Interview 1:

1) How did you train the custom model?
2) Elaborate on the project?
3) Why did you choose the process?
4) Can you explain the architecture?
5) Math under architecture?
6) Can you think of a better model for the project?
7) Brief me about logistic regression? (back end, the math involved, formulae)
8) Default probability when you say the class is positive in logistic regression?
9) How to change the threshold? What performance metrics can be used? (answer: we will consider ROC and AUC curve)
10) How do we test the performance of an algorithm?
11) Kappa score?
12) Have you heard about imbalanced data?
13) Which parameters to be considered to declare it as a good model? (precision, recall, f1 score, accuracy, etc..)
14) Which parameter would you consider, when you are predicting the stock market? (answer: we need to try to reduce false negative)
15) Any other algorithms?
16) Any other knowledge of Statistics?
17) Equation of logistic regression?
18) Scenario-based Question.
    There are 10 features, independent features, there is one independent feature. In the dependent feature, you have o/p of yes/no. One more dependent feature(second dependent) if o/p is zero there are multiple categories concerning that 0,1,2,3 and if i/p data set is one then categories is 4,5,6. And data is an unbalanced data set. How will you build the model?
    O/P must be 0-6.
    Step 1: Use a binary classification.
    Step 2:
19) How did you choose how many neurons in layers?
20) Are there any libraries that automate to find the best parameters?
21) At a particular time loss is not decreasing, what might be the problem, and how to deal with it?
22) What are the advantages of flutter compared to others?
23) What made you learn DL/ML?
24) Does AI use cases in our interesting things?

Interview 2:

1) Tell me about your projects.
2) Deep learning: what are all the topics you know?
3) What was the data set, What was the flow, what is the model you used for a particular project?
4) Difference between iterator and generator in python with example.
5) What is right skewness? (what is skewness)
6) How do you relate the points with mathematical expression? By using Mean, Median, Mode how do you correlate?
7) What we have to do if our data is skewed?
8) What standard scalar does?
9) What is standardization and what is normalization? (Definition)
10) Difference between vector and scalar with example?
11) What is the direction in the vector?
12) What is the meaning of second-order differential? How is the difference between differentiation?
13) What is the layman's definition of $\frac{dy}{dx}$?
14) What is the slope and what is the scope of the slope?
15) Which machine algorithm are you confident in?

16) What is an ensemble technique?
17) What we will do if we are using an ensemble technique is logistic regression?
18) What is the data selection technique when it comes to random forest classifiers? How can you get a training data set for the random forest?
19) Do we divide the data set into row-wise or column-wise? (In a random forest we can divide in both row-wise and column-wise)
20) What is bootstrapping?
21) What happens in cross-validation?
22) Difference between bootstrapping and cross-validation?
23) What are the parameters in the Random forest classifier?
24) Gini indexing (Definition)
25) In ensemble technique with any other technique?
26) What is the neural network, how it works, why neural networks?
27) Components in neural networks?
28) What is the need for activation function in a neural network?
29) How to choose an activation function?
30) What optimizer does?
31) How to decide the learning rate?
32) How SGD works? What is the working mechanism of the optimizer?
33) Why does the graph fluctuate more in SGD?
(answer: since the single amount of data can be passed using SGD, the graph will be fluctuating
34) Momentum in terms of a neural network?
35) What are the use cases of stemming and lemmatization?
36) Use cases of NLTK?
37) In sentiment analysis, what should be used stemming or lemmatization? Question-answer application?
(In Question-Answer application, if we use stemming we cannot get an accurate word, so we use lemmatization
In Sentiment analysis, we should go for stemming because the importance of the word is mostly put up in some word itself, not in the whole sentence.)
38) The basic difference between lemmatization and stemming.
39) In stemming and lemmatization, which is fast? What is the backend of lemmatization?

