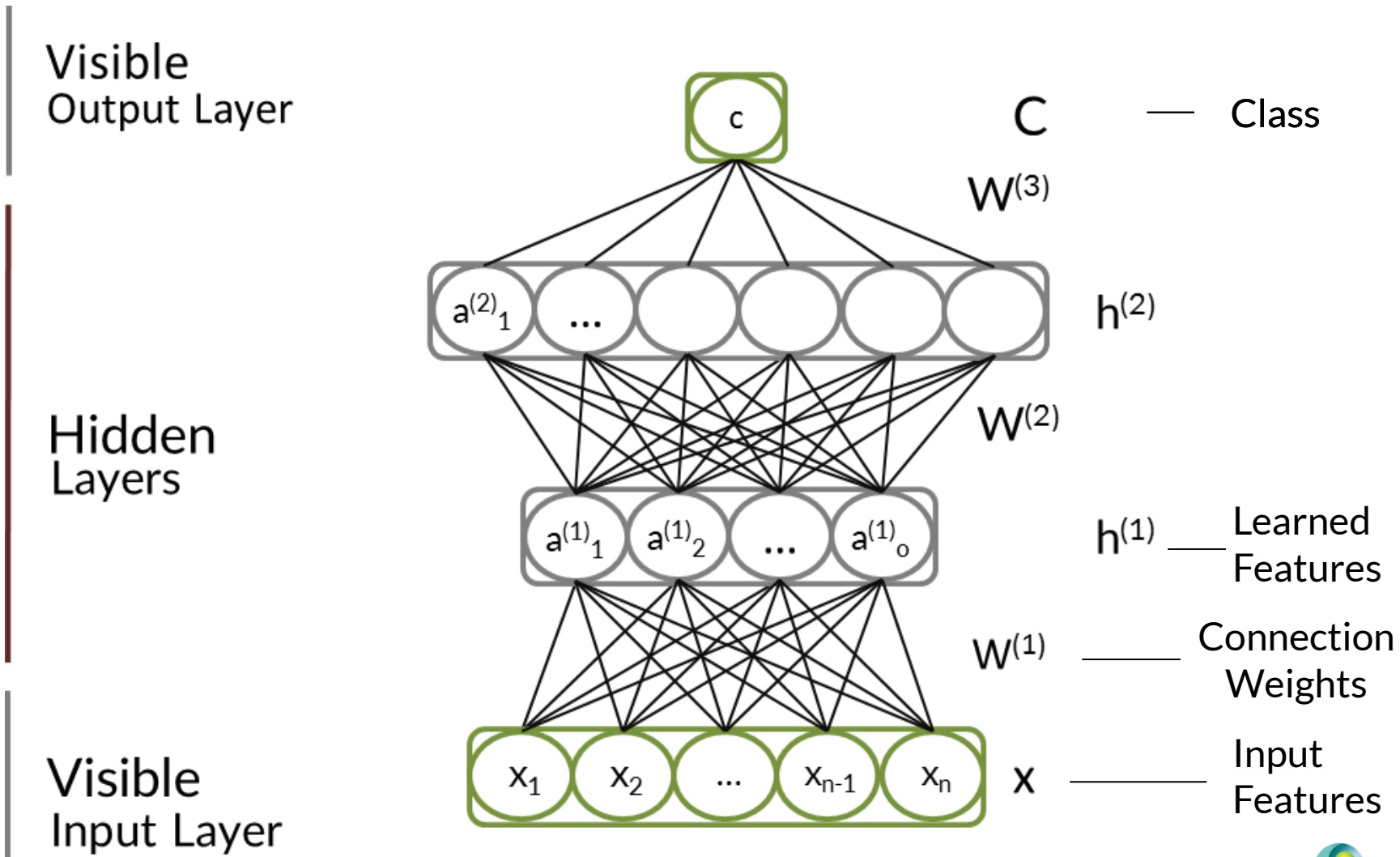


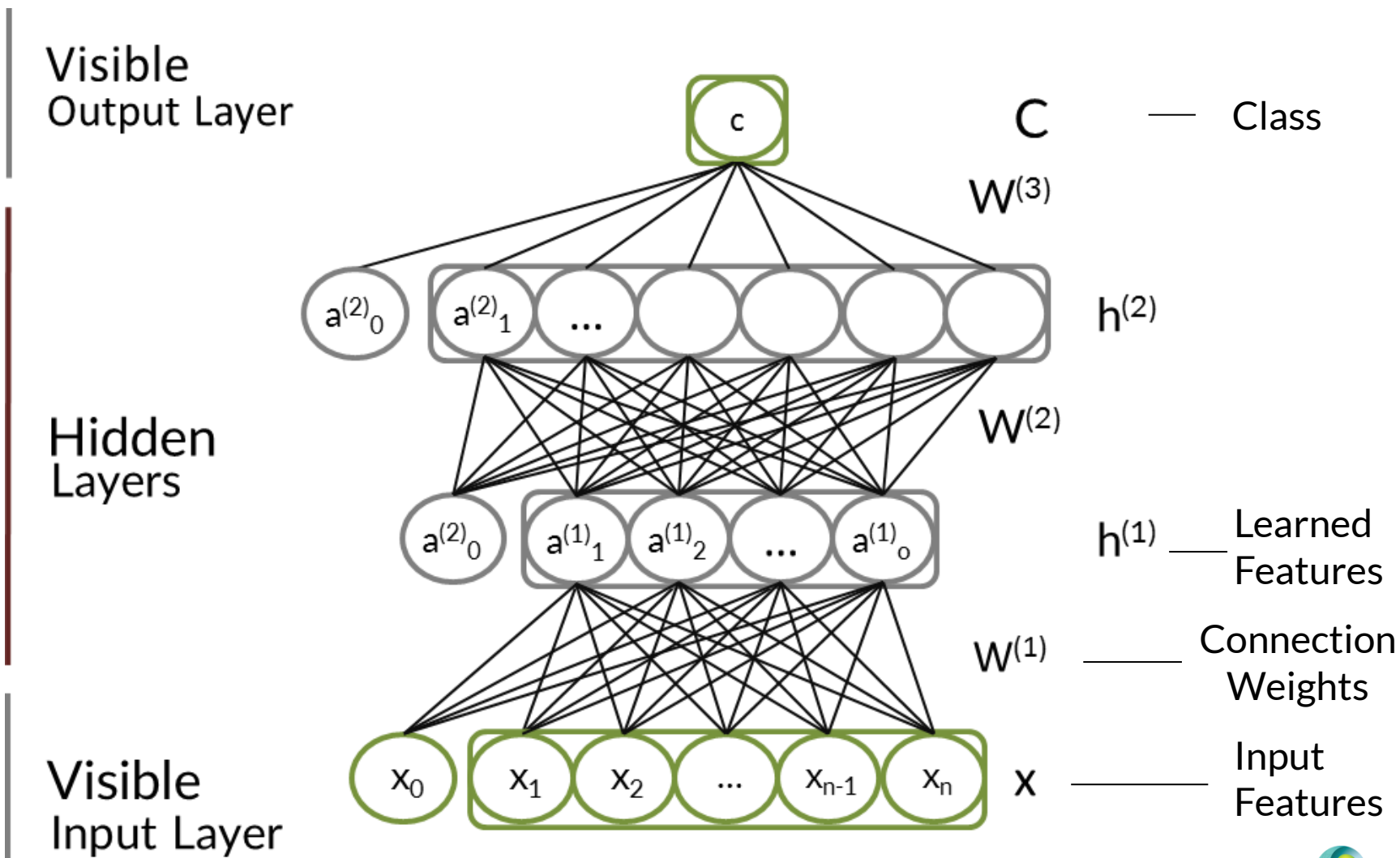






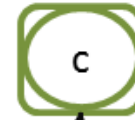






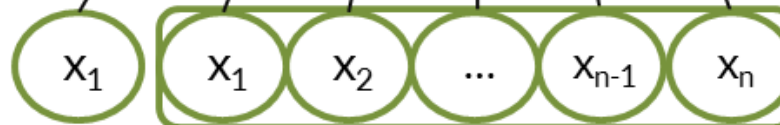
# ALGORITHM 1 REGRESSION

Visible  
Output Layer



$C$  — Class

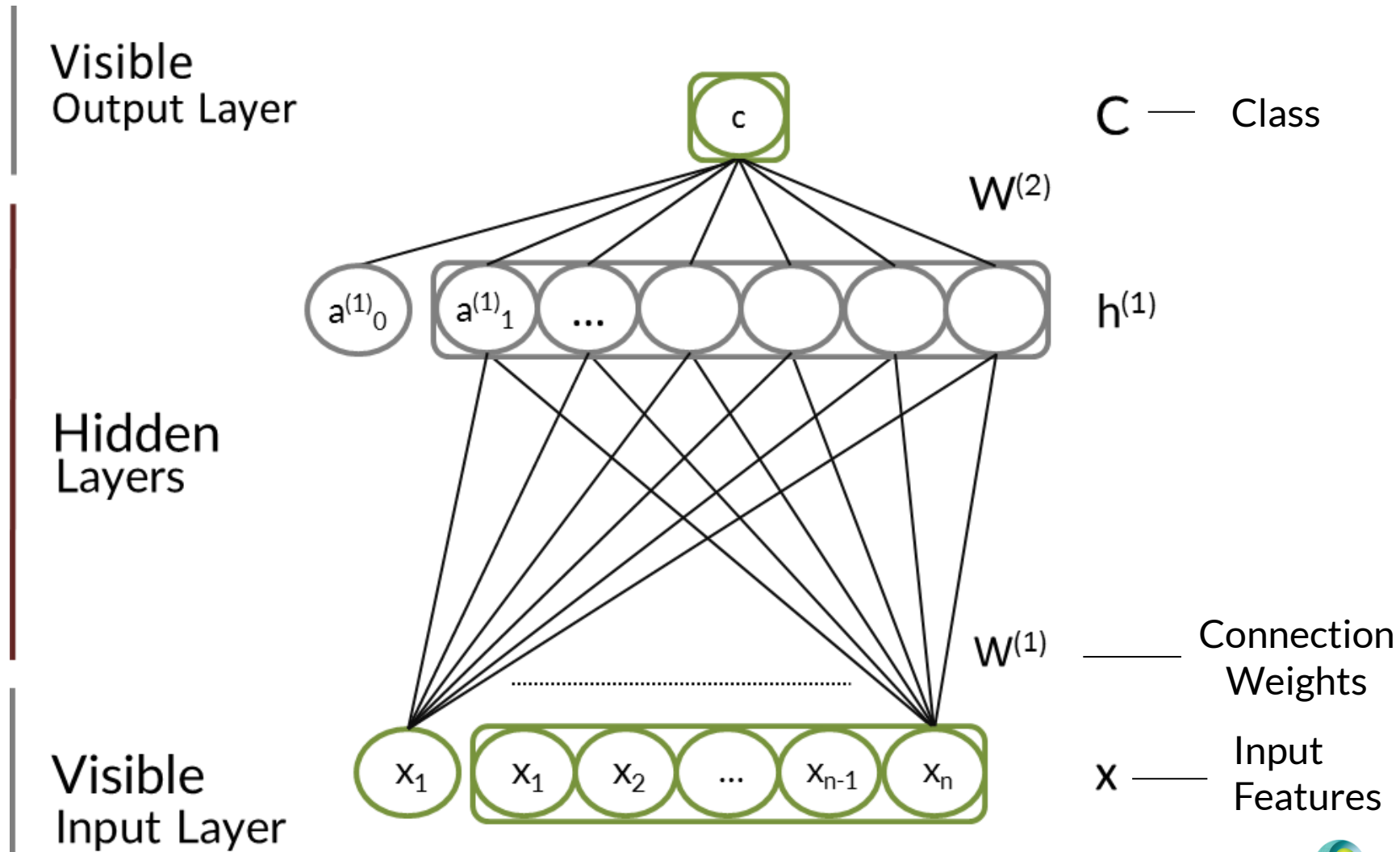
Visible  
Input Layer



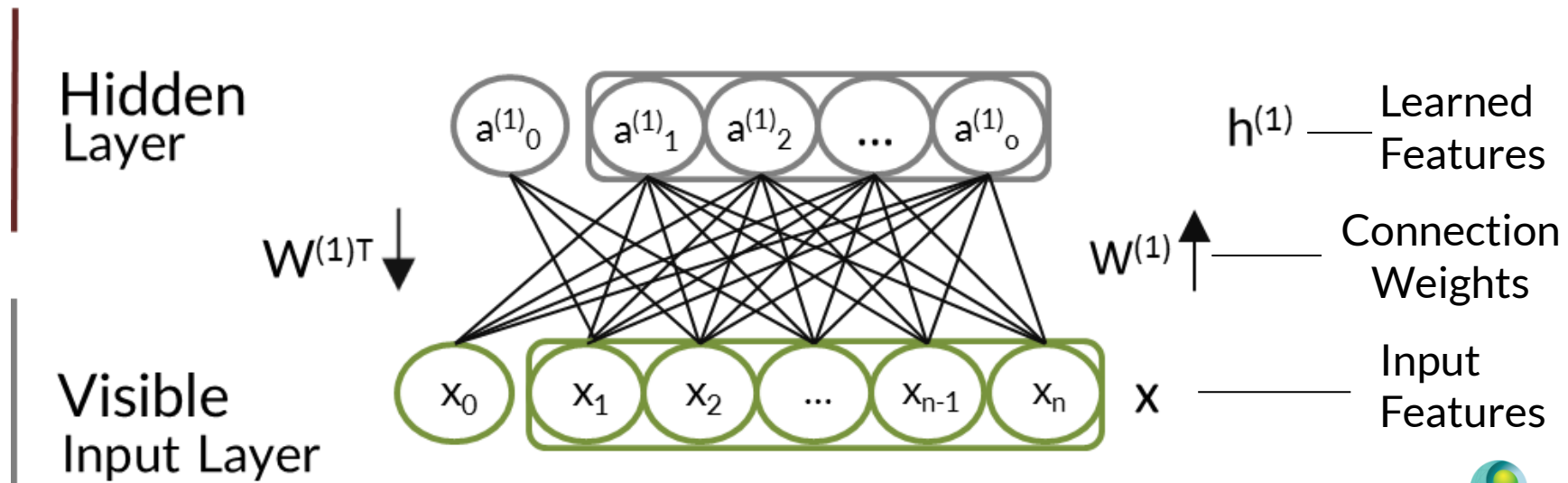
$W^{(1)}$  — Connection Weights

$X$  — Input Features

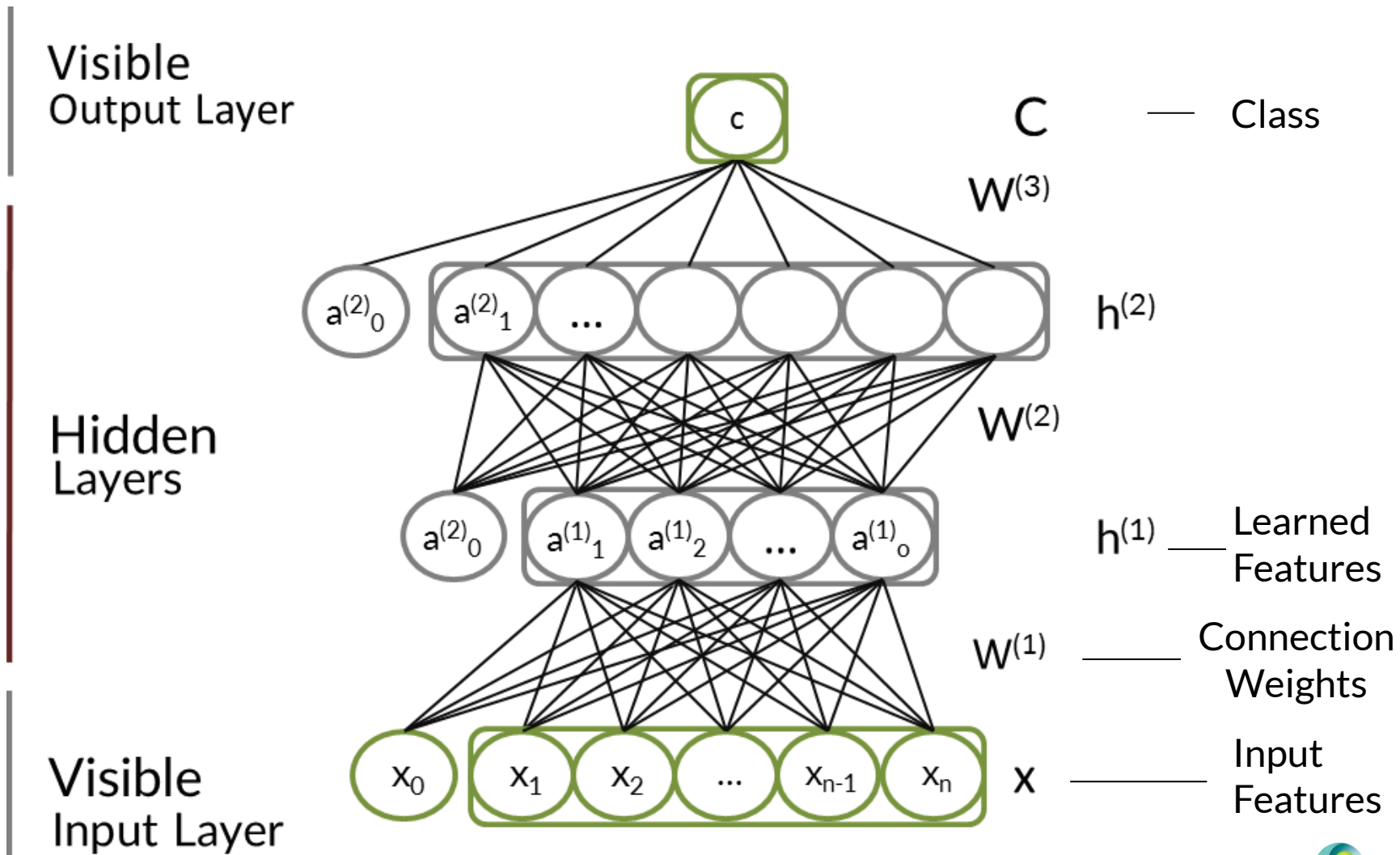
# ALGORITHM 2 MLP



# ALGORITHM 3 RBM



# ALGORITHM 4 DBN



# CONFIGURABLE DEEP NETWORK FRAMEWORK

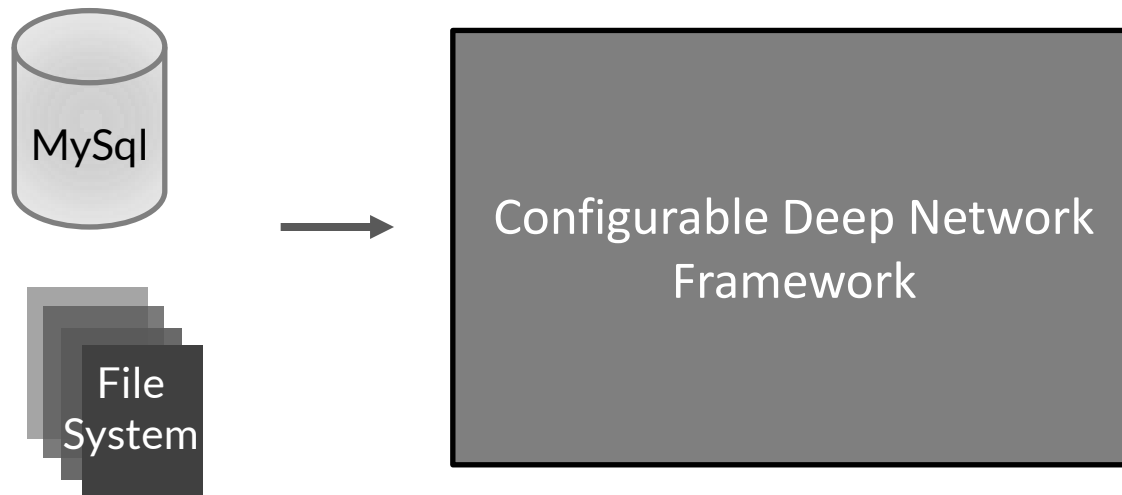


In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979

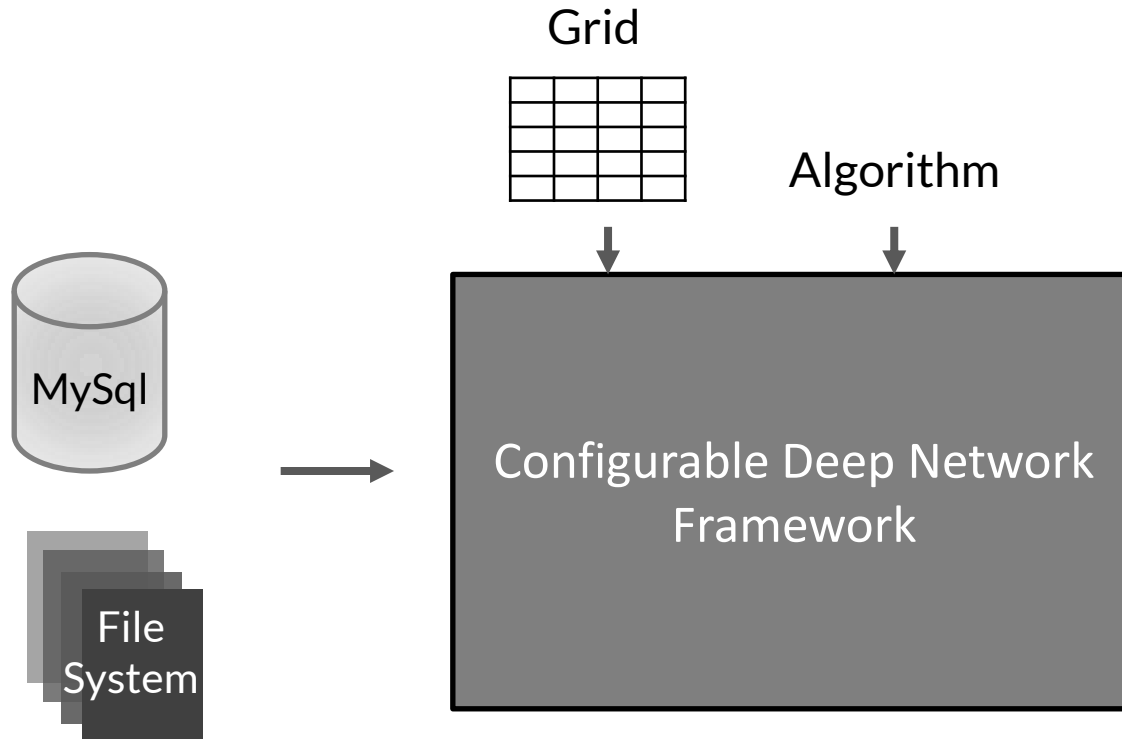




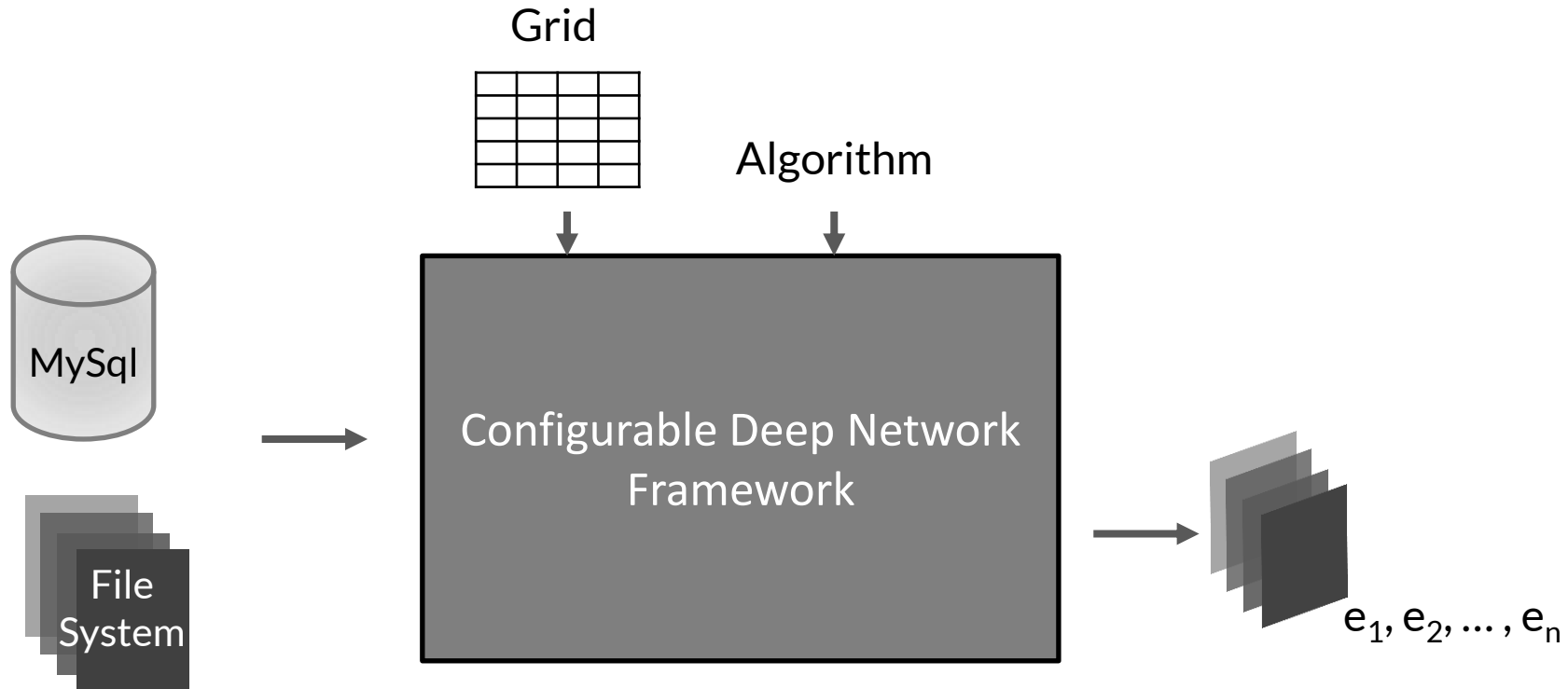
# CONFIGURABLE DEEP NETWORK FRAMEWORK



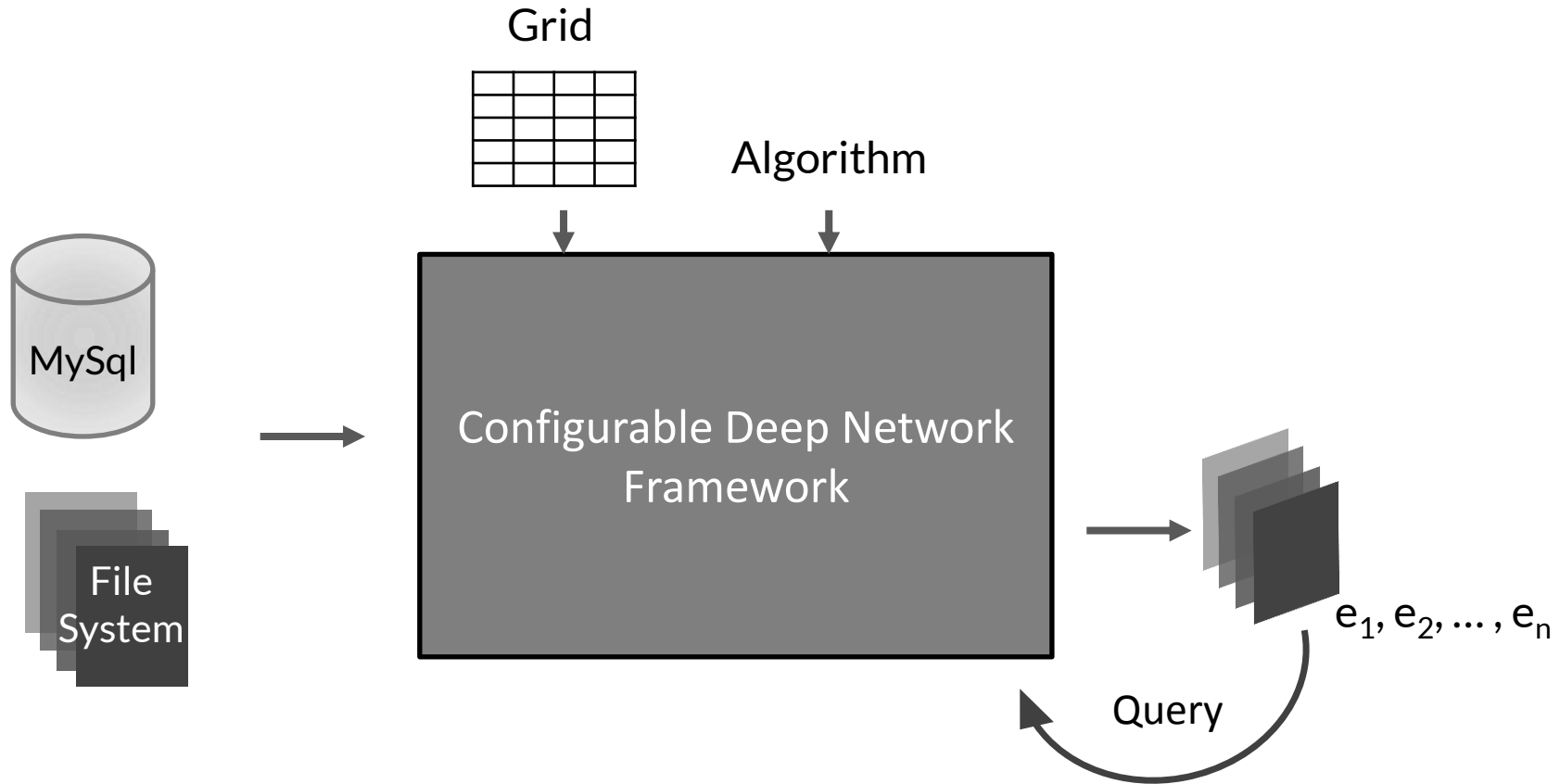
# CONFIGURABLE DEEP NETWORK FRAMEWORK



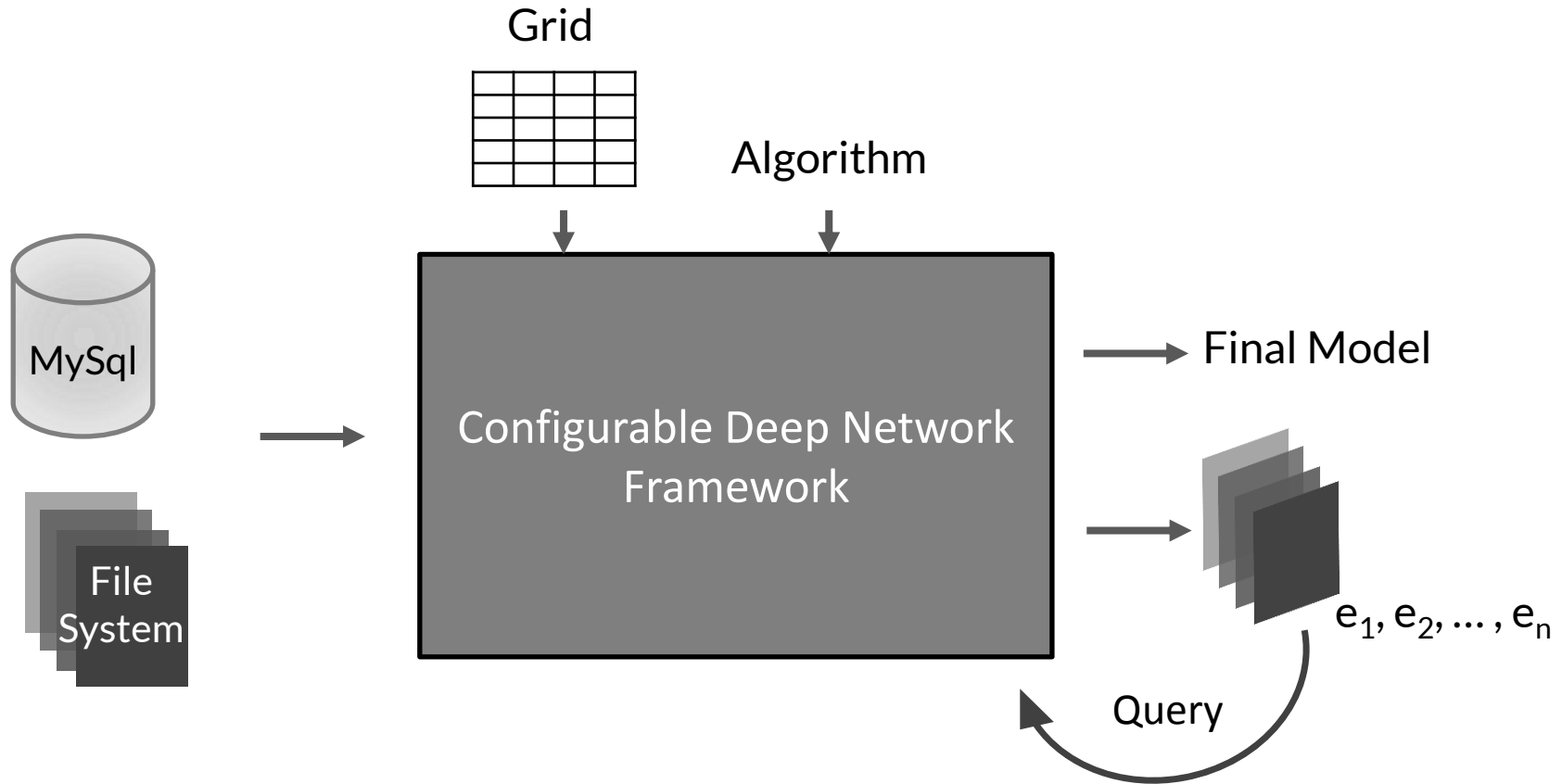
# CONFIGURABLE DEEP NETWORK FRAMEWORK



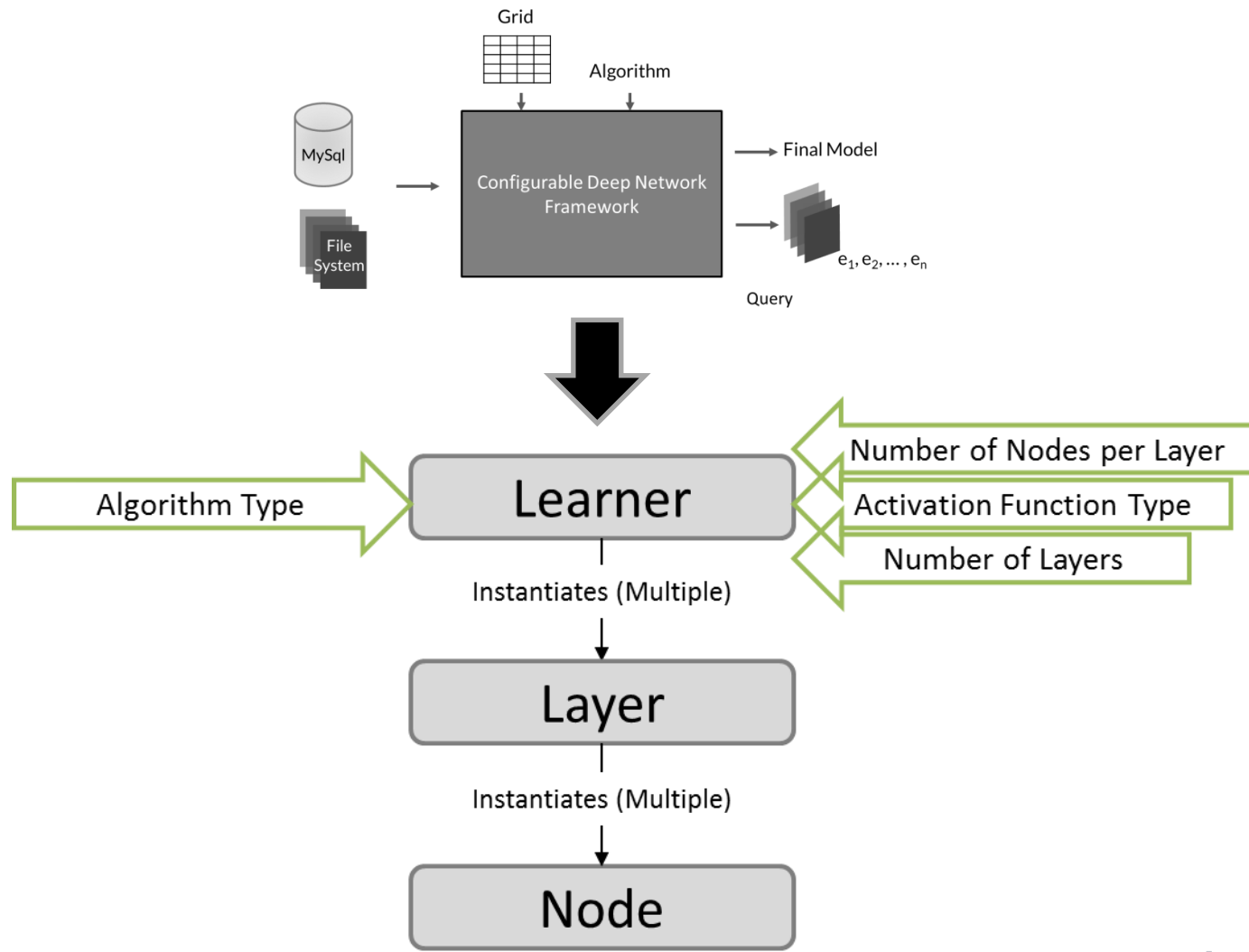
# CONFIGURABLE DEEP NETWORK FRAMEWORK



# CONFIGURABLE DEEP NETWORK FRAMEWORK

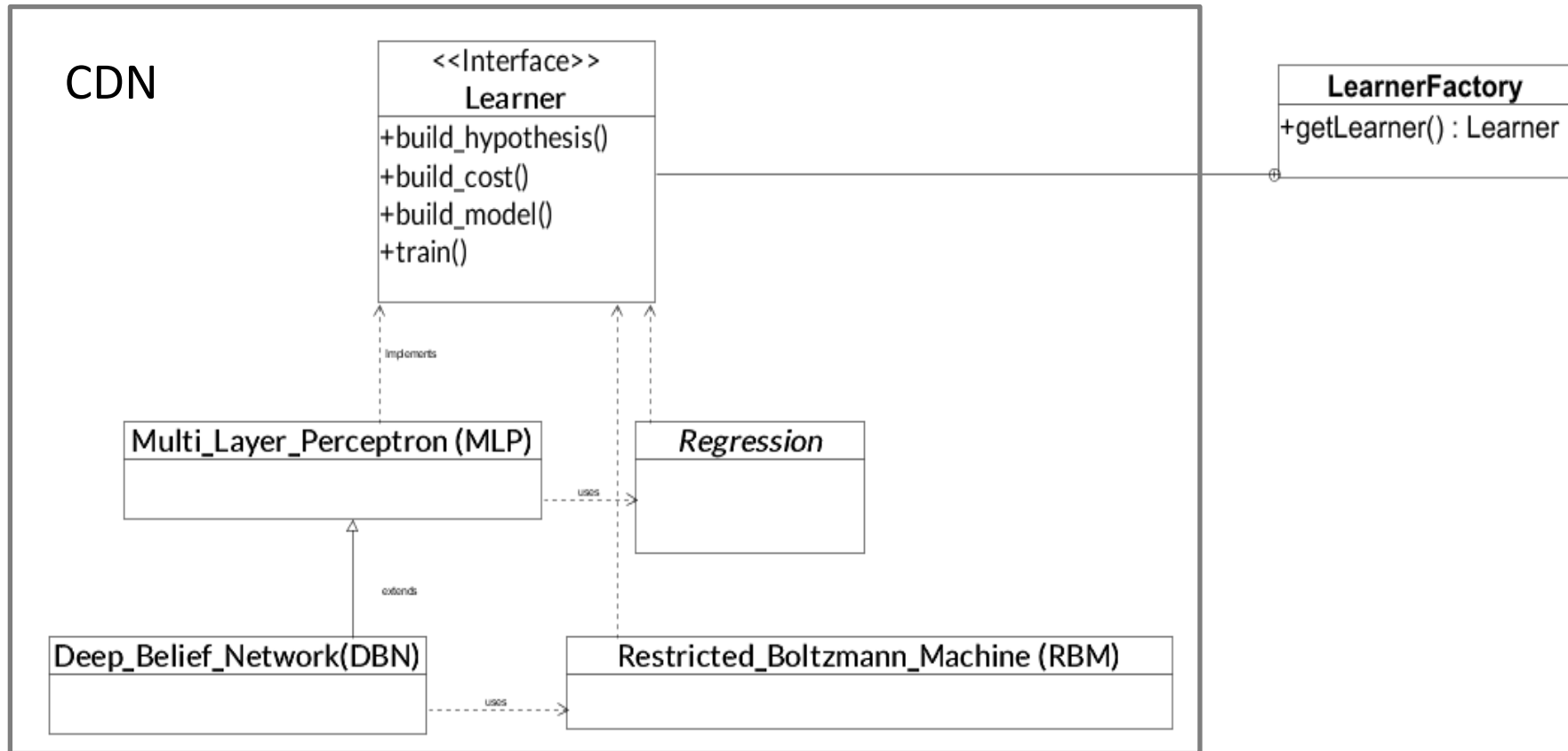


