Machine learning answer sheet

01) A

02) B

03) B

04) C

05) A

06) A

07) D

08) C

09) A

10) A

11) B

12) A & D

13) Regularization is a technique used to replace the error by fitting the function appropriate on the given training set and avoid overfitinfg

14) There are three main regularization technique 1) Ridge regression, 2) Lasso (L1 Norm), 3) DROPPOUT

15) Linear regression most often uses mean square error (MSE) to calculate the error of the models. MSE is calculated by measuring the distance of the observed y- values from the predicted y-values at each values of x squaring each of this distance calculating the mean of each of the squared distance.

Python answer sheet

1) C
2) B
3) C
4) A
5) D
6) C
7) A
8) C
9) A
10) D

STATISTICS ANSWER SHEET

1) A
2) A
3) B
4) D
5) D
6) A
7) B
8) A
9) C
10) A normal distribution is a type of continuous probability distribution in which most data point cluster toward the middle of the range while the rest taper off symmetrically towards either extreme. The middle of the range is also known as the mean of the distribution.
11) One way to impute missing values in a time series data is to fill them with either the last to fill them with either the last or the next observed values pandas have fillna() function which has method parameter where we can choose “ffill” to fill with the next observed values or “bfill” to fill with previously observed values.
12) A/B testing also called split testing or bucket testing compares the performance of two version of content to sec which one appeals more to visitors/ viewers it test a control (A) version against a variant (B) version to measure which one is most successful based on your key metrics.
13) Mean imputation is typically consider terrible practice since it ignore feature correlation. Consider the following scenario we have a table with age fitness score & an eight year old has a missing fitness score.

14) Linear regression analysis is used to predict the value of a variable based on the values of another variable, the variable year want to predict is called the dependent variable. The variable year are using to predict the over variable values is called the independent variable.

15) The two major areas of statistics are known as descriptive statistics, which described the properties of sample & population data inferential statistics, which uses those properties to best hypothesis & draw conclusions.