(Note: Interviewer don't expect more complex things, No too many technologies in resume)
(Very good in at basics is important)

Interview 3:

1) What is the objective of the particular project, what technology was used in the project?
2) Which algorithm was used for the project? Why did you use it?
3) (food and price prediction) How does a client benefit from this project?
4) In machine learning, what is your take on a particular algorithm?
5) How do you finalize algorithms for particular projects?
6) Out of all machine learning, which algorithm can you explain?
7) Let's say we have categorical data, in that case, what will be the approach to build the decision tree? Why do you choose this approach? (like ID3)
8) What is the formulae of entropy?
9) What is your understanding of entropy? (in layman way)
10) Definition of Information gain.
11) Suppose if we have to create a tree, we need to define parent nodes for each and every label. So based on information gain how to decide which one is the node to the particular label?
12) What is MSE? How does it happen in regression?
13) Will there be the effect of imbalanced data set on decision trees? (in classification problem) why?
(answer: no, because of its hierarchy model)
14) Outliers? What did you do in the box plot?
15) What type of distribution data? How should you remove the outliers?

16) Will neural networks have the effect of outliers?
    (answer: wherever the gradient descent comes into the picture. There will be the effect of outliers)
17) Compare Resnet 50 with VGG.
18) VGG-16 and Resnet difference in research papers
    (answer: If we increase the number of layers, our error is not going to decrease. It increases after certain layers. They are looking for a better feature extractor. So they proposed Resnet and they also found that even if we change kernels sizes it's not going to help. They have introduced techniques like LRN, batch norm, and some other)
19) How object detection is different from classification?
20) What are Daily tasks in general?
21) Scenario-based Question
    Suppose you want to predict the growth of rice in India, based on some different parameters. How do you predict?
    (answer:
        1) Clustering for finding common geographic locations
        2) There will be some kind of outliers in clustering, we need to find some way
        3) We need to create that many numbers of regression models [number of regression models = number of clusters]
22) How do you update yourself?

Interview 4:

1) What are the projects worked on?
2) What is the model you have used for a particular project?
3) What is the object detection model you used?
4) (Object detection) Compared to Yolo V3. V4 how the Yolo act is better?
5) How did you start with ML?
6) If you have given data with some irrelevant points, You need to find outliers using a statistical approach and how you deal with them. What is the method/approach you are going to use? (answer: IQR, box plot, Z-score, skewness test)
7) How will you treat the outliers without removing data? (Answer: Normalization)
8) Explain the meaning of Normalization?
9) What is the difference between normalization and standard deviation?
10) In machine learning, In linear regression, we used to consider RMS error. Why sometimes RMS is the worst in terms of calculations for regression?
    (answer: In some case, loke data is varied like some data is in between 0-5 and other is 0-10000, In these cases, RMS is not the same that might be the reason)
11) What to do in such cases? (answer: we can do feature scaling)
12) Talk about the Error/ loss function in logistic regression.
13) What is your favorite algorithm?
14) Let's say I'm a kid, You have to explain logistic regression and linear regression. How does the prediction happen? Teach me. (Tell me the process happening inside them) (I'm not in a situation of understanding math)
15) Even if we can use logistic regression in multiclass classification, why it is not that good as it is used for binary classification?
16) In logistics have you heard of one vs all?
17) How did you do the train, test split in the projects with respect to time series?
18) Why do we take the time sequence when it comes to the time series data set?
    (answer: In such cases, we will miss the direct effect/indirect effect of previous data on current data, Usually in time series problems data should be considered in sequential order)
19) How do you decide the number of clusters in K means clustering?
20) Are there any algorithms that won't get impacted by an imbalanced data set?
    (answer: Random Forest due to Bootstrapping)
21) Scenario-based question

Let's say you are a placement cell coordinator, Every company visiting your college has some set of different requirements. As a coordinator, you have thousands of resumes and you have seen the specific skills required to the company and allow/map the person to that particular company. Relevant to the company requirements. Your job is to build a system where you have to build a model that everyone will dump their resume. Based on machine learning or statistical techniques. (Don't use NLP, Deep Learning)
(answer: Hamming distance approach, cosine similarity approach, clustering-based approach
Example approach:
We can take one sample resume with company requirements and we calculate the distance between the resumes, the JD's with minimum distance is good enough for the company)

22) You can change the categorical data into numerical data by using One-Hot Encoding, what if you have a category of pin code.
(answer: mean encoding approach, target mean encoding approach)
23) Can you explain one of the approaches?
24) Why do we need conversion?
25) How did you choose the number of layers and number of neurons in the architecture?
(answer: Keras tuner or trial and error method)