# CONFIGURABLE DEEP NETWORK FRAMEWORK



# CONFIGURABLE DEEP NETWORK

## ALGORITHM ARCHITECTURE



# EXPERIMENTS DATASET PREPARATION

Subset of the Data – dimensions

What the variables are

What the predictor is

Purpose





EXPERIMENT

# REGRESSION

To Choose:



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



EXPERIMENT

# REGRESSION

To Choose:

learning rate  $\alpha$



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



## EXPERIMENT

# REGRESSION

To Choose:

learning rate  $\alpha$

weight decay term  $\lambda$



## EXPERIMENT

# REGRESSION

To Choose:

learning rate  $\alpha$

weight decay term  $\lambda$

training iterations  $t$



EXPERIMENT

# REGRESSION

The Grid:



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



## EXPERIMENT

# REGRESSION

The Grid:

$\alpha, \lambda$ :

[0.001, 0.003, 0.009, ..., 0.1, 0.3, 0.9]



# REGRESSION

The Grid:

$\alpha, \lambda$ :

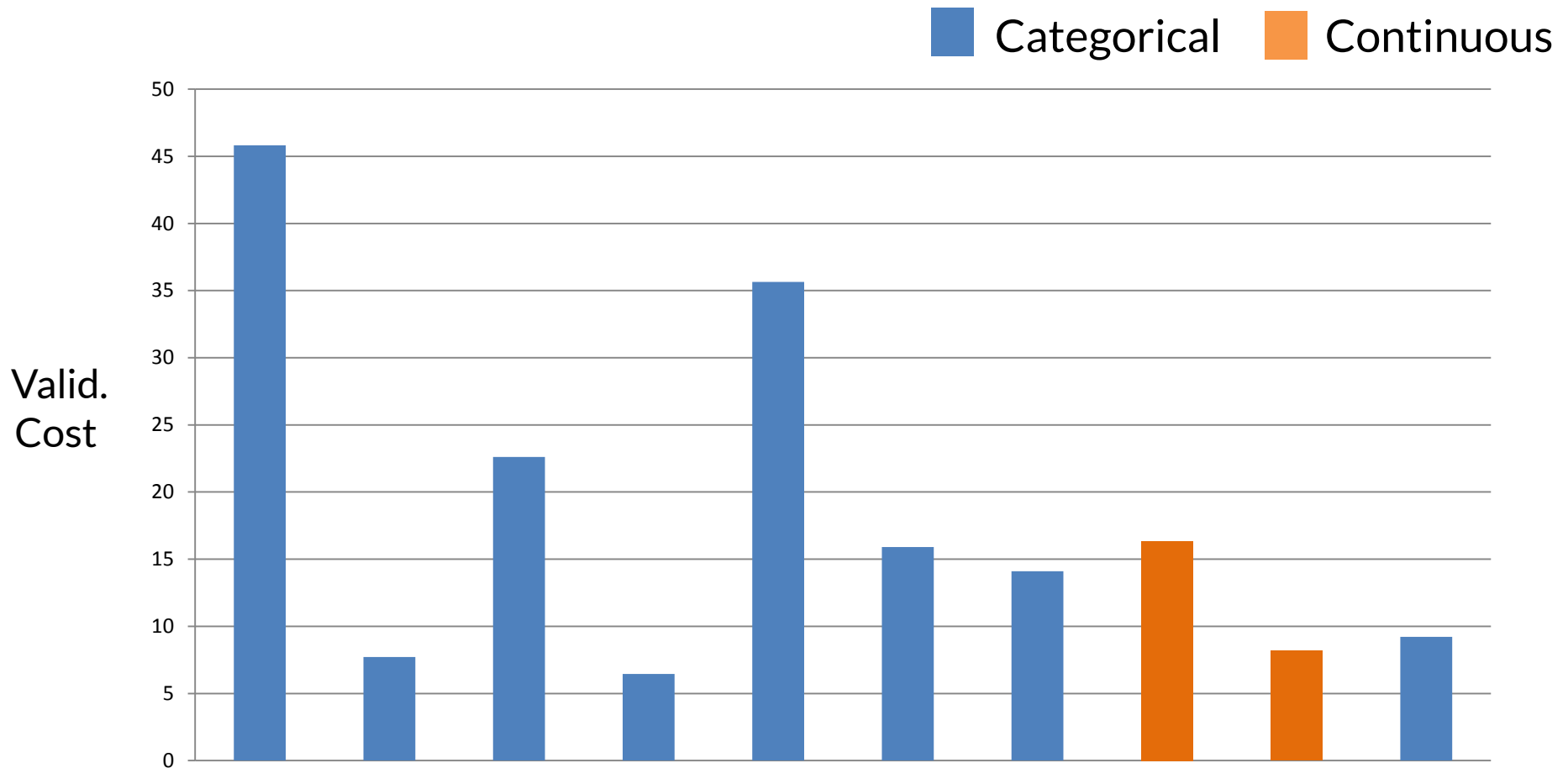
[0.001, 0.003, 0.009, ..., 0.1, 0.3, 0.9]

t:

[100, 1000, 10000]

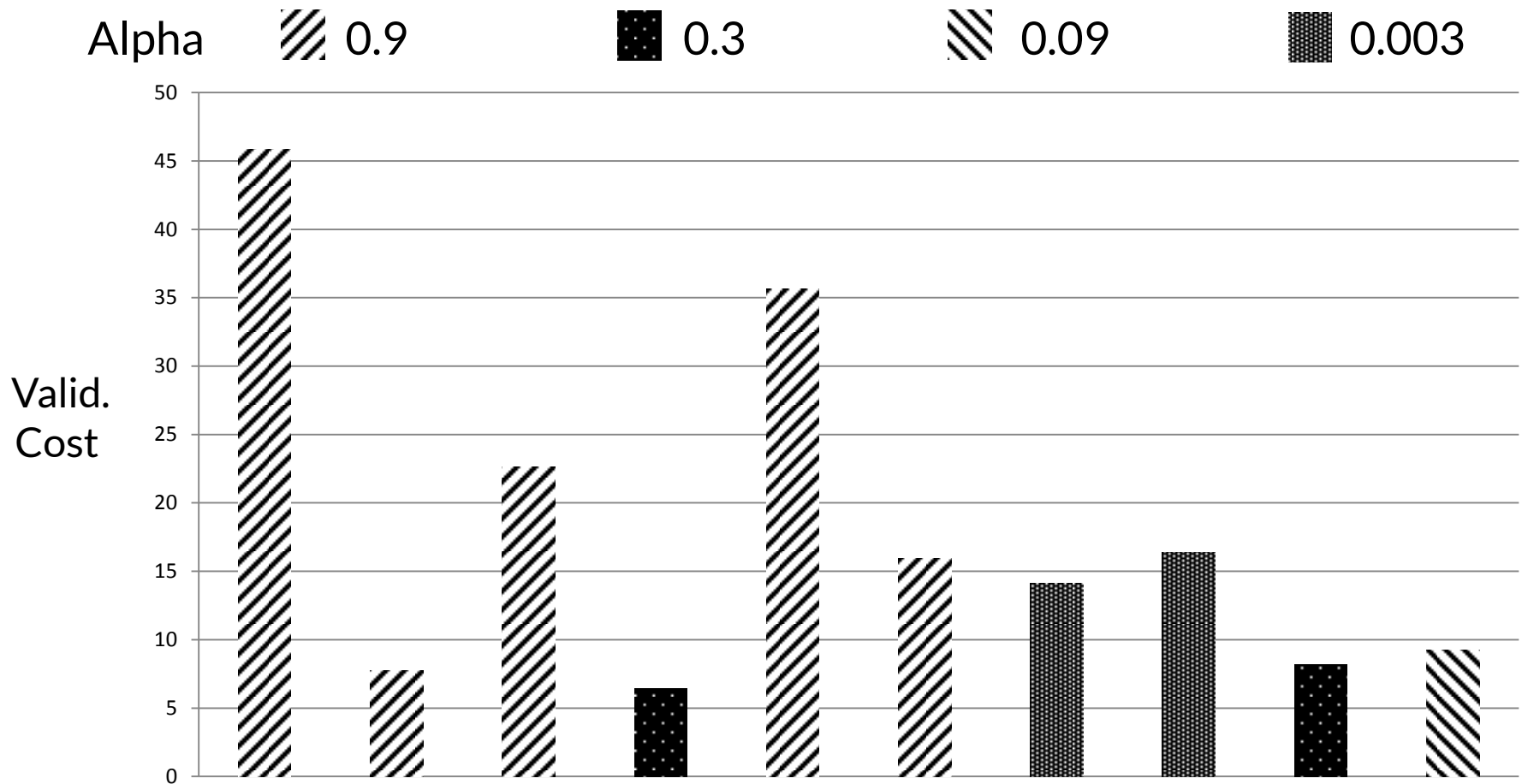


# EXPERIMENT RESULTS REGRESSION

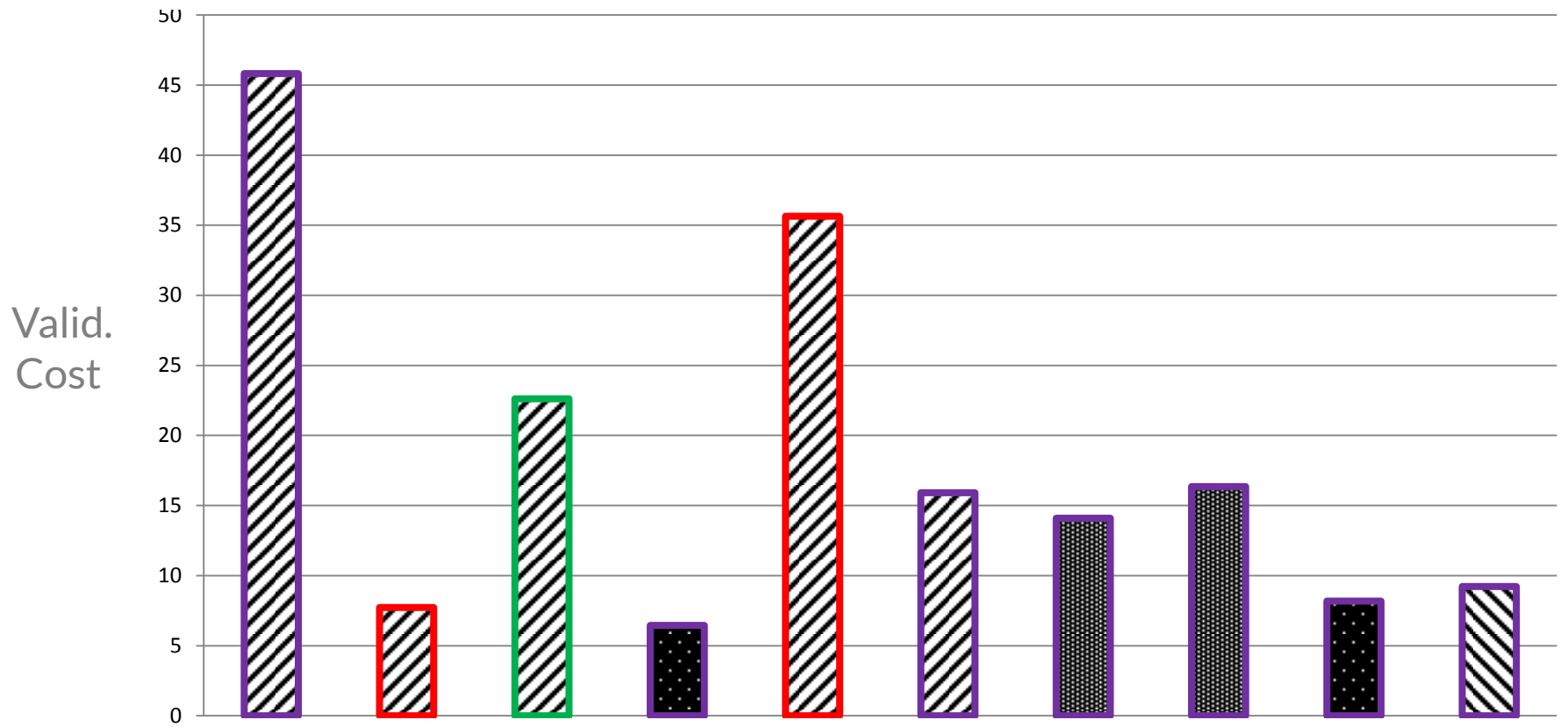
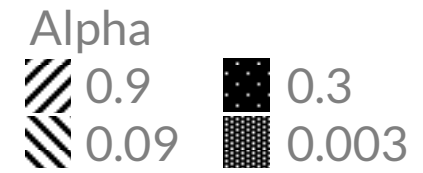




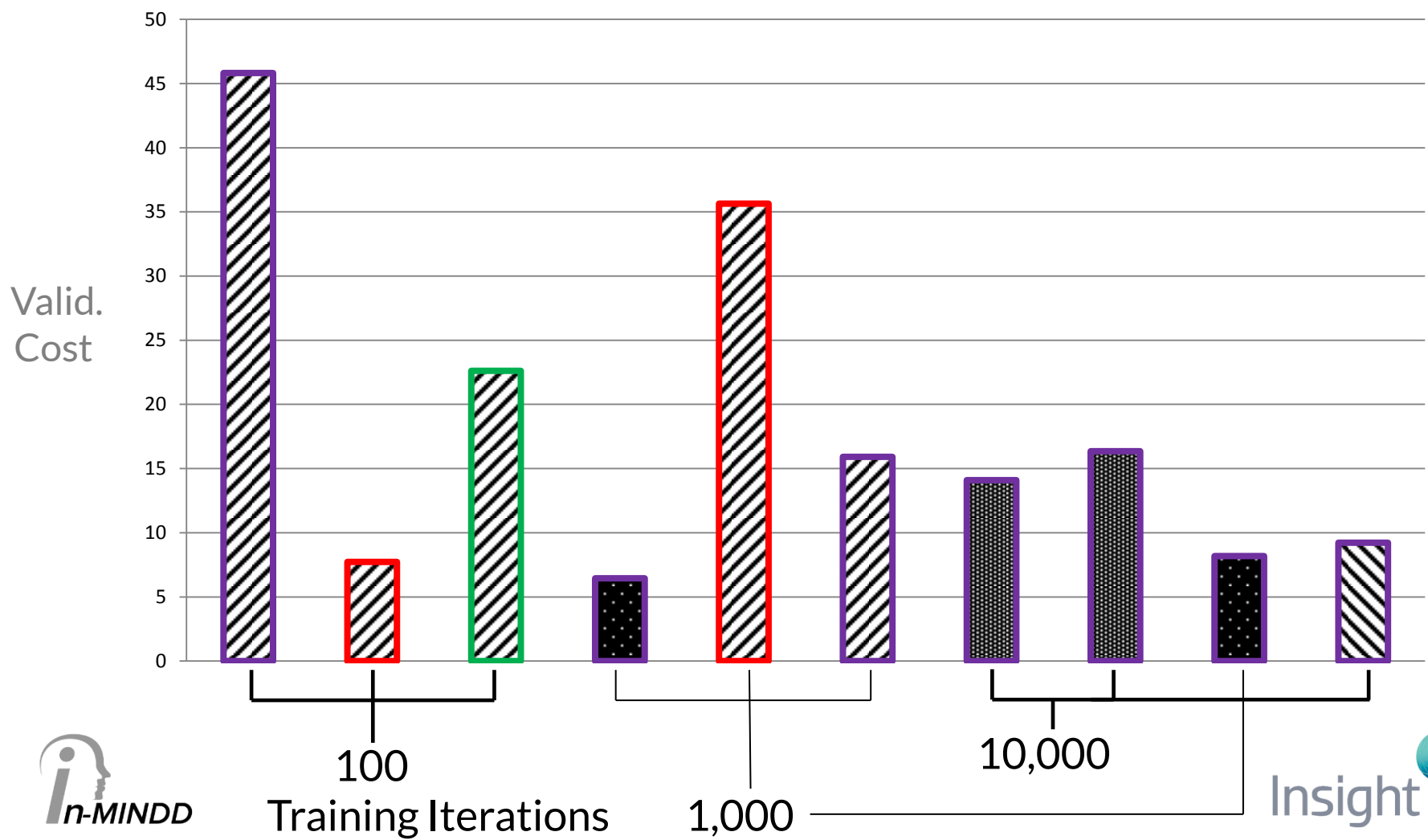
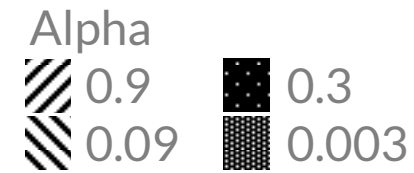
# EXPERIMENT RESULTS REGRESSION



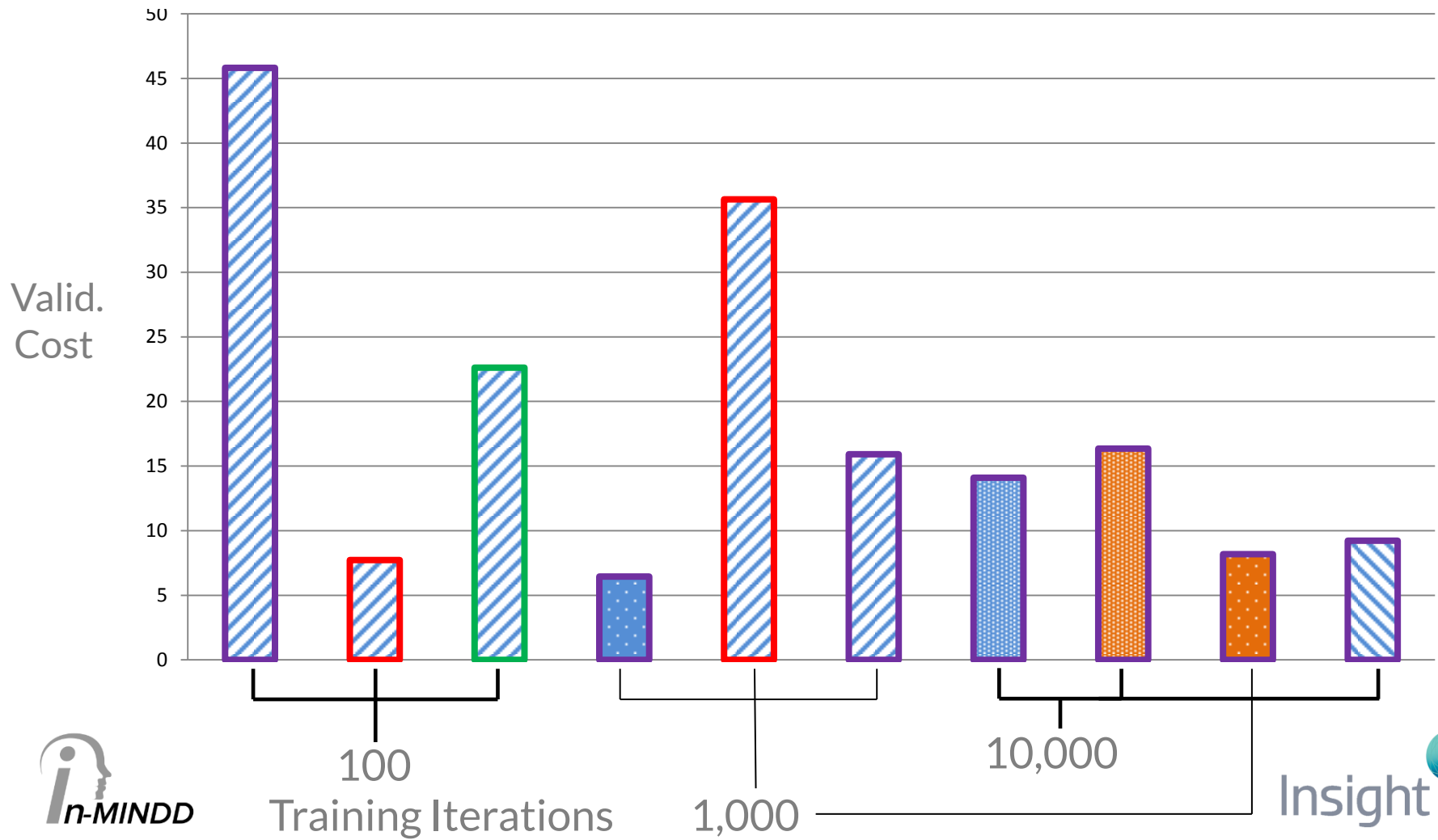
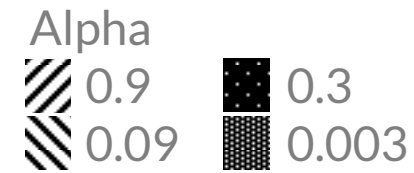
# EXPERIMENT RESULTS REGRESSION



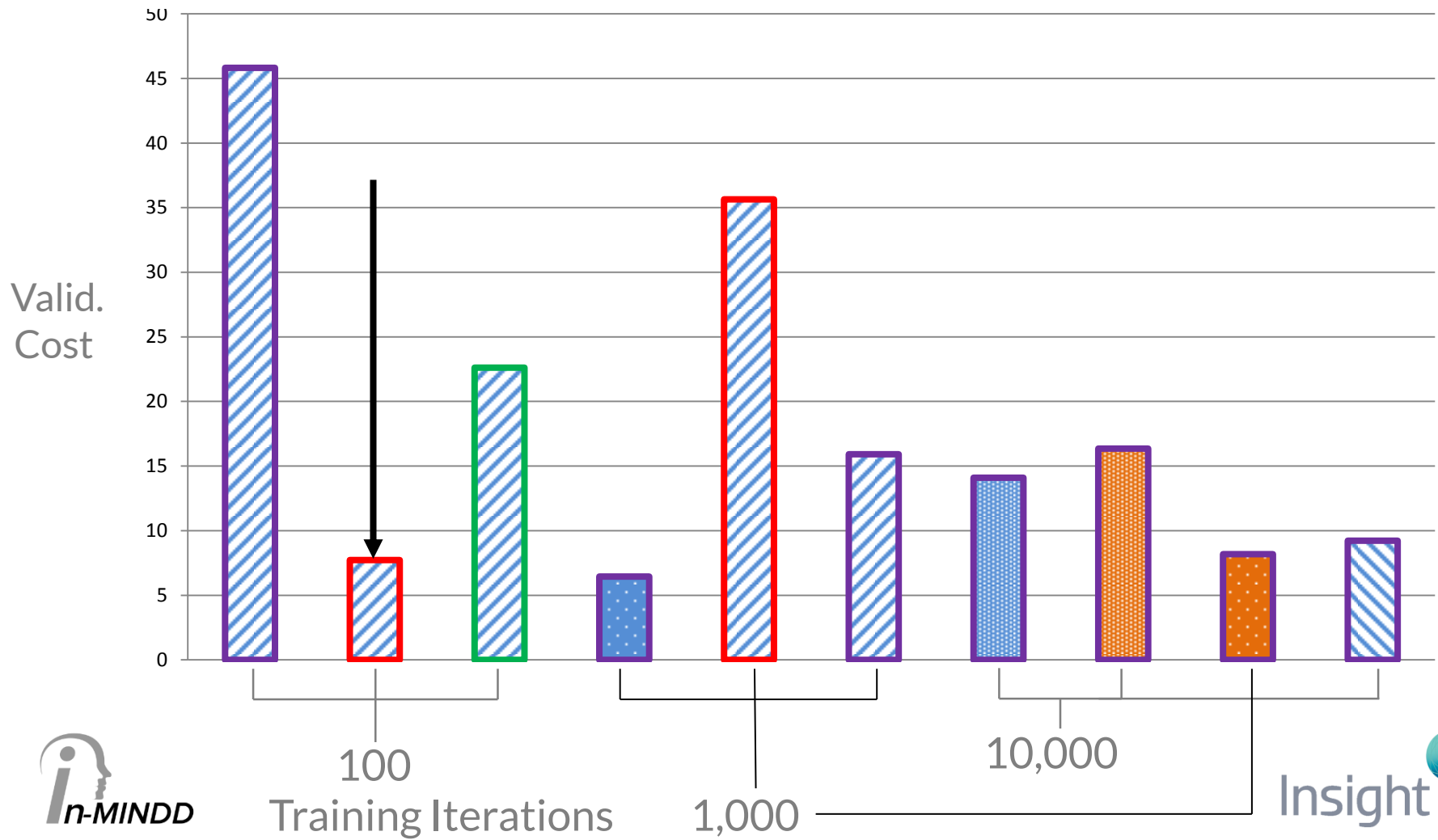
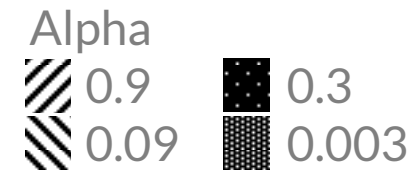
# EXPERIMENT RESULTS REGRESSION



