Answers to PFA file 2,3

21)

**d) Both a and b**

a)$\beta_0$ ,$\beta_1$,...,$\beta_r$ are the regression coefficients:

This statement is true. In linear regression, $\beta_0$ is the intercept, and $\beta_1$,...,$\beta_r$ are the coefficients corresponding to the predictors $x_1$,..., $x_r$, respectively.

b) Linear regression is about determining the best-predicted weights by using the method of ordinary least squares (OLS):

This statement is true. Linear regression aims to find the best-fitting linear relationship between the dependent variable $y$ and the independent variables $x_1$,..., $x_r$ by minimizing the sum of the squared differences between observed and predicted values using OLS.

c) $E$ is the random interval:

This statement is unclear and doesn't relate to the standard terminology of linear regression. $E$ denotes the error term or residual in regression models, representing the difference between observed and predicted values. It is not referred to as a "random interval."

d) Both a and b:

Given that statements a) and b) are true based on standard linear regression concepts

---

22)

**d) The value $R^2$ = 1, which corresponds to SSR = 0**

In linear regression:

$R^2$ (R-squared) is a statistical measure that represents the proportion of the variance for the dependent variable that's explained by independent variables in a regression model.

SSR stands for the Sum of Squared Residuals, which measures the total sum of the squared differences between the observed values and the values predicted by the model.

A perfect fit in linear regression would mean that the model explains all the variance in the dependent variable, implying that the Sum of Squared Residuals (SSR) is 0 because there's no discrepancy between observed and predicted values.

23)

**b) $\beta_0$**

In simple linear regression, the point where the estimated regression line crosses the $y$-axis (i.e., the intercept) is represented by the coefficient $\beta_0$.

---

24)

**d) The top-left plot**

---

25)

**d) d, b, e, a, c**

---

**26)**

The optional parameters for Linear Regression in scikit-learn are:

**b) fit_intercept**

- Determines whether to calculate the intercept (constant term) for the model.
- Set to True by default, meaning an intercept is included.
- Set to False if your data is already centered or you specifically don't want an intercept.

**c) normalize**

- Specifies whether to normalize the input features (X) before fitting the model.
- Set to False by default.
- Normalizing can sometimes improve model performance, especially when features have different scales.

**d) copy_X**

- Controls whether to create a copy of the input features (X) before fitting the model.
- Set to True by default to prevent modification of the original data.
- Set to False if you want to avoid the overhead of copying and are sure you won't modify X.

**e) n_jobs**

- Specifies the number of processor cores to use for parallel computation.
- Set to 1 by default, meaning a single core is used.
- Set to -1 to use all available cores for potentially faster fitting, especially for large datasets.

27)
**c) Polynomial regression**

Polynomial regression is an extension of multiple linear regression but allows for modelling relationships between variables as polynomial equations. This involves transforming the original features x into polynomial features, which include not only the original features but also their powers.

---

28)
**C) You need more detailed results.**
This is true. Statsmodels often provide more detailed statistical results, including p-values, confidence intervals, and other statistical metrics that provide deeper insights into the data and model.

---

29)
**b) Numpy**

---

30)
**b) Seaborn**

---

41)
**d) Collinearity**

---

42)
**b) Random Forest**
It is an ensemble learning method that combines multiple decision trees in order to improve the predictive performance and control over-fitting. It uses bagging to create an ensemble of decision trees where each tree is trained on a random subset of the data.

43)

**c) Decision Trees are prone to overfit: This is correct.**
One of the significant disadvantages of decision trees is that they can easily overfit the training data, especially when they are deep or complex.

---

44)

**c) Training data**
Machine learning algorithms learn from the training data to build a model that can make predictions or decisions without being explicitly programmed to do so.

---

45)

**c) Anomaly detection**

---

46)

**a) Support Vector**

---

47)

**d) Both a and b**
a) Statistical learning theory- It focuses on understanding the principles of learning from a statistical perspective, including concepts like bias-variance tradeoff, regularization, and empirical risk minimization.

b) Computational learning theory- It is concerned with the efficiency and computational aspects of learning algorithms, such as time complexity, sample complexity, and the feasibility of learning certain functions.

---

48)
**c) Both a and b**

---

49)
**c) 3**
In radial basis function (RBF) neural networks, there are typically three types of layers:
**Input Layer:** This layer receives the input data.
**Hidden Layer:** This layer contains the radial basis functions that transform the input into a higher-dimensional space.
**Output Layer:** This layer produces the network's output based on the transformations performed by the hidden layer.

---

50)
**d) KMeans**

---