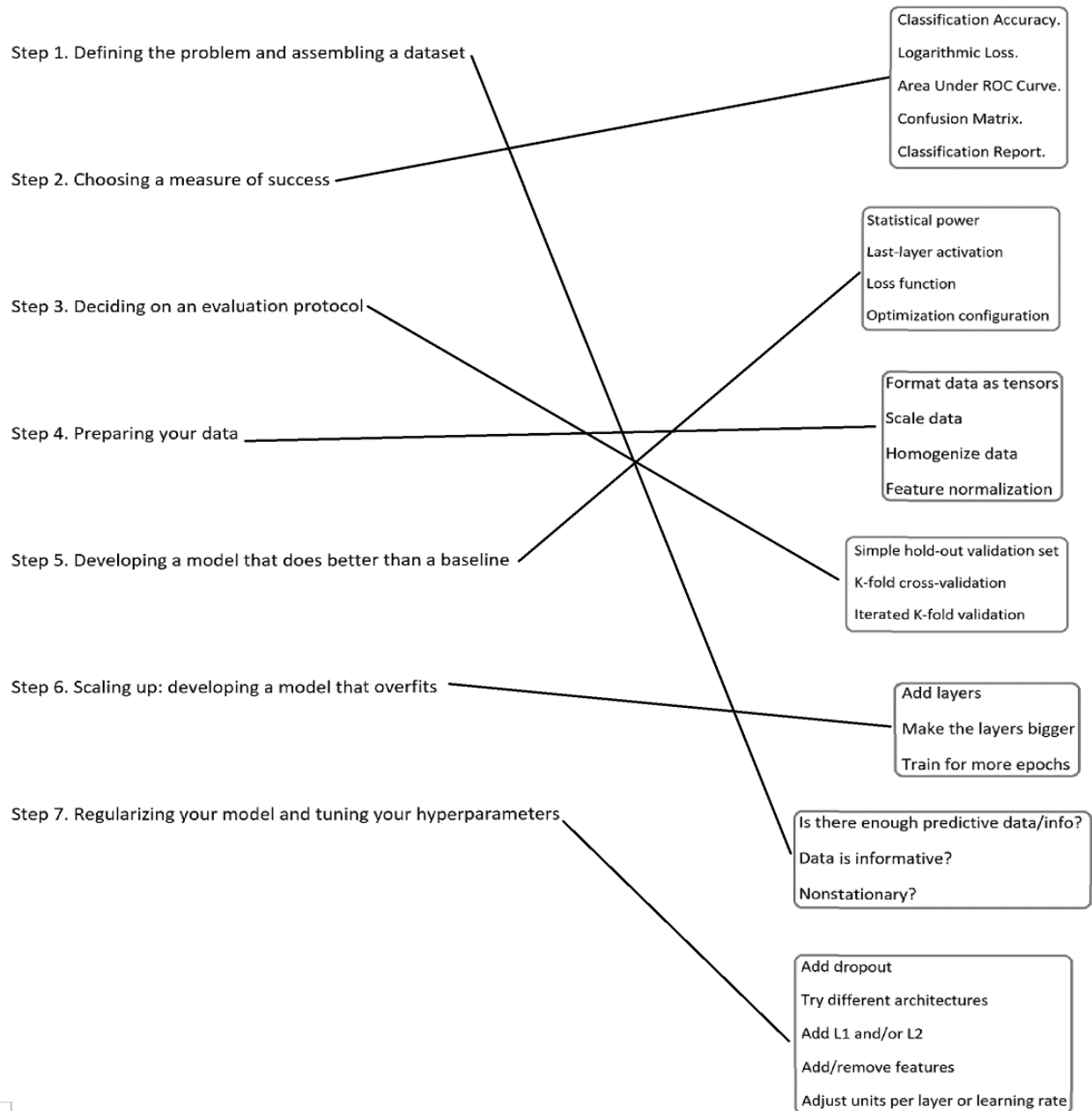


Chapter 4 Pt. 1 ANSWER KEY
Fundamentals of Machine Learning

1) (111-115)



2) (114)

Problem type	Last-layer activation	Loss function
Binary classification	sigmoid	binary_crossentropy
Multiclass, single-label classification	softmax	categorical_crossentropy
Multiclass, multilabel classification	sigmoid	binary_crossentropy
Regression to arbitrary values	None	mse
Regression to values between 0 and 1	sigmoid	mse or binary_crossentropy

3) (100)

c. Redundancy in your data

b. The arrow of time

a. Data representativeness

4) (101) e. All the above

5) (101) b. Tensors of floating-point data or integers

6) (102) b. Normalization

7) (101-102) b. Normalization

8) (102) e. All the above

9) (102-103) e. None

10) (102-103) e. All the above

11) (104) b. overfitting

12) (104) b. In every machine-learning problem

13) (104) d. Generalization refers to how well the trained model performs on data it has never seen before and optimization is the process of adjusting the model to maximize performance on training data

14) (104) b. The machine-learning engineer cannot control

15) (104) c. Get more training data

16) (104) a. The process of fighting overfitting

17) (104) b. Reduce the size of the model

18) (104) a. The number of learnable parameters in a model

19) (106) b. The smaller model overfits more slowly

20) (106) d. Too much capacity

21) (107) d. Simple is better

22) (107) d. All the above

23) (107) a. Adding a cost proportional to the absolute value of the weight coefficients

24) (107) d. b & d

25) (108) a. They cannot be used in the same instance

26) (109) b. Also drop units during test time

27) (110) a. After the output of the layer and before the next

28) (110) f. All the above