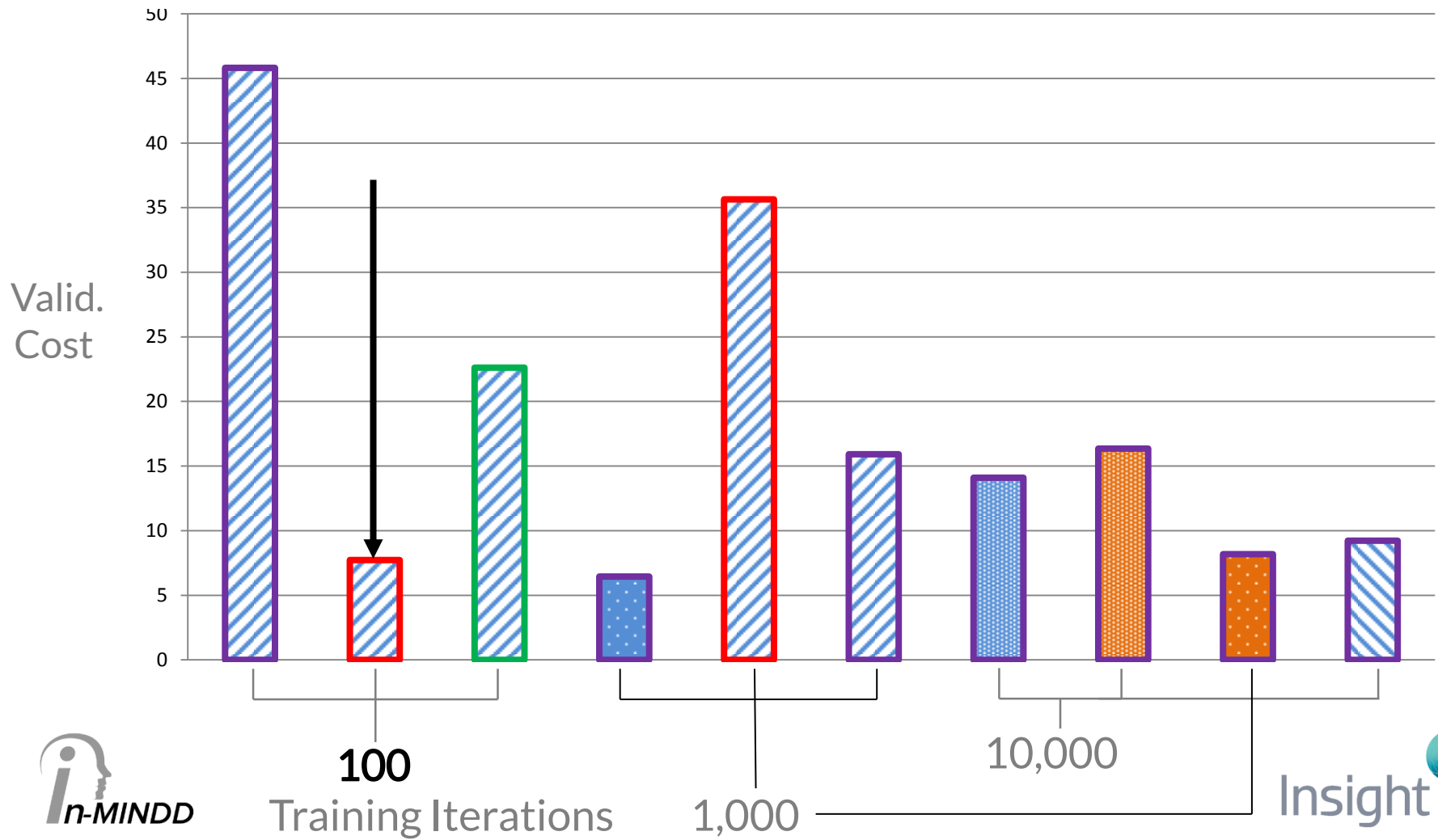
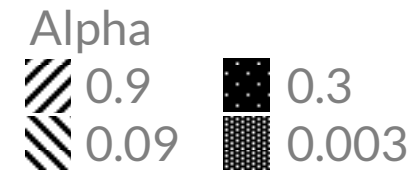
# EXPERIMENT RESULTS REGRESSION



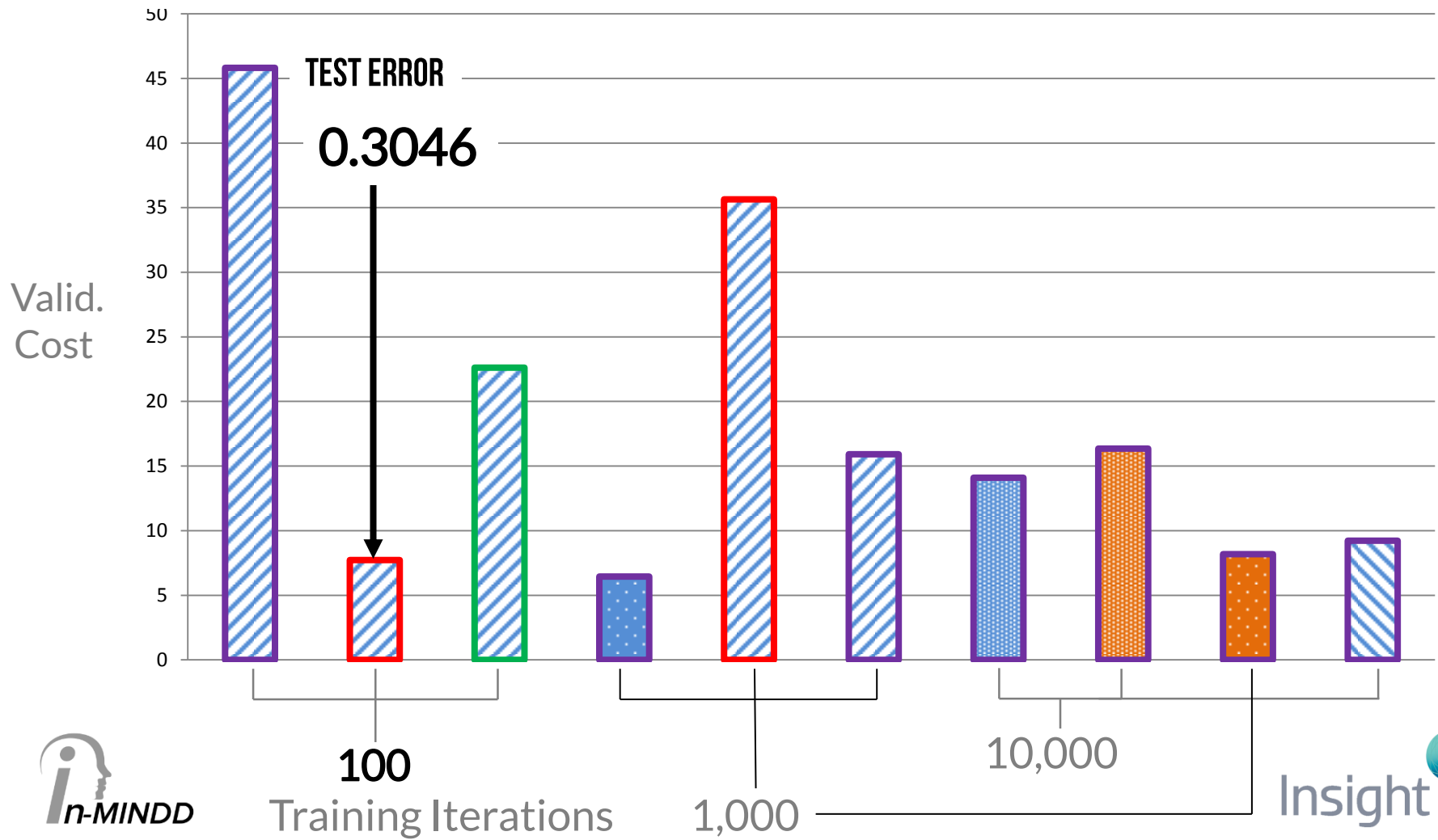
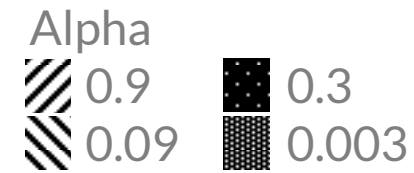
# EXPERIMENT RESULTS REGRESSION



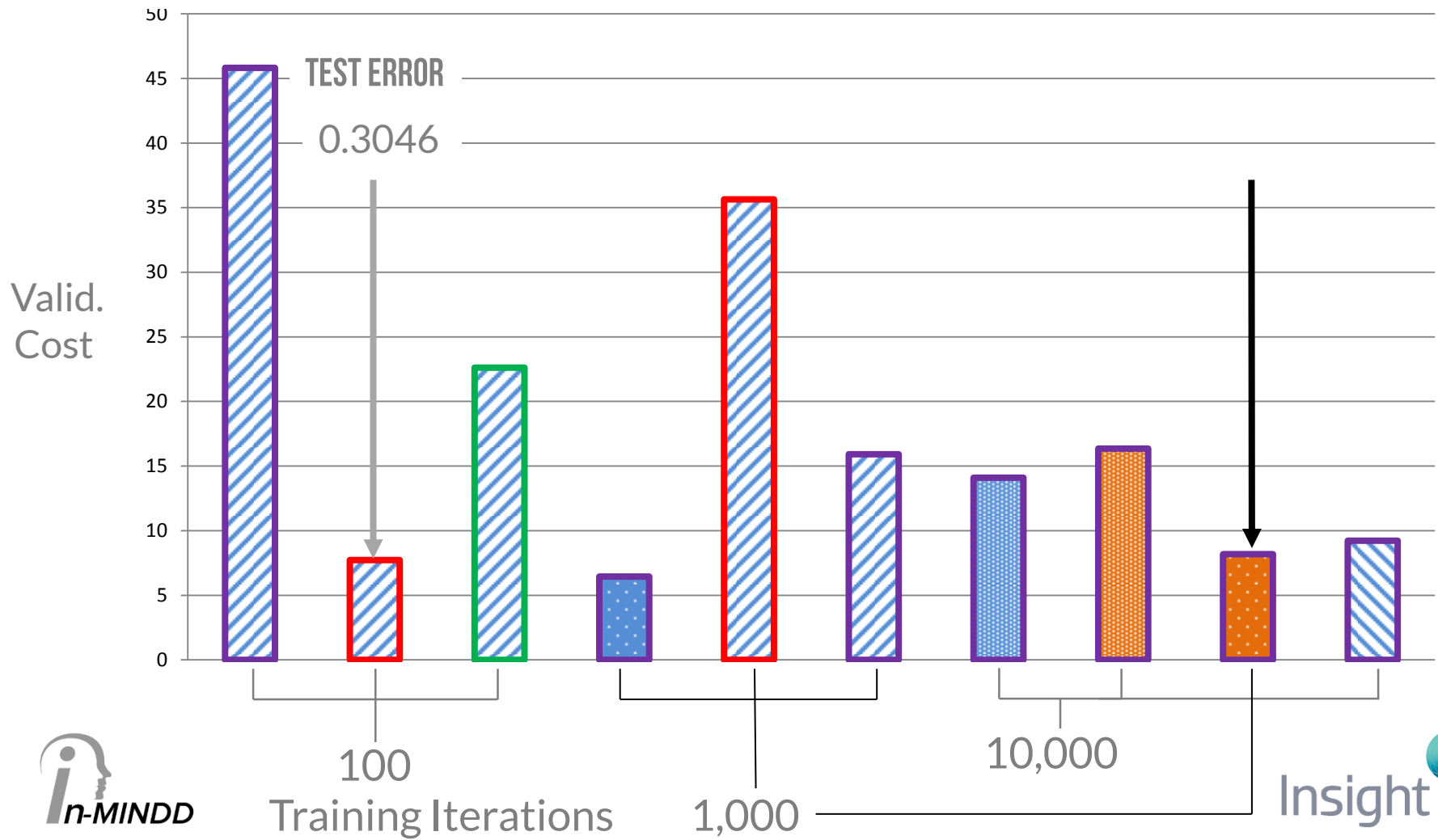
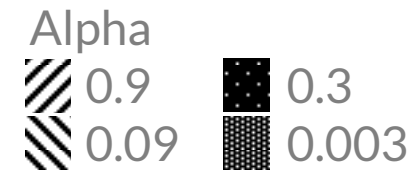
# EXPERIMENT RESULTS REGRESSION



# EXPERIMENT RESULTS REGRESSION

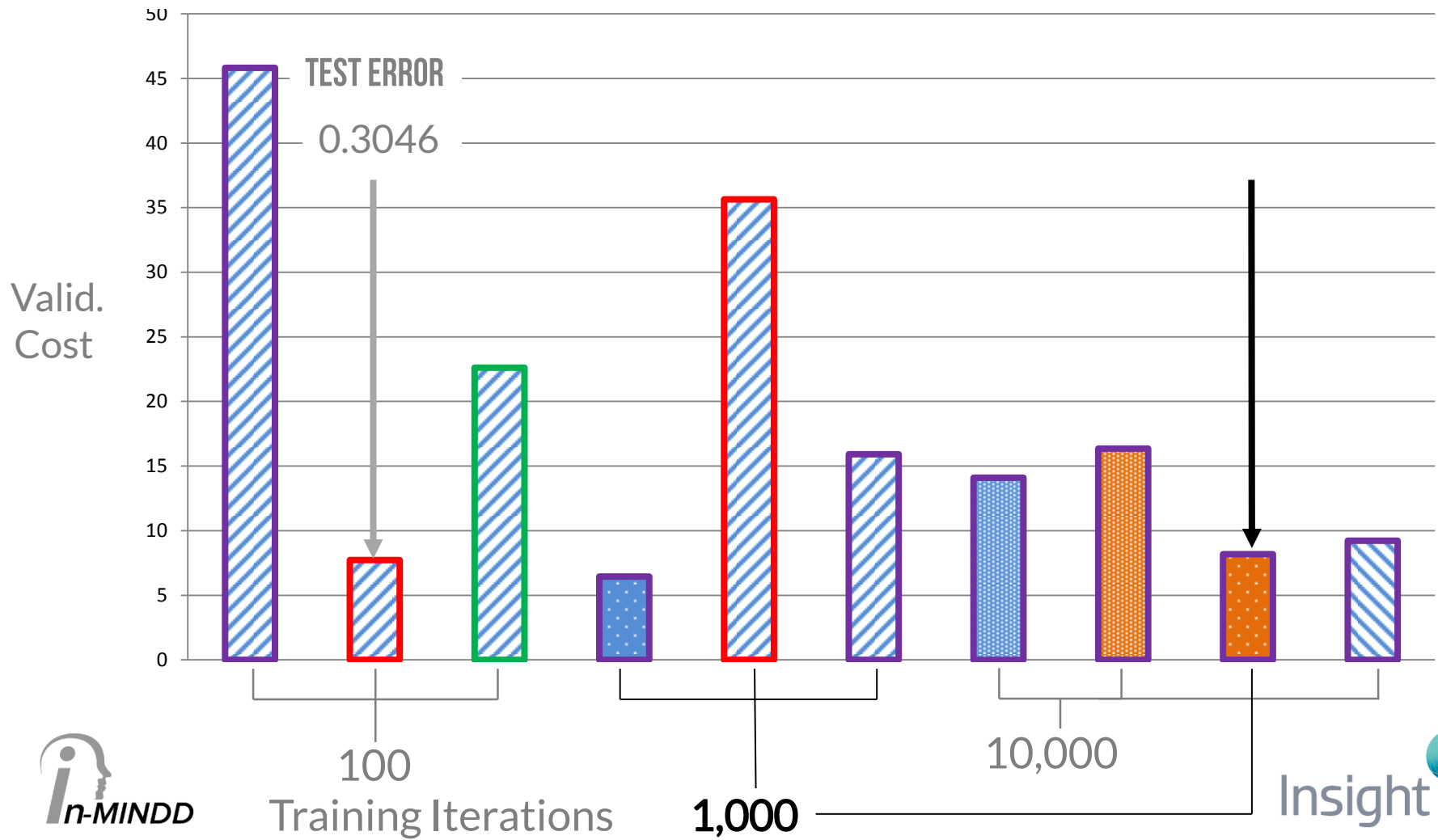
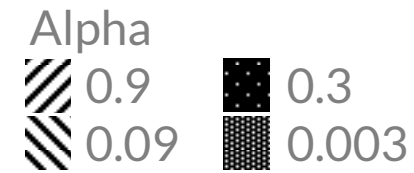


# EXPERIMENT RESULTS REGRESSION

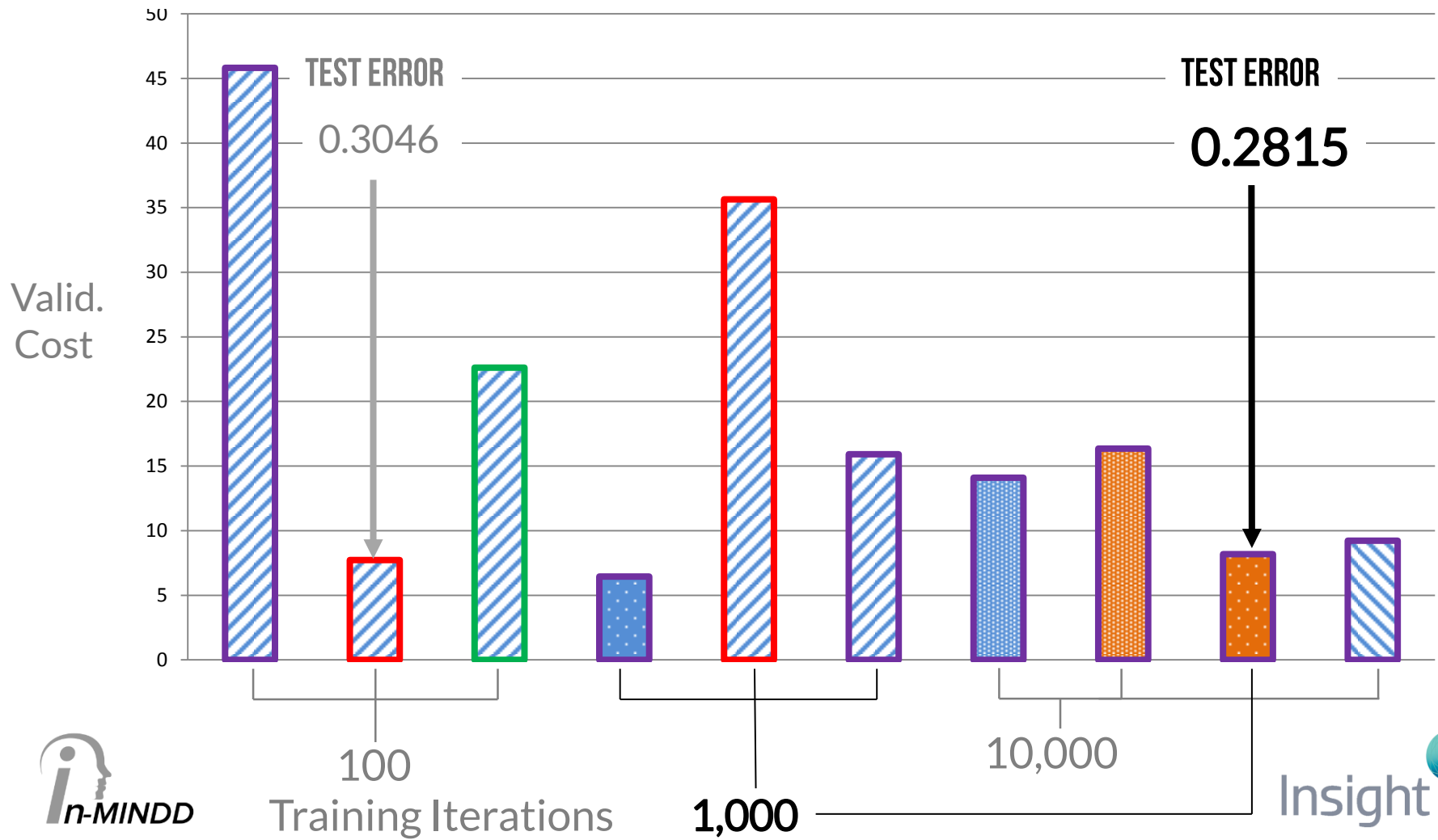
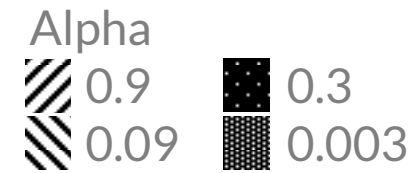




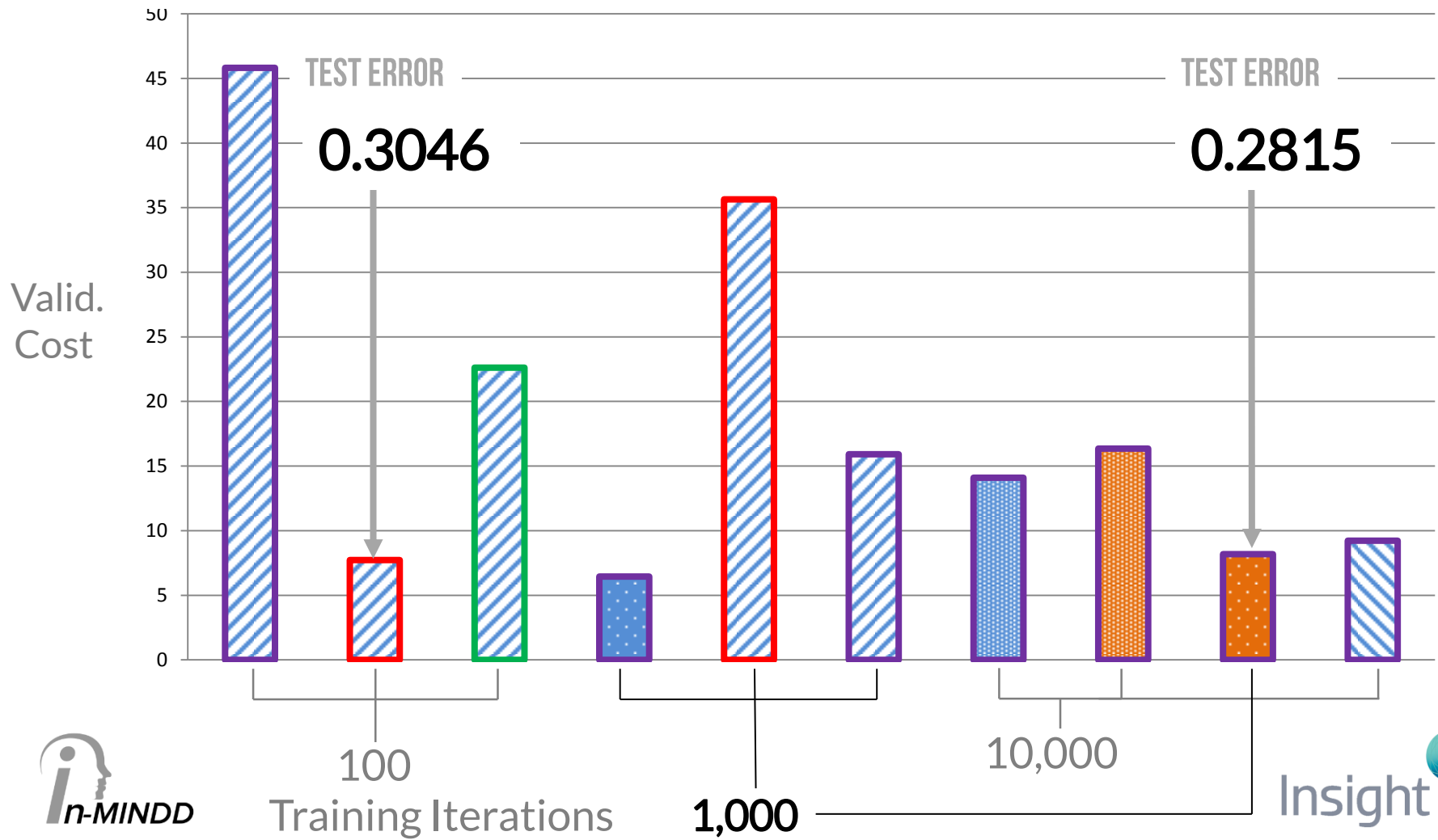
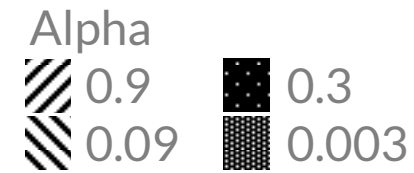
# EXPERIMENT RESULTS REGRESSION