Interview 5:

1) Difference between sample and population data in Statistics?
2) Difference between the sample mean and population mean?
3) What is the importance of standard normal distribution?
4) Let's say standard deviation is one, what are we talking about(population or sample data)?
(answer: sample data)
5) I have collected two samples and I'm going to do a survey to know the people below the poverty line and give some benefits. How to statistically classify a true sample or a false sample?
(answer: Z-statistics) (approach if you can)
6) Difference between PDF and CDF, describe their use cases?
7) By using CDF, can we calculate percentile?
8) What are the favorite machine learning algorithms?
9) Difference between bagging and boosting?
10) How do you decide the degree of the polynomial in polynomial regression without overfitting?
11) Difference between overfitting and underfitting?
12) How is the bias and variance in overfitting?
13) Can you implement linear regression other than using a sklearn library? How can you calculate the performance of the model?
14) What is the difference between R squared and adjusted R square?
15) Can R squared be negative?
16) Have you heard of lasso regression?
17) Why do we need regularisation?
18) If there are outliers in your regression problem, what will you do?
19) Tell me about SVM.
20) What are the limitations of SVM?
21) If it's non-linear, is it ok to use SVM?
(answer: In this case, SVM takes high computation cost, there are some methods like linear kernel, polynomial kernel etc.)
22) Difference between soft margin and hard margin?
23) What are all the boosting algorithms you heard?
24) Difference between gradient boosting and XG boosting?
25) How does XGboost handle outliers?
26) What are the clusters you know?
27) In which case you shouldn't use KNN?
28) Try to explain KNN.
29) KNN is a Top-down approach, what will you say yes or no?

30) K-mean is a top-down or bottom approach?
31) Can you please talk about Euclidean distance, hamming distance, manhattan distances?
32) I'll give categorical data, you cannot use any other distance except euclidean distance. What will you do? (Usually, for categorical data we use hamming distance)
    (answer: Turning categorical data into vectors with the help of One-Hot encoding)
33) Suppose you are in New York, If you want to calculate distance between two blocks, what will you use Euclidean distance or Manhattan distance?
    (answer: Manhattan)
34) Have you heard of DBscan?
35) I have random/Zig Zag data, I have to use a clustering algorithm, Can we use K-means clustering?
36) Name activation functions, loss functions, optimizers that you know.
37) What is the reason behind using Adam optimizer? advantages.
38) Name CNN algorithm?
39) Why to go with RNN rather than CNN in NLP?
40) How LSTM works?
41) Do you have any idea of Encoder-Decoder?

Interview 6:

1) Tell me about yourself
2) What is data leakage?
3) Suppose you have to do standard Scaling, how do you proceed with that?
4) What happens if we apply fit_transform?
5) What is your understanding of API?
6) What are all the frameworks you will use to build an API?
7) Difference between get request and post request?
8) Develop a method and try to call that method through an API.(15 minutes)
9) Difference between standardization and normalization?
10) Minmax scalar range, and what can you call it?
11) Tell me something about the Central limit theorem.
12) To apply the central limit theorem, what is the sample you will take from the population?
13) Which is your favorite algorithm in ML?
14) How is DBSCAN different from K-mean clustering?
15) Is there any method to automate the process to automate and know the number of clusters?
    (Answer: elbow method, melocator in scikit learn)
16) Explain Navy Bias theorem and classifier.

Interview 7:

E&Y interview Process(Data Scientist):

Round 1:

1) Use case-solving(Ex: Image Classification) (15-20 minutes)

2) Technical interview(45 minutes)
    a) CNN's, RNN's
    b) Some machine learning algorithms
    c) Performance metrics
3) The complex problem in Deep Learning(Basically using transfer learning)(15-20 minutes)

Round 2: (15-20 minutes)

4) Managerial round
5) Technical manager round
   (Some problems, Use cases are given)

Round 3:

6) Director round
   (brief about all the things we worked on, the domain we worked on, and the challenges we faced. Overprint of the projects we have done.)

(Note: Interview is not that difficult, a normal person who knows machine learning and deep learning can do it. Need GOOD Communication skills with respect to any projects. Be confident (revise the things you have done))

(Better to go with referral)