**Linear Regression**

Linear regression falls under supervised machine learning and is a model where the algorithm is to find the linear relationship between a dependent and independent variable (Analytics Vidhya, 2021). In other words, the most fitting linear line that exists between a dependent and independent variable. It is most useful in situations where there are no discrete set of possible values rendering an unlimited number of possibilities. An example is to try to determine the expected log in time for new users of an app. And to do so collect data of one’s log in time on the app for a week, and write a browser plugin which will use the data to predict the log in time for new users of the app.

**A decision** **tree**

Decision Trees come under supervised machine learning, indicating a stipulated input and its corresponding output in the training data. A decision tree is drawn upside down, which means its roots are at the top. Its algorithm is made up of decision nodes and leaves, where the leaves are the final outcomes which depict the leaf at the end of a branch which no longer splits. The decision nodes indicate where the data is split indicating where a tree splits into branches. Decision trees are best used where data needs to be classified (Gupta, 2017). For example, where one needs to determine whether a person is fit or unfit based on one’s age, eating habits and physical activity, as shown below:

Is a person fit?

Age >= 40

yes? No?

Eats a lot of junk food? Exercises daily

yes? No? yes? No?

UNFIT FIT FIT UNFIT  **Ref**: Xoriant (2022)

**K-means**

K-means falls under clustering which is an unsupervised machine learning algorithm. It uses inferences from datasets that have no labels or known outcomes. It groups similar data points to determine patterns. To do this, K-mean figures out a fixed number (k) of clusters in a dataset (Education Ecosystem, 2018). A cluster being a group of data points put together due to similarities. The fixed number (k) is the number of centroids needed in a dataset, with a centroid being the center of the cluster, that is averaging of the data to find the centroid (Education Ecosystem, 2018). K-means can be used to cluster and predict different subgroups such as finding various clustered segmentations like market segmentation (Education Ecosystem, 2018).

**Naive Bayes**

Naive Bayes is a supervised machine learning model mostly used for large volumes of data. It works via a fast and less complicated classification algorithm (Dwivedi, 2020). It is modelled after the Bayes theorem which is based on conditional probability where it is ascertained that something will happen, given an occurrence. Similarly, Naive Bayes uses an algorithm that works as a classifier using probabilities to segregate different objects based on particular features a variable contains (Dwivedi, 2020). Naive Bayes classifiers uselabelled data and are therefore pre-categorized for ease of classification. A classic example of Naive Bayes is email Spam filtration.

**REFERENCES**

Analytics Vidhya (2021) *All you need to know about your first Machine Learning model – Linear Regression.* [Online] Available at <https://www.analyticsvidhya.com/blog/2021/05/all-you-need-to-know-about-your-first-machine-learning-model-linear-regression/#:~:text=In%20the%20most%20simple%20words,the%20dependent%20and%20independent%20variable>. [Accessed on the 18th of Sept. 2022].

Dwivedi, R. (2020). *What Is Naive Bayes Algorithm In Machine Learning?* [Online] Available at <https://www.analyticssteps.com/blogs/what-naive-bayes-algorithm-machine-learning> [Accessed on the 19th of Sept. 2022].

Education Ecosystem (2018). *Understanding K-means Clustering in Machine Learning.* [Online] Available at <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1> [Accessed on the 18th of Sept. 2022].

Gupta, P. (2017) *Decision Trees in Machine Learning*. [Online] Available at <https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052> [Accessed on the 19th of Sept. 2022].

Xoriant (2022) *Decision Trees for Classification: A Machine Learning Algorithm.* [Online] Available at <https://www.xoriant.com/blog/decision-trees-for-classification-a-machine-learning-algorithm#:~:text=Introduction%20Decision%20Trees%20are%20a,namely%20decision%20nodes%20and%20leaves>. [Accessed on the 19th of Sept. 2022].