# EXPERIMENT RESULTS REGRESSION



# EXPERIMENT RESULTS REGRESSION



EXPERIMENT

**RBM**

To Choose:



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# RBM

To Choose:

layer 1 nodes  $h^{(1)}_n$

## EXPERIMENT

**RBM**

To Choose:

layer 1 nodes  $h^{(1)}_n$

pre-training epochs  $e$



The Grid:

$h^{(1)}_n$ :

[10, 30, 337, 900, 1300, 2000]

## EXPERIMENT

**RBM**

The Grid:

$h^{(1)}_n$ :

[10, 30, 337, 900, 1300, 2000]

e

[1, 5, 10, 15, 20]

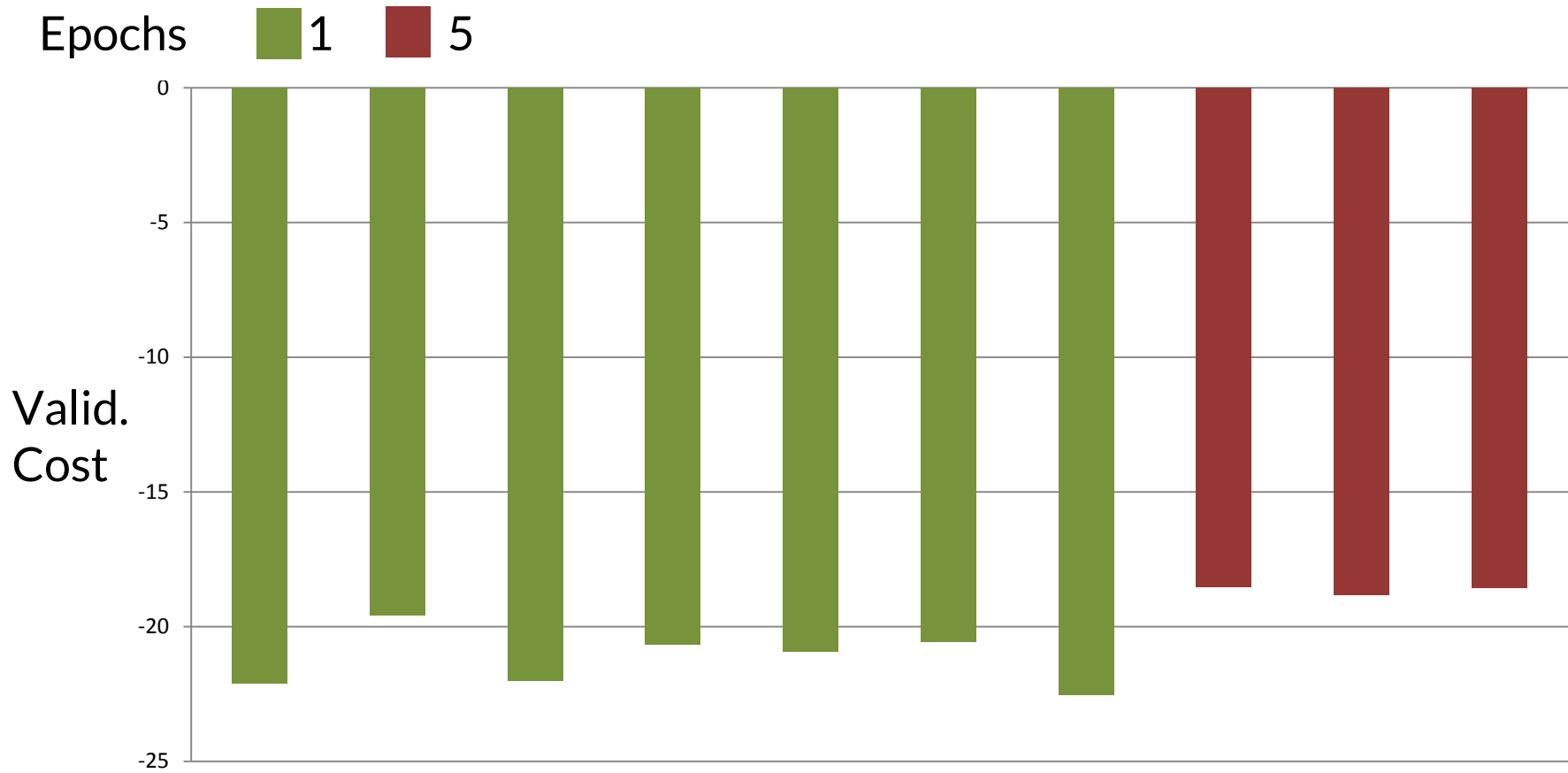




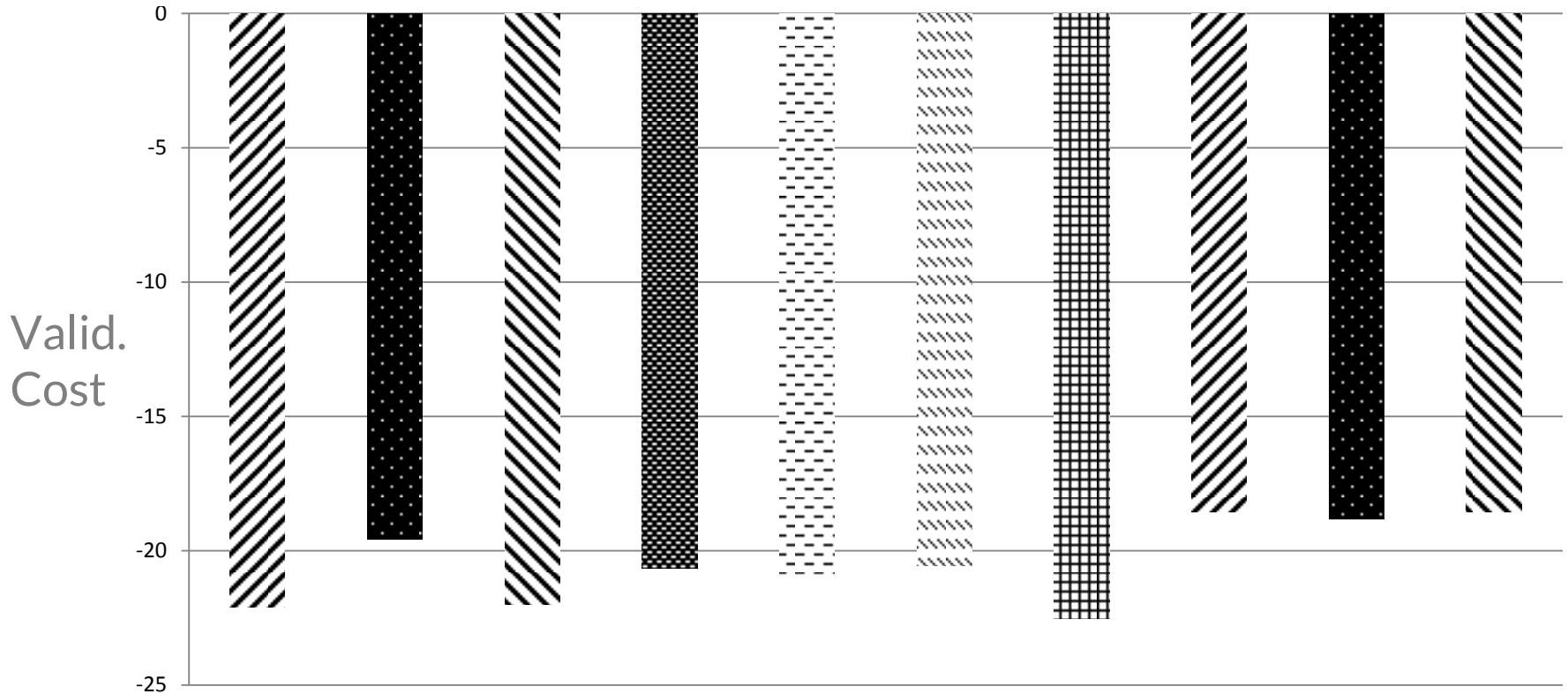
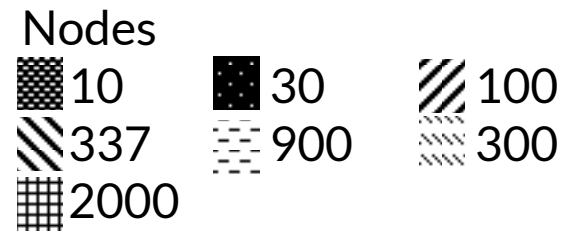
Parameter Initialisation:

$$-4 \frac{\sqrt{6}}{fan\_in + fan\_out}, +4 \frac{\sqrt{6}}{fan\_in + fan\_out}$$

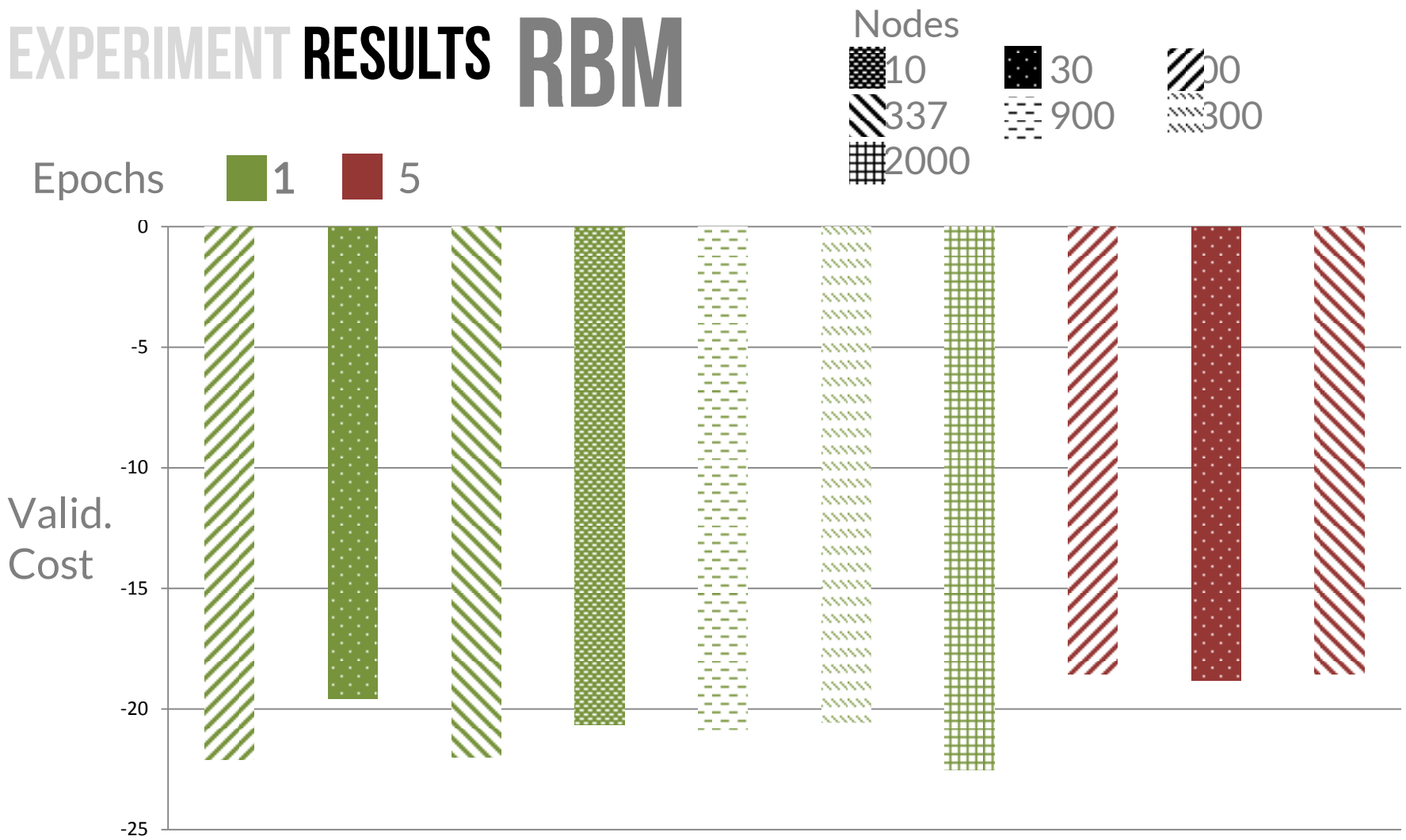
# EXPERIMENT RESULTS RBM



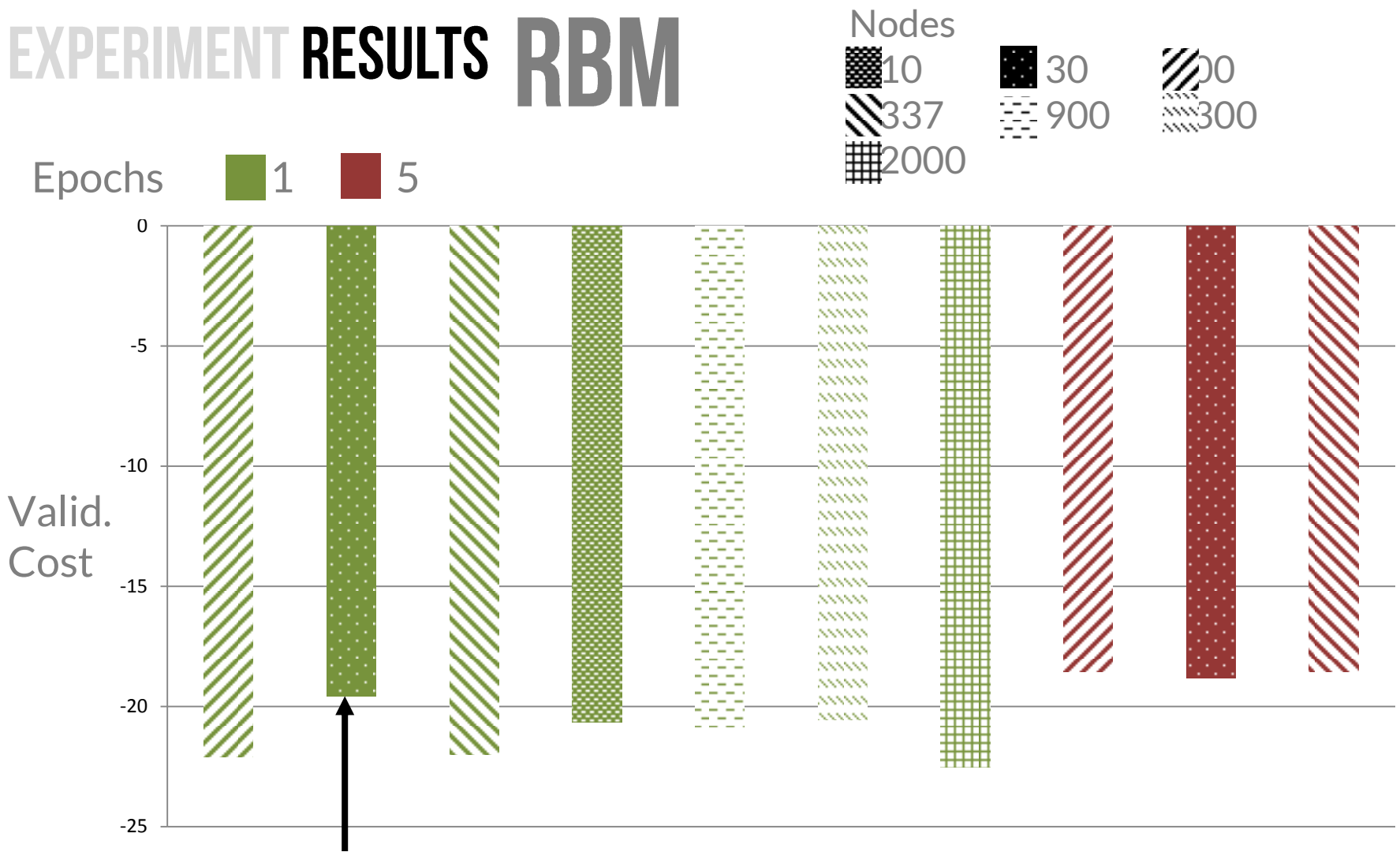
# EXPERIMENT RESULTS RBM



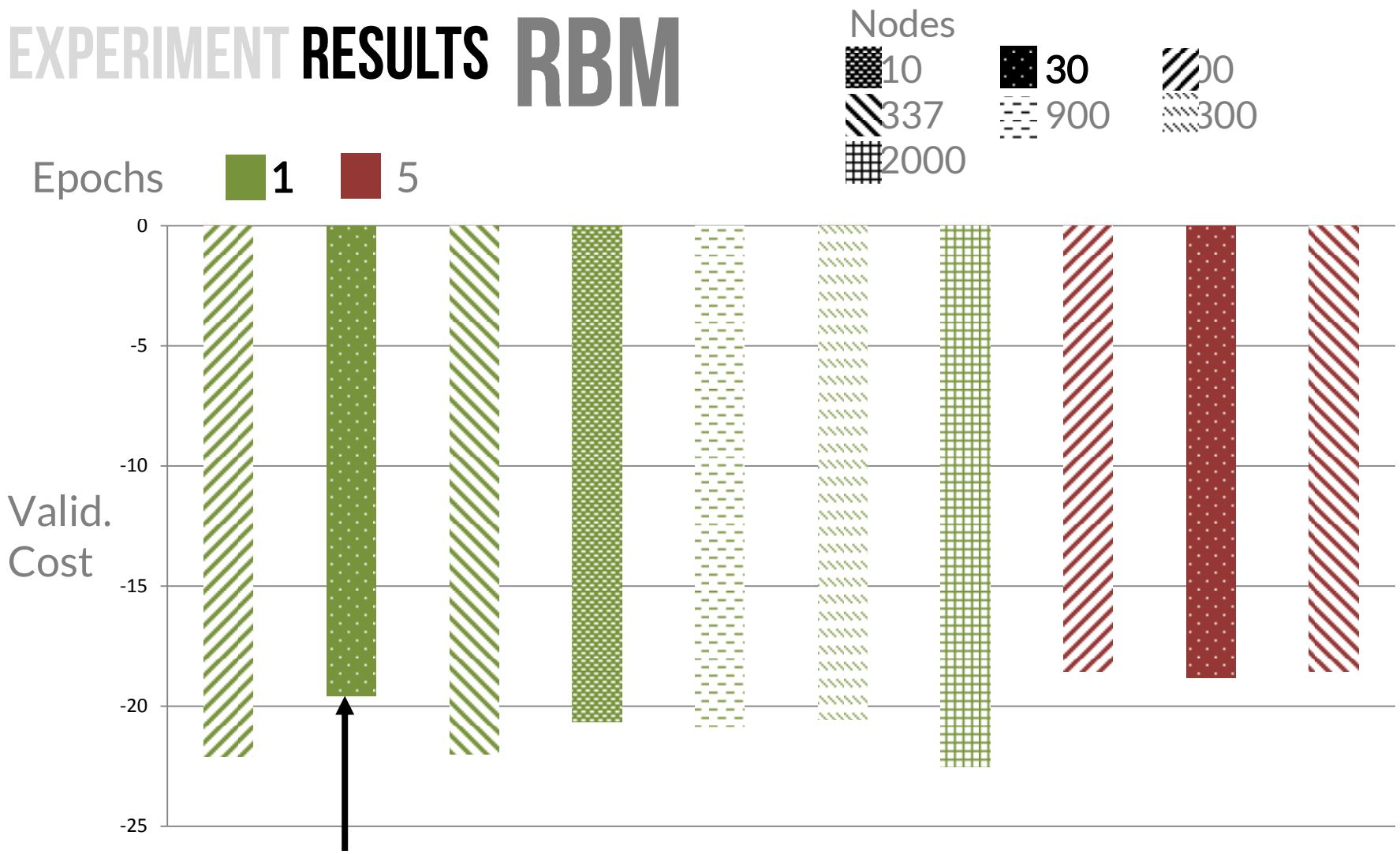
# EXPERIMENT RESULTS RBM



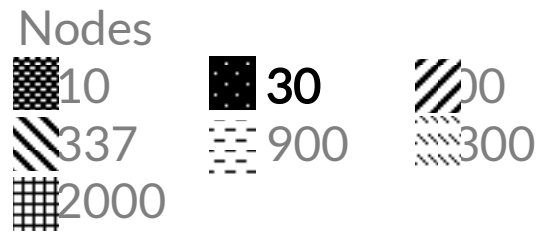
# EXPERIMENT RESULTS RBM



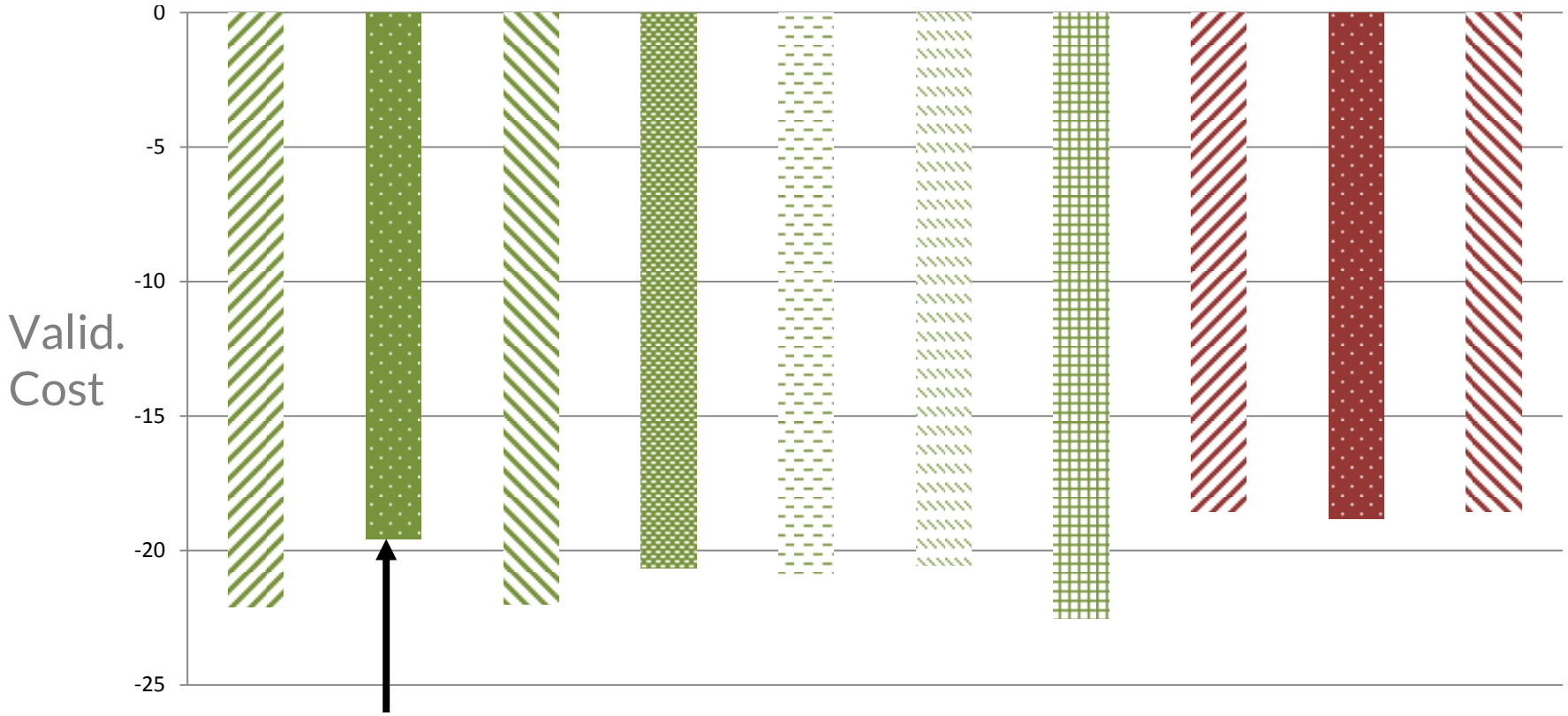
# EXPERIMENT RESULTS RBM



# EXPERIMENT RESULTS RBM



Epochs  1  5



**MONITORING COST**

**-19.580**



## EXPERIMENT

**MLP**

To Choose:

