

# STINTSY / MACHLRN – Midterm Exam Reviewer

AY 2024 – 2025 Term 2

1. What is machine learning?
2. What is the difference between machine learning and traditional AI algorithms?
3. What is the difference between supervised and unsupervised learning?
4. What are instances? What are features? What are labels?
5. How do we train a machine learning model?
6. How do we determine which instances are the nearest neighbor to a test instance in the kNN algorithm?
7. How do we perform kNN for classification? for regression?
8. What are the distance metrics used in kNN?
9. How do we select the value of k in kNN?
10. What is the difference between classification and regression?
11. Where do we use linear regression?
12. What is the difference between parameters and hyperparameters?
13. Why do we divide the dataset into training, validation, and testing?
14. What is the use of a hypothesis function?
15. What is the use of a loss / objective function?
16. What are weights? What do they indicate?
17. How do we interpret the value of the loss function of linear regression?
18. What is the difference between the analytical solution and gradient descent?
19. What is the difference between gradient descent, stochastic gradient descent, and mini-batch gradient descent?
20. Why do we compute the gradient?
21. What is the effect of the value of the learning rate in training a model?
22. How do we update the weights of any model?
23. How do we train a linear regression model?
24. How do we compute the loss function for a linear regression model?
25. How do we perform inference using a linear regression model?
26. What is the bias error? What is the variance error?
27. What characterizes an underfitted model? What characterizes an overfitted model?
28. Why do we increase the degree of the input values?
29. How do we solve a bias error in a model? How do we solve a variance error in a model?
30. What is the effect of doing regularization to a model?
31. What is the difference between ridge and lasso regression?
32. How do we train a logistic regression model?
33. How do we compute the loss function for a logistic regression model?
34. How do we perform inference using a logistic regression model?
35. How do we represent the y of an instance in a multinomial logistic regression model?
36. How do we train a neural network?
37. How do we compute the loss function for a neural network?
38. How do we perform inference using a neural network?
39. What are intermediate features in a neural network?
40. What are the different parts of a neural network?
41. How do we design the input layer? output layer?
42. How do we set the number of layers and neurons in the hidden layers?
43. How do we select the appropriate activation functions per layer?
44. What scenario should we use a sigmoid / tanh / softmax / ReLU / leaky ReLU activation function?
45. How do we perform forward propagation? backward propagation?

Don't forget to review your notebooks. Use our in-person quizzes to review for potential solving problems.

Formula List provided in the exam.

$h(x) = \theta^T x = X\theta$	$-\ell(\theta) = -\frac{1}{n} \sum_{i=1}^n \left[ y^{(i)} \log \left( h(x^{(i)}; \theta) \right) + (1 - y^{(i)}) \log \left( 1 - h(x^{(i)}; \theta) \right) \right]$
$\mathcal{L}(\theta) = \frac{1}{2n} \sum_{i=1}^n \left( h(x^{(i)}) - y^{(i)} \right)^2$	$h(x) = \text{softmax}(\theta^T x) = \frac{e^{\theta^T x}}{\sum_{j=1}^k (e^{\theta^T x})^{(j)}}$
$\theta := \theta - \alpha \cdot \frac{\delta \ell(\theta)}{\delta \theta}$	$-\ell(\theta) = -\frac{1}{n} \sum_{i=1}^n y^T \log \left( h(x^{(i)}; \theta) \right)$
$h(x) = \sigma(\theta^T x) = \frac{1}{e^{-\theta^T x} + 1}$	