Section A – Data Cleaning

No	Question	Answer
1	When merging datasets with different measurement systems, ensuring _____ between units is crucial.	
2	In data clearing "___in ____ out" stresses that analysis quality depends on input data.	
3	_____ are systematic problems introduced during the data-cleaning process that can be corrected.	
4	Irrecoverable information loss during acquisition (e.g., power outage) is a(n) _____.	
5	Replacing missing data with the mean of existing values is called _____ imputation.	
6	Filling missing values via linear regression or curve fitting is known as _____.	
7	Always preserve both the _____ data and the cleaned dataset.	
8	Converting "John Smith" to "smith, john" is an example of name _____.	
9	Outlier detection is often used as a "_____ check."	
10	Functions like *unidecode()* help remove inconsistent characters during _____ repair.	

Section B – Machine Learning Concepts

No	Question	Answer
11	Transforming raw variables into more informative ones is called _____.	
12	Predicting a numeric value such as salary is an example of _____.	
13	A model that fits training data perfectly but fails on new data suffers from _____.	
14	Dropout and weight penalties are examples of _____ methods.	
15	Training multiple diverse models to improve accuracy is known as _____ learning.	
16	A simple reference model for comparison is called a _____ model.	
17	Independent, unchanging data samples are described as _____ data.	
18	Parameters tuned before training are called _____.	
19	Comparing two versions statistically to see which performs better is _____ testing.	
20	Normalization typically rescales numeric features to the range _____.	
21	The difference between train and test performance measures model _____.	
22	Biased sampling in data collection can cause _____ bias.	

23	A model trained once and not updated further is a _____ model.	
24	A model that continually retrains with new data is a _____ model.	
25	The loss function measuring divergence between predicted and true probabilities is _____.	

Section C – Deep Learning

No	Question	Answer
26	Deep Learning is a specialized subset of _____.	
27	Neural Networks are universal _____ of functions.	
28	A network with more than one hidden layer is termed a _____ NN.	
29	Each layer's output becomes the next layer's _____.	
30	CNNs are mainly used for _____ classification.	
31	RNNs handle data with strong _____ dependencies.	
32	LSTM is an improved version of a _____ network.	
33	In a GAN, the second model acts as a _____ distinguishing real and fake data.	
34	Reusing a pre-trained model for a related task is called _____ learning.	
35	The "wide" part of a Wide & Deep model helps _____ specific examples.	
36	The "deep" part captures more _____ features.	
37	Linear classifiers combine inputs with _____ and _____ to form predictions.	
38	Activation functions introduce _____ into neural models.	
39	Adding neurons or hidden layers increases model _____.	
40	Transfer Learning leverages _____ features learned from prior tasks.	

Section D – Python Data Structures

No	Question	Answer
41	Lists are ordered and _____, meaning they can be changed.	
42	Tuples differ from lists because they are _____.	
43	Dictionaries store data as key–_____ pairs.	
44	Sets contain only _____ elements.	
45	A *deque* allows fast operations at _____ ends.	
46	*Counter* is a subclass of _____ that tallies occurrences.	
47	A *frozenset* is a set that is _____.	
58	The *heapq* module implements a _____ queue.	
49	In a max-heap, each parent node's value is _____ than its children's.	
50	Converting a list into heap order is called _____.	

Roll Number _____ **Name** _____