Last layer nodes  $h^{(1)}_n$





## EXPERIMENT

## MLP

To Choose:

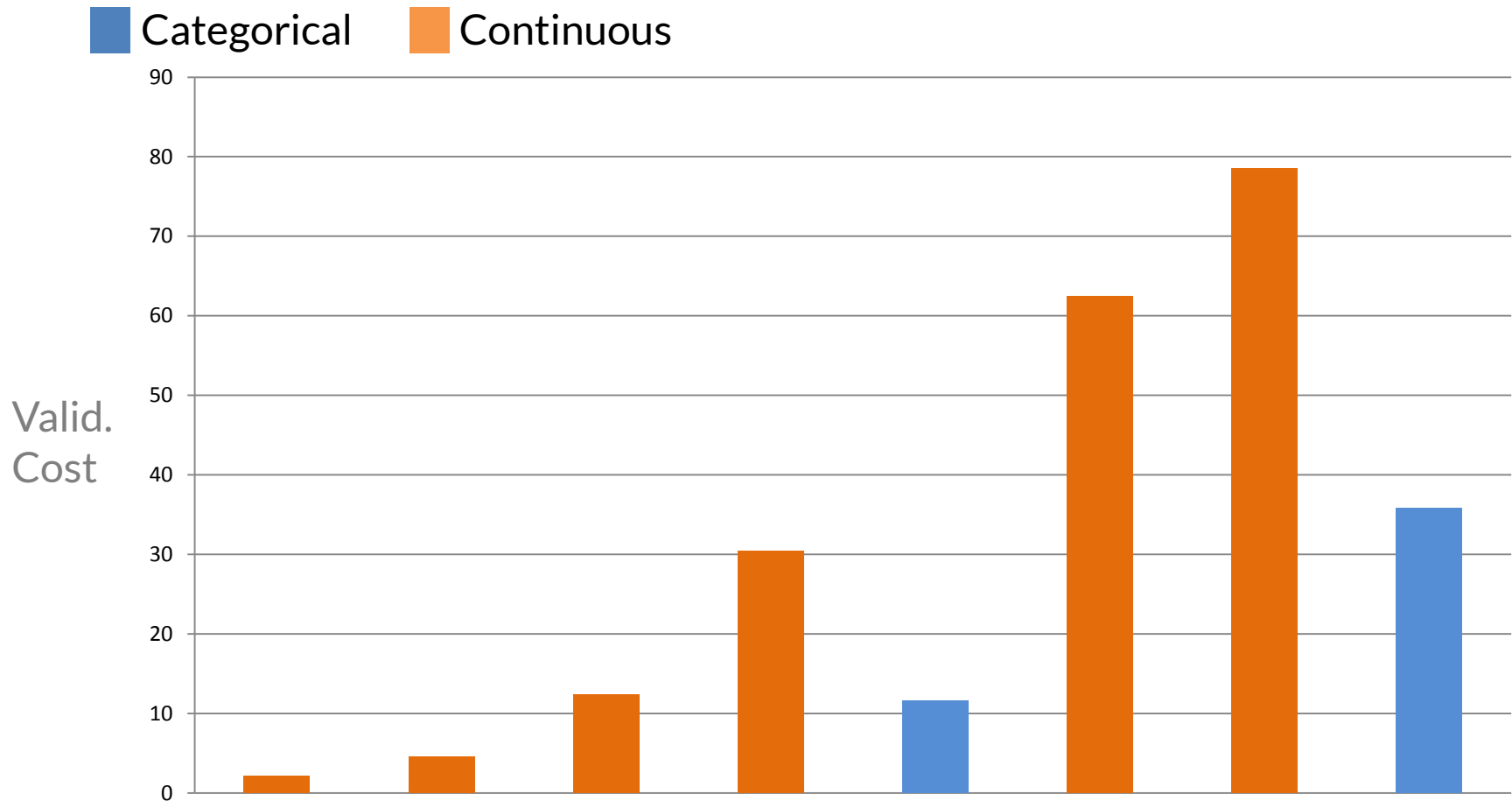
Last layer nodes  $h_n^{(1)}$

The Grid:

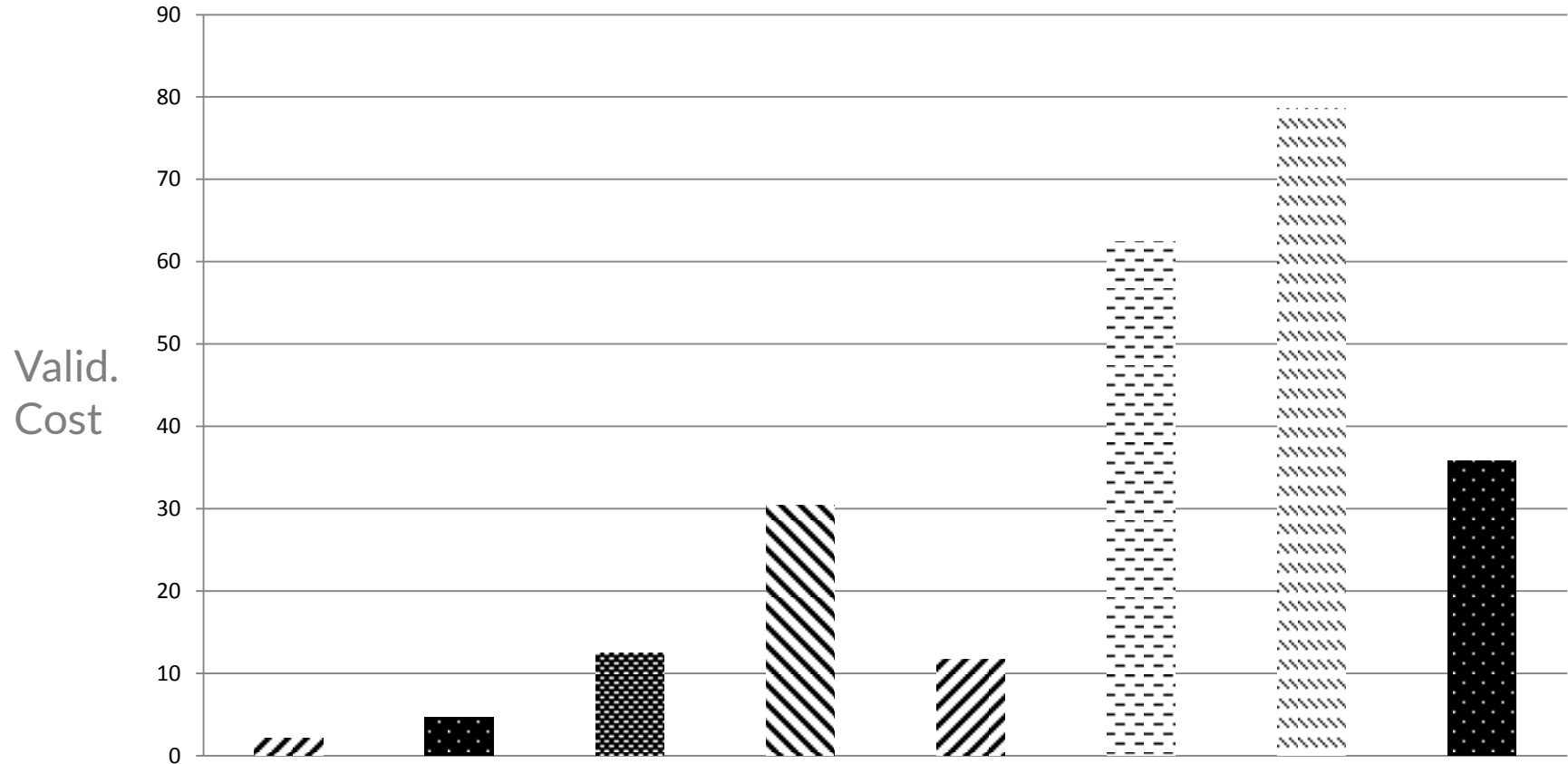
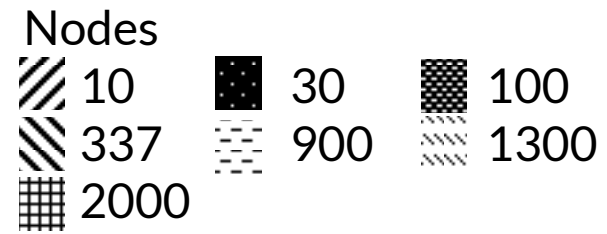
[10, 30, 337, 900, 1300, 2000]



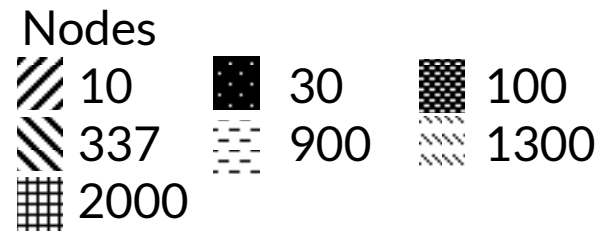
# EXPERIMENT RESULTS MLP



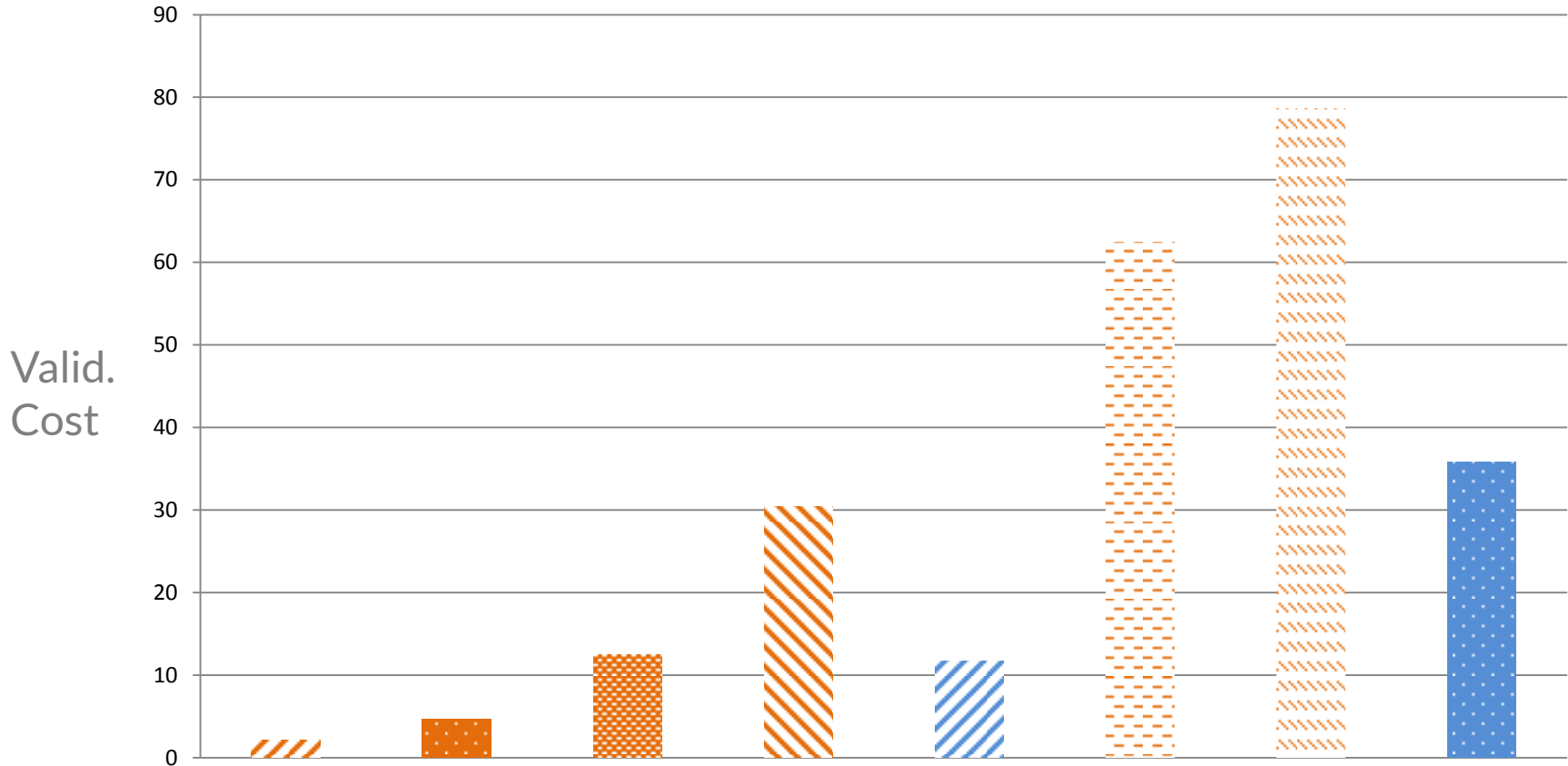
# EXPERIMENT RESULTS MLP



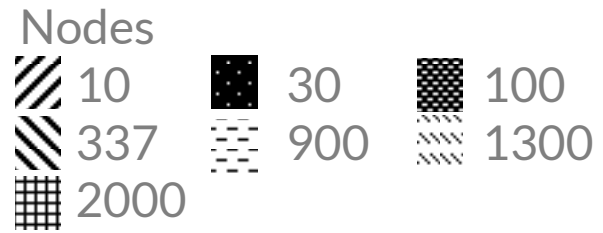
# EXPERIMENT RESULTS MLP



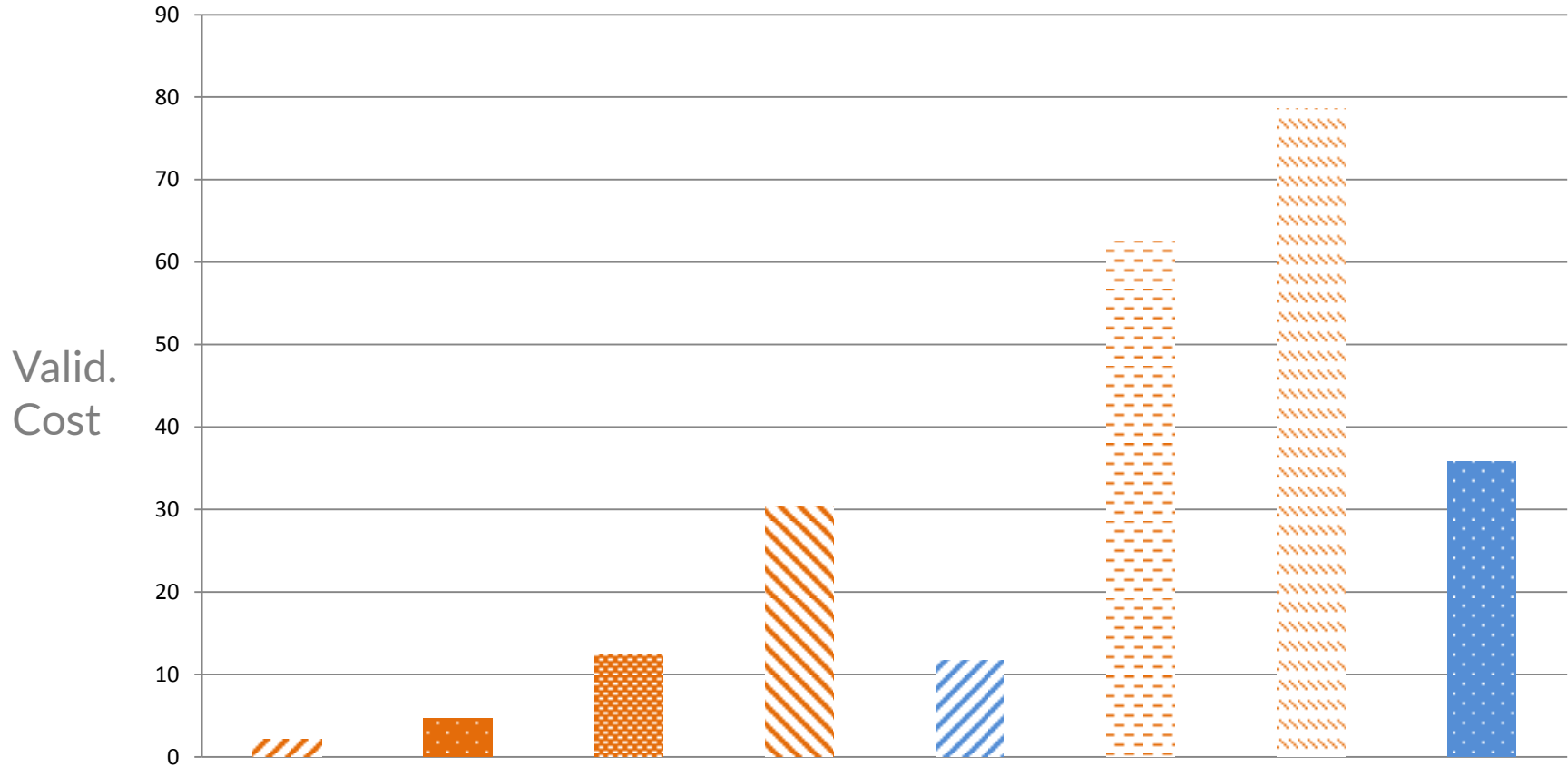
 Categorical  Continuous



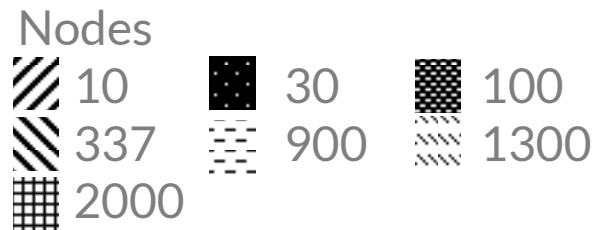
# EXPERIMENT RESULTS MLP



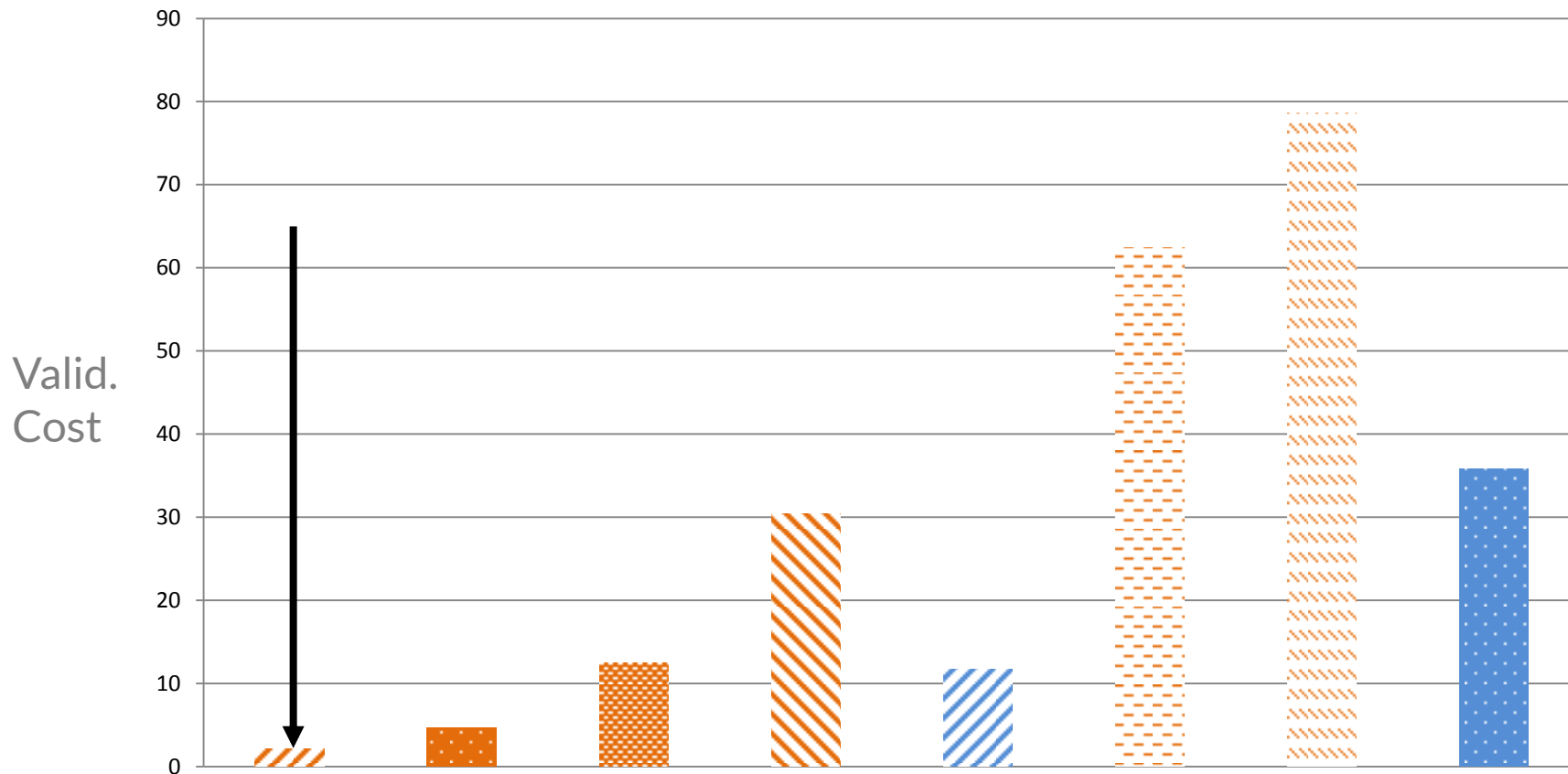
 Categorical  Continuous



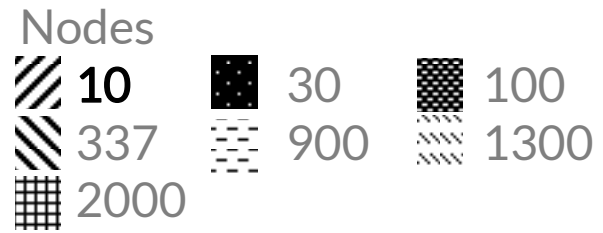
# EXPERIMENT RESULTS MLP



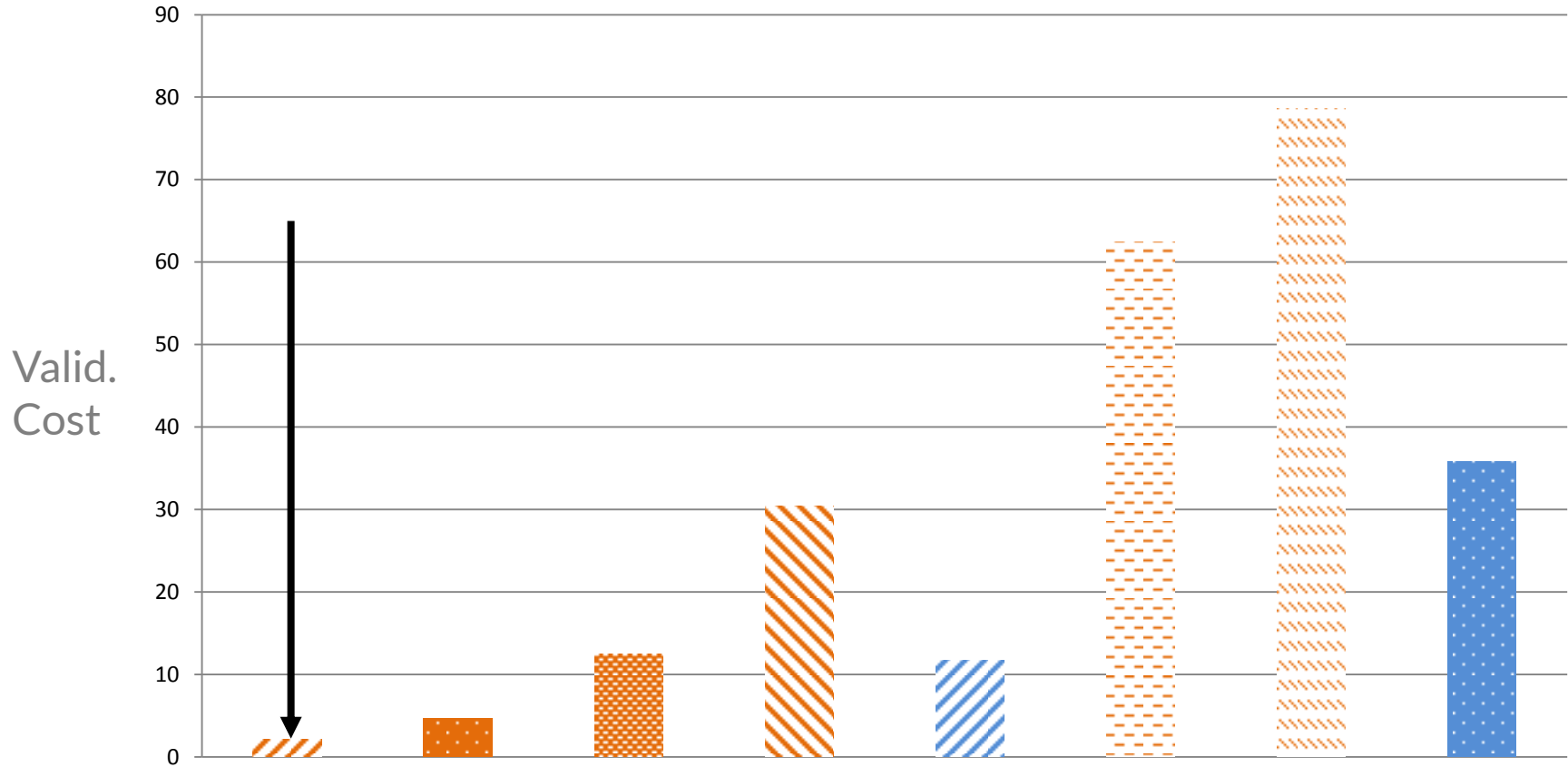
 Categorical  Continuous



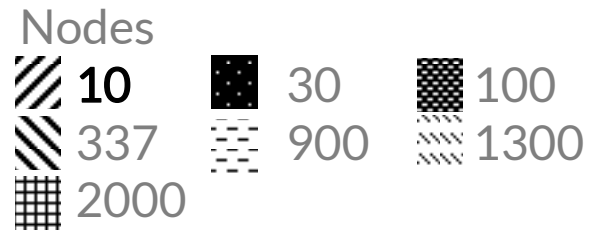
# EXPERIMENT RESULTS MLP



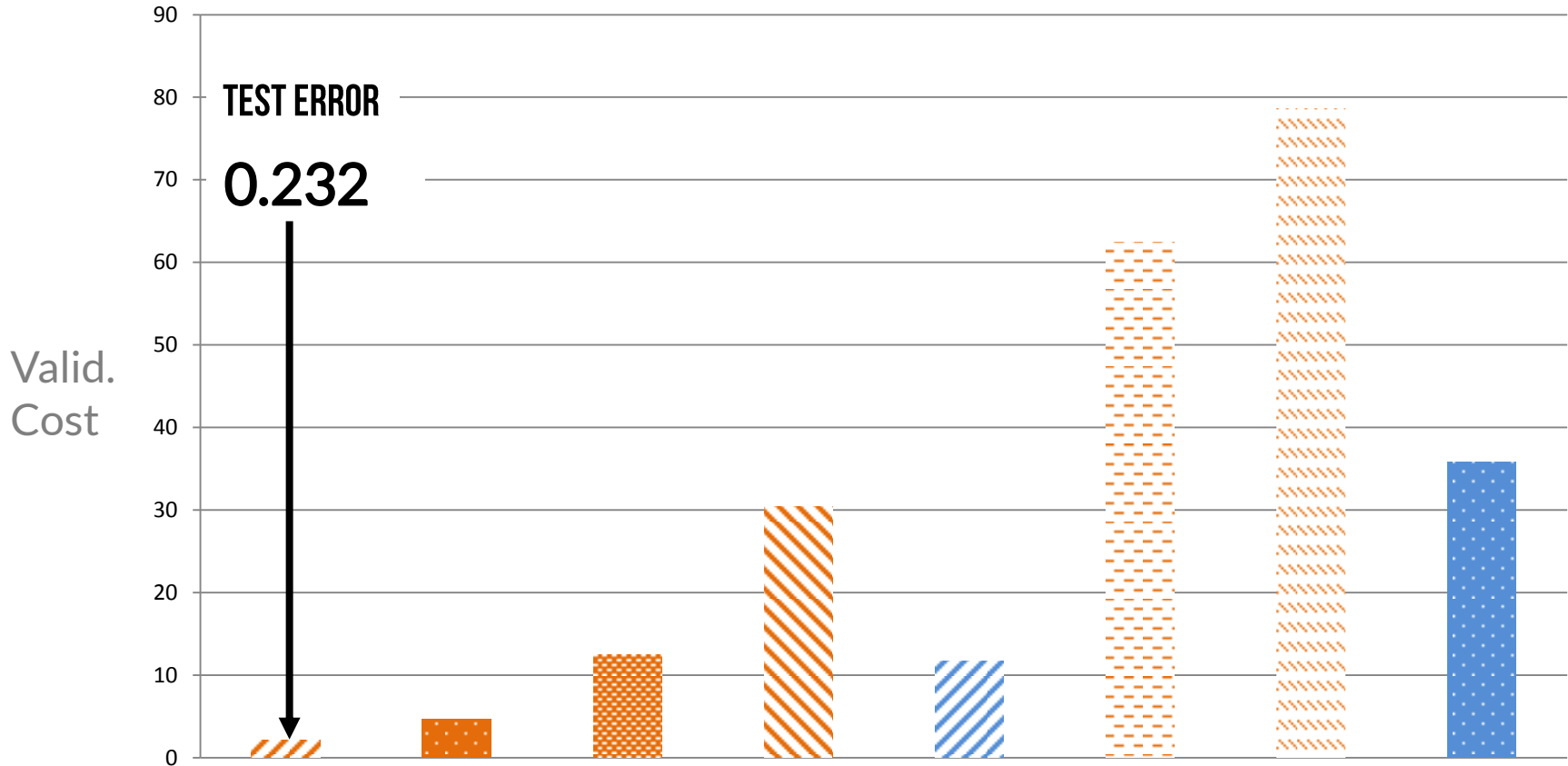
 Categorical  Continuous



# EXPERIMENT RESULTS MLP

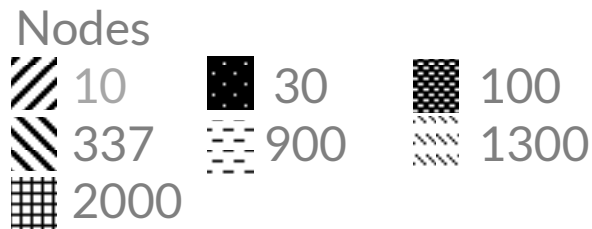


 Categorical  Continuous

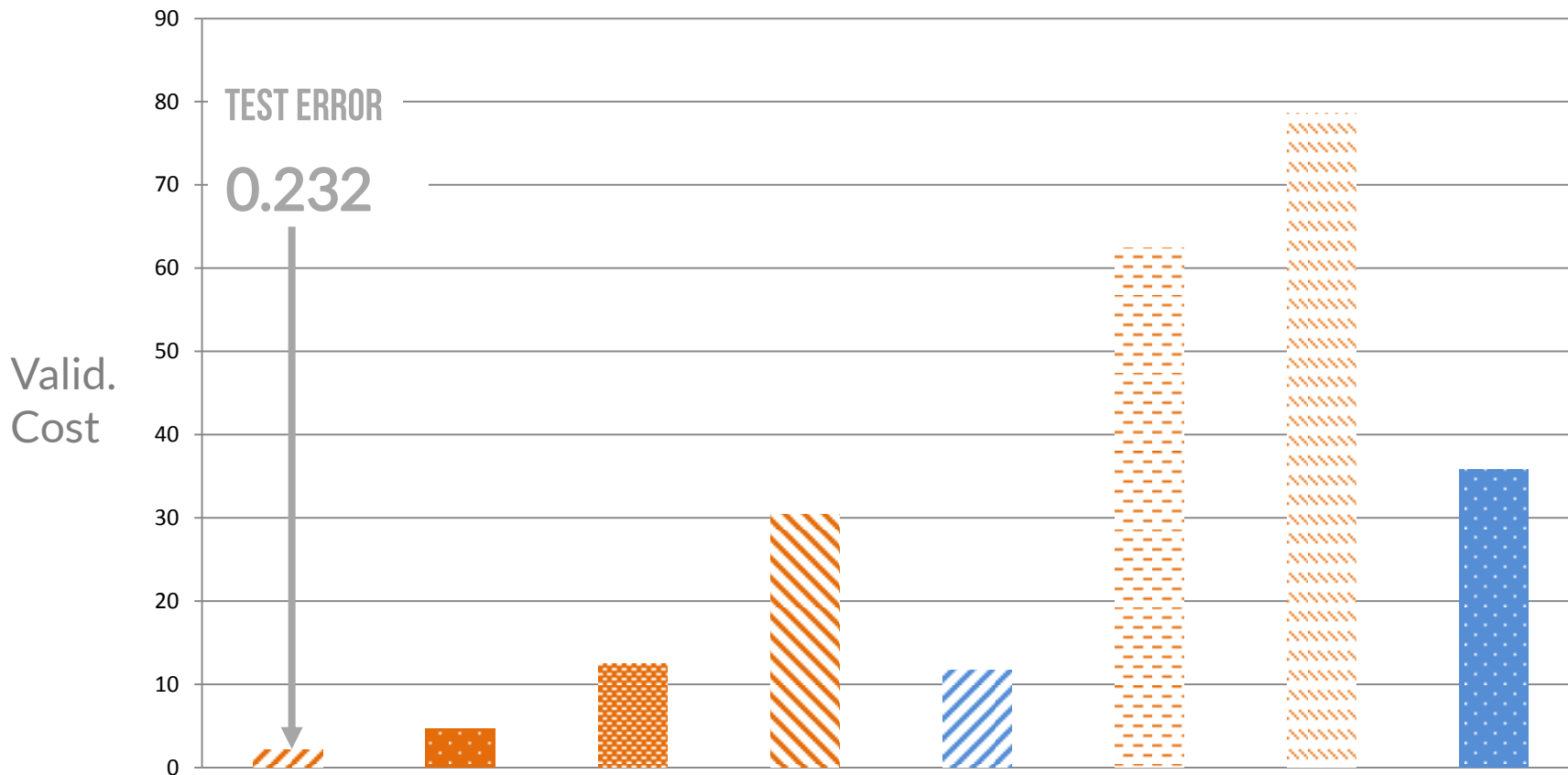




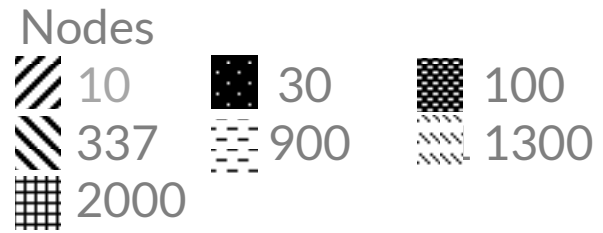
# EXPERIMENT RESULTS MLP




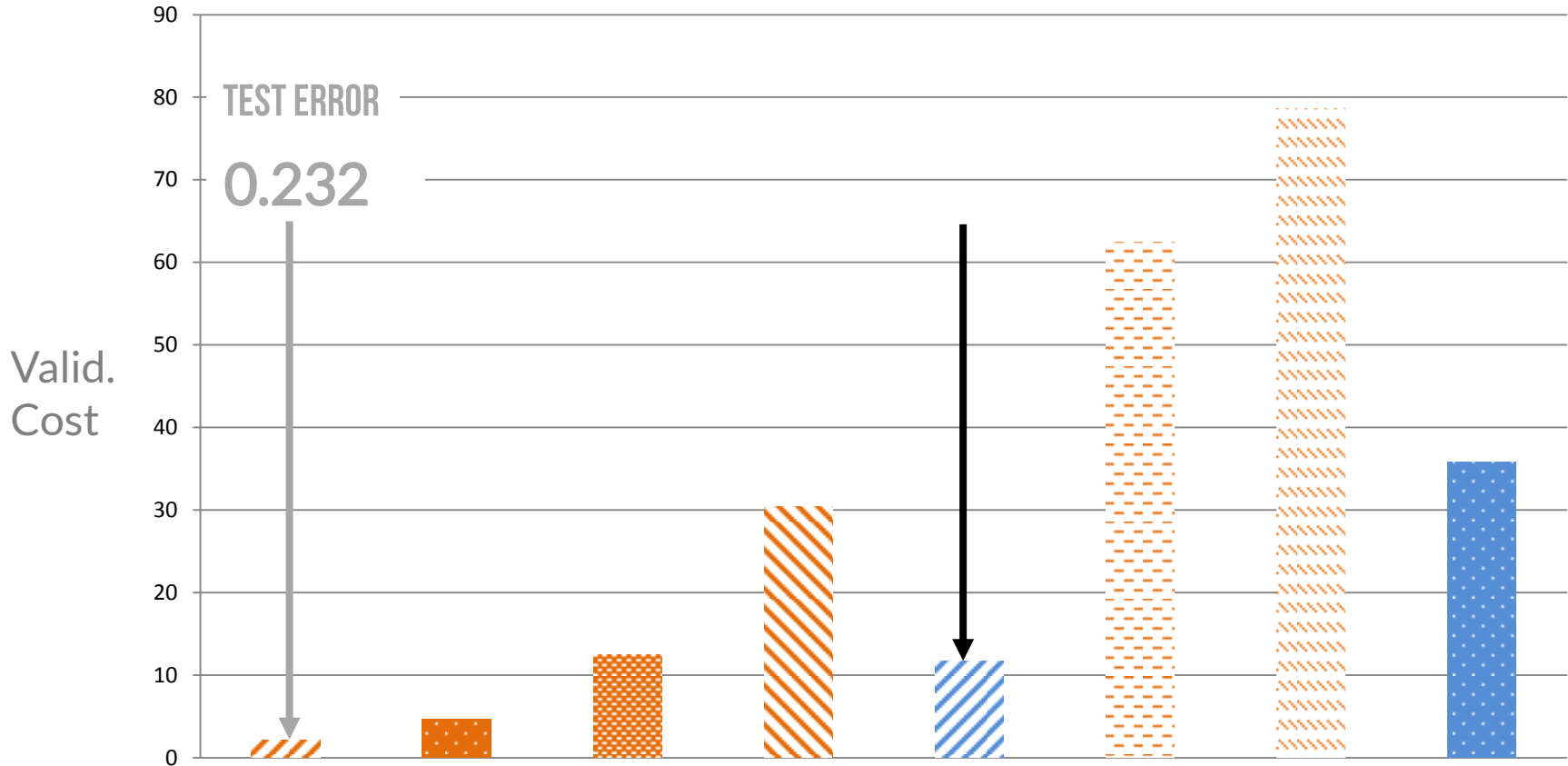
 Categorical  Continuous



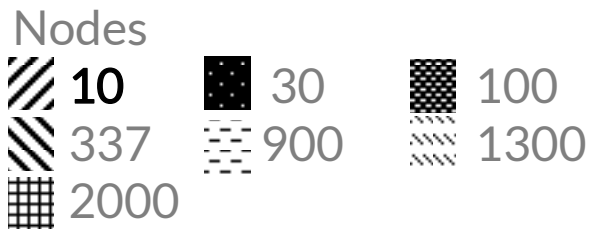
# EXPERIMENT RESULTS MLP



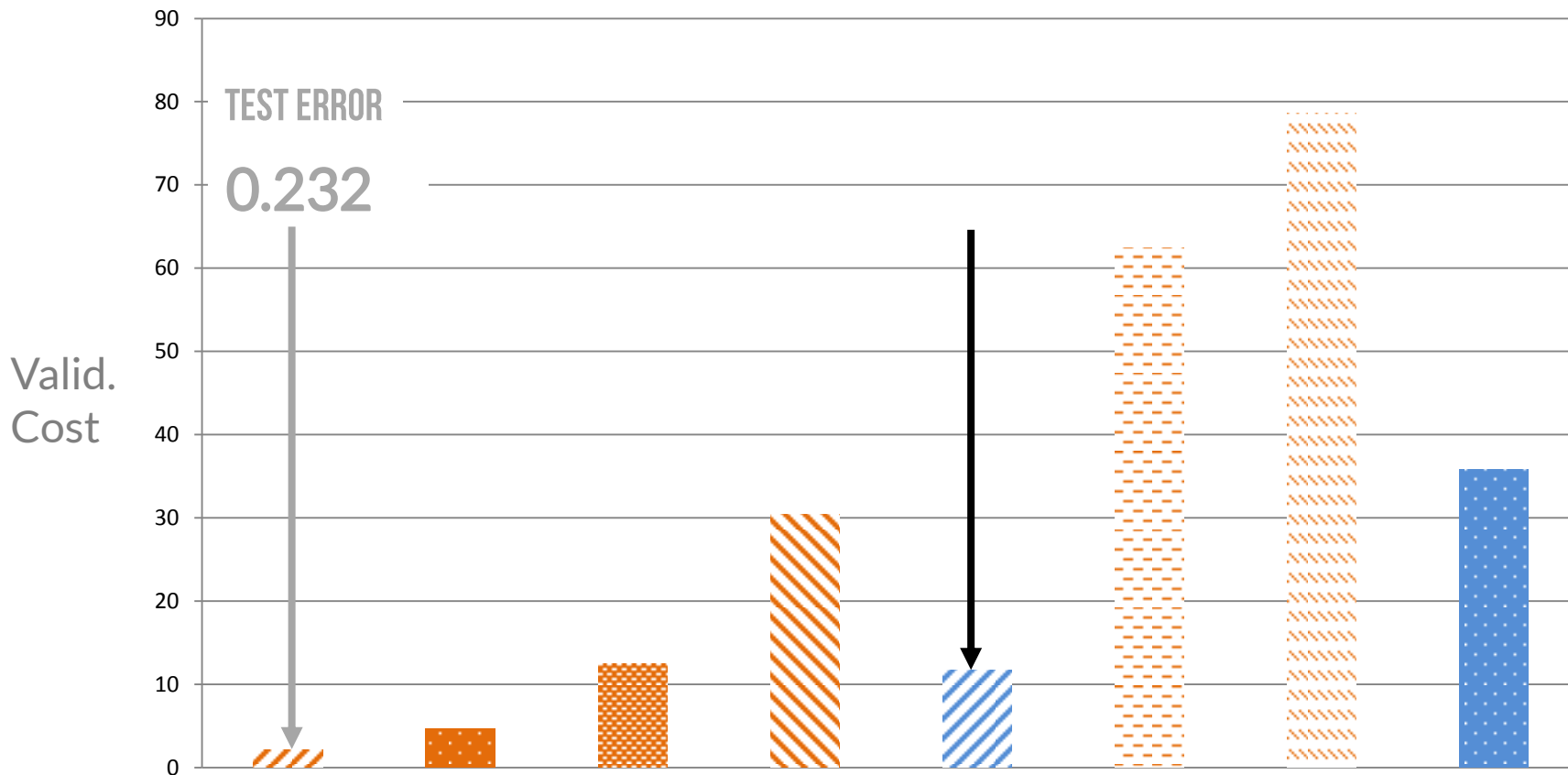
 **Categorical**     **Continuous**



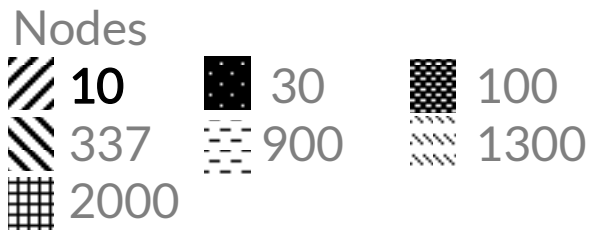
# EXPERIMENT RESULTS MLP



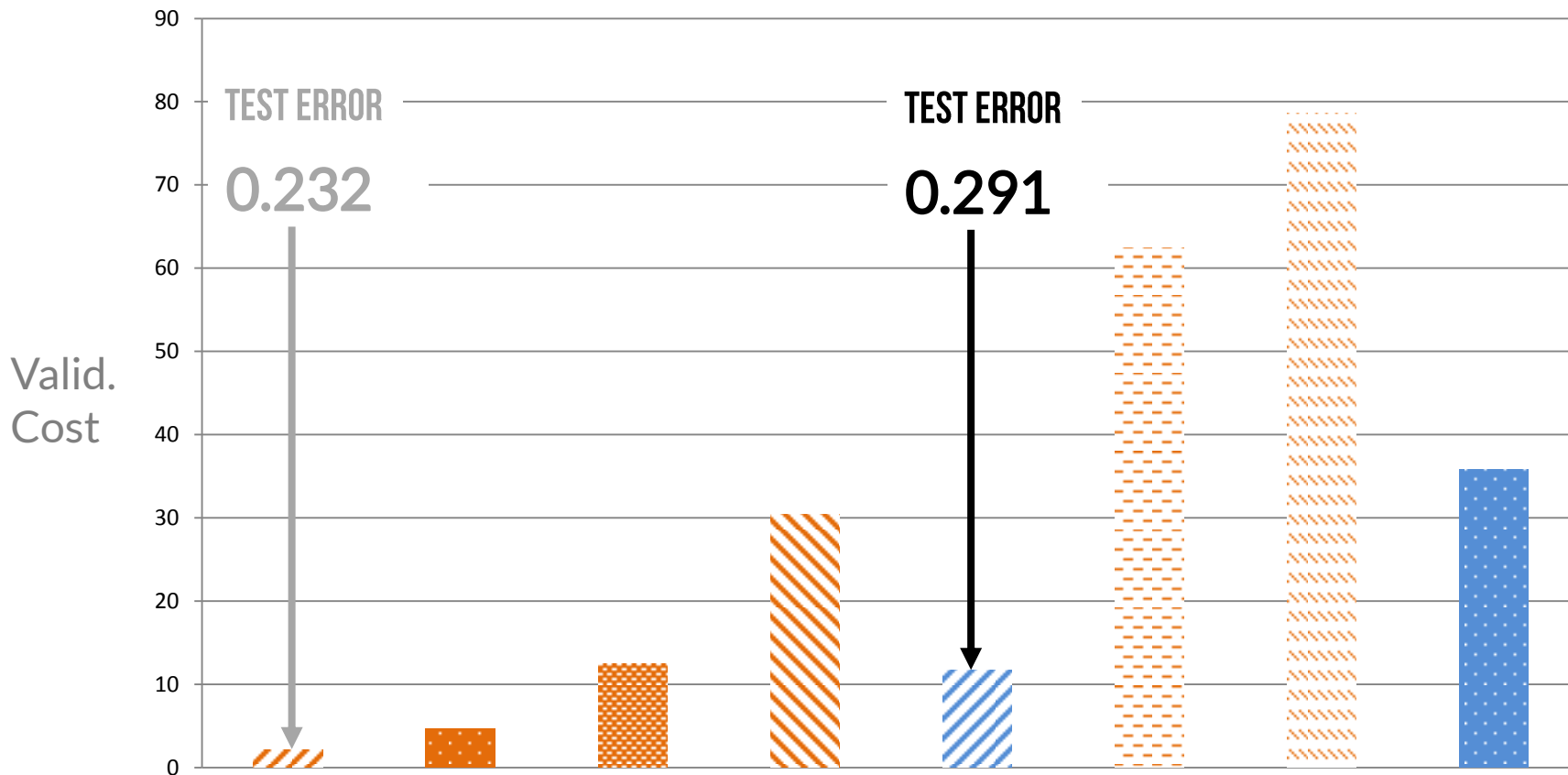
 Categorical  Continuous



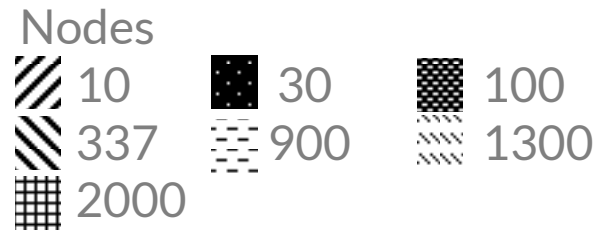
# EXPERIMENT RESULTS MLP



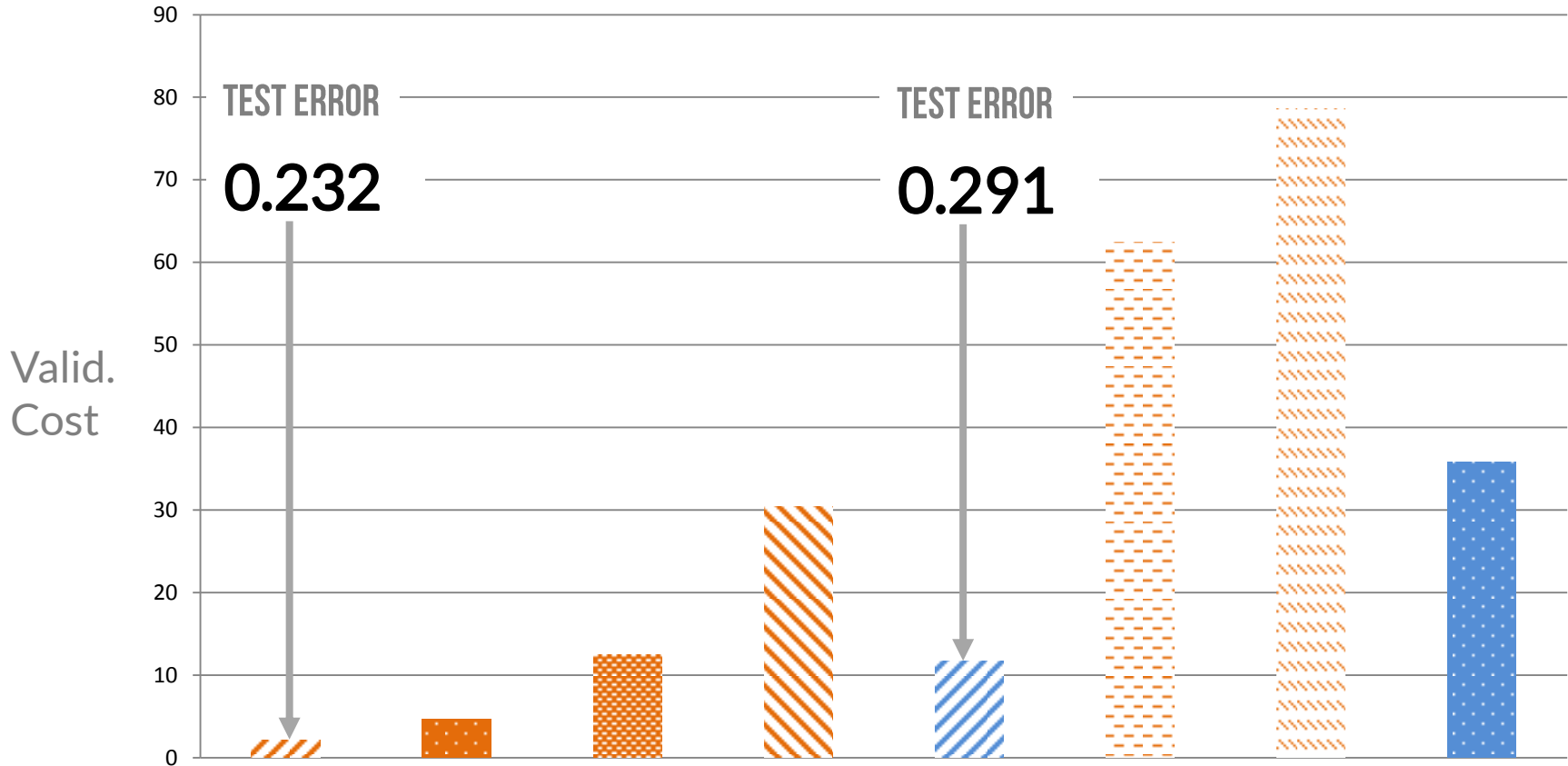
 Categorical  Continuous



# EXPERIMENT RESULTS MLP



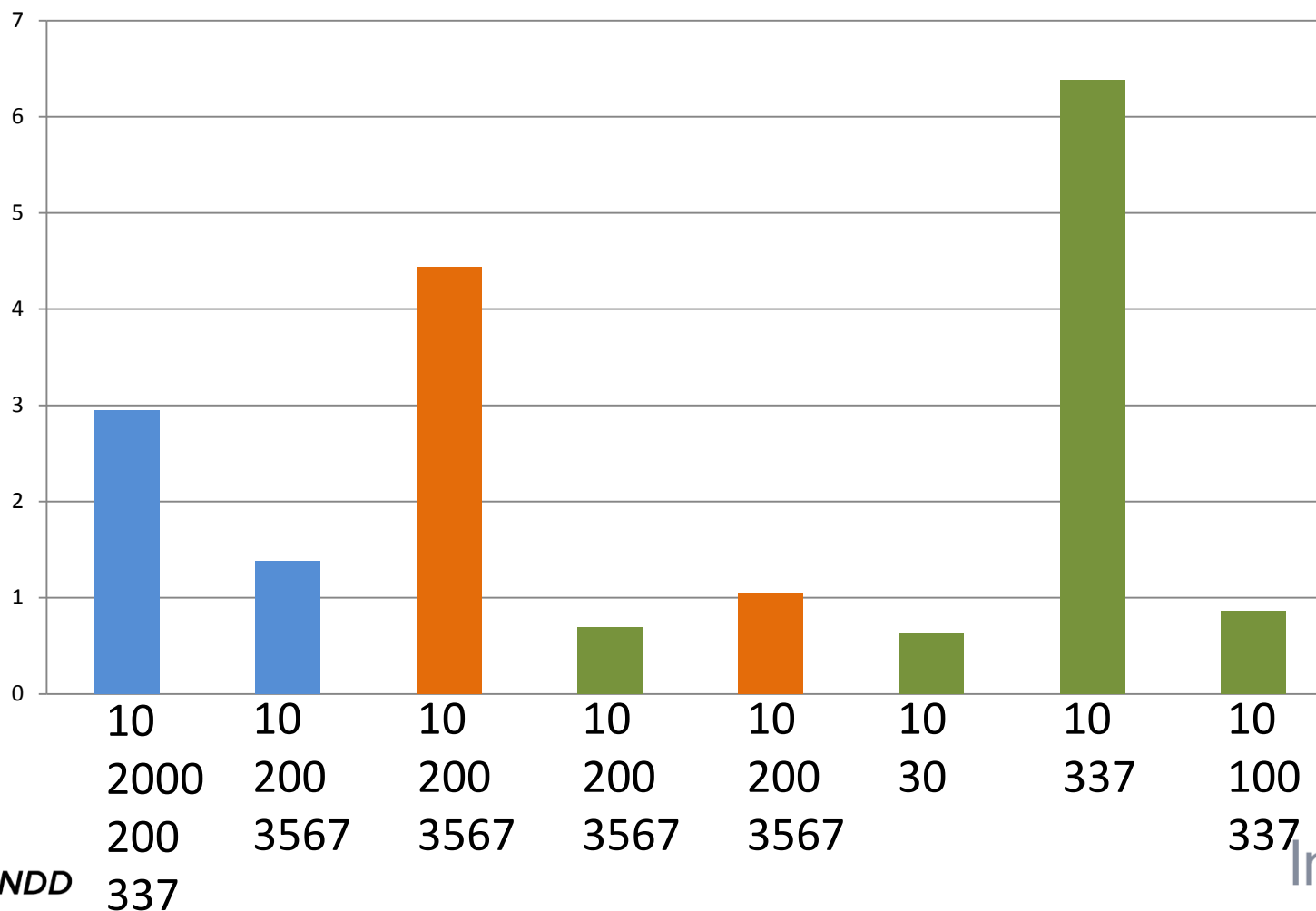
 Categorical  Continuous



# EXPERIMENT RESULTS DBN

Lambda @ 0.03

Alpha ■ 0.001 ■ 0.01 ■ 0.9



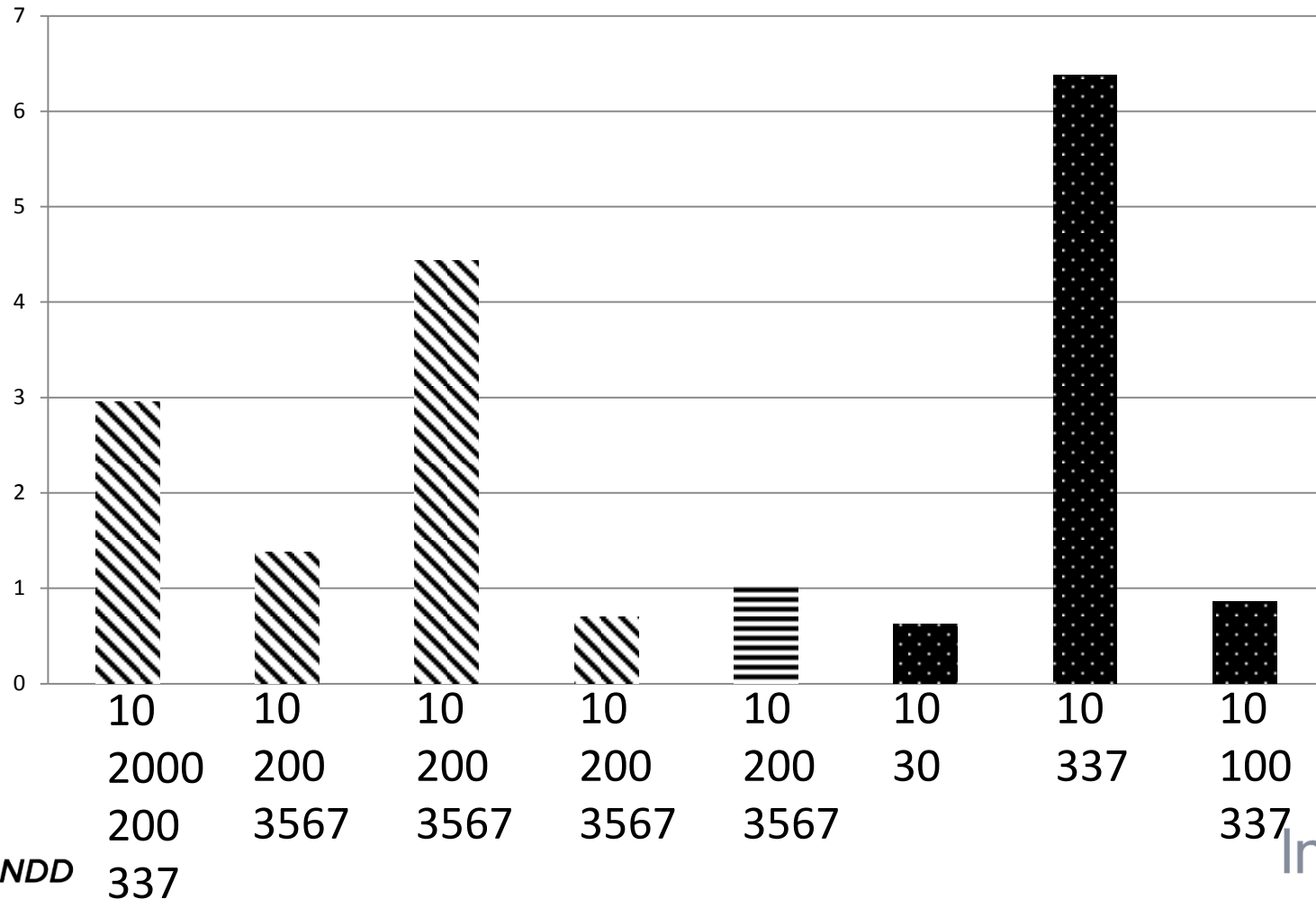
In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# EXPERIMENT RESULTS DBN

Lambda @ 0.03

Step  3000  1000  100





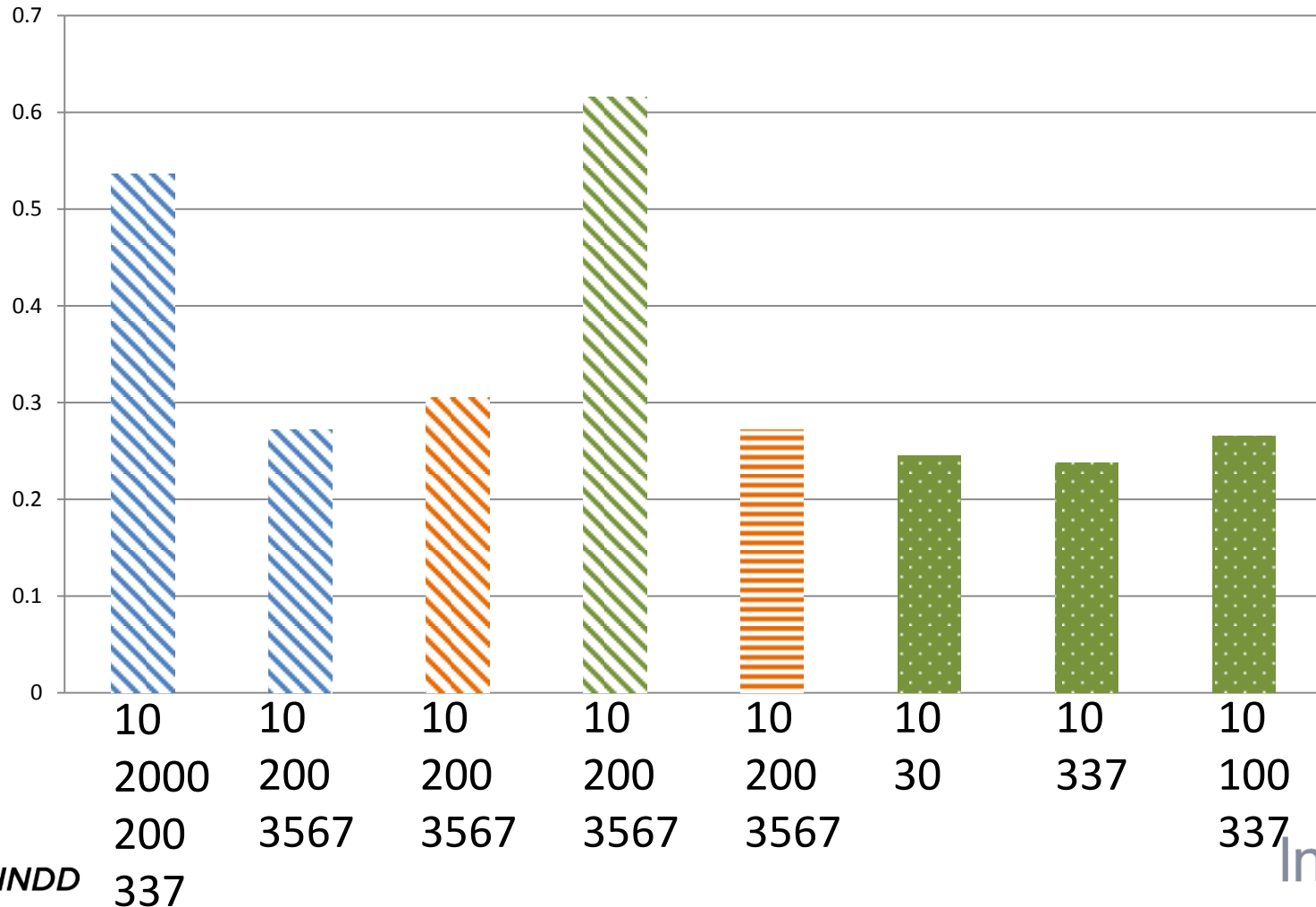


# EXPERIMENT RESULTS DBN

Lambda @ 0.03

Alpha ■ 0.001 ■ 0.01 ■ 0.9

Steps  3000  1000  100



In-MINDD is funded under the European Union Seventh Framework Programme, Grant Agreement Number 304979



# FUTURE WORK

Activation functions

Algorithms

Inference

Framework – to Mongo and input from

Visualising learning

Implementing Early Stopping

Mini-batch Stochastic Gradient Descent



# EXPERIMENTS CONCLUSIONS

Much easier to model when you have one extensible network that can handle many type of data

Constituent models can be used to select a starting point for deep learning configurations



# QUESTIONS?

# EXPERIMENT RESULTS DBN

Lambda @ 0.03

Alpha ■ 0.001 ■ 0.01 ■ 0.9

Steps 3000 1000 100

