**Literature Review**

2.1 Machine Learning:

2.1.1 What is ML

Machine Learning is a type of Artificial Intelligence (AI) that allows a computer to predict outcomes without being specifically programmed to do so. This type of AI is becoming more popular and widely used in fields such as sales forecasting [1].The main objective of Machine Learning is to study, engineer and improve mathematical models. These models can be trained more than once, given related-context data to be able to predict future data and make necessary decisions [2].

If a machine is well trained, it will be able to learn very complex scenarios and predict the expected data with high precision. The choice of the correct predictors is very important when training a Machine Learning model [3]. Besides in the forecasting of data, due to being versatile, Machine Learning can also be used to identify objects, detect fraud, detect malware and filter spam. Machine Learning is made up of different algorithms, each having their own ways of learning the data. Some commonly used algorithms are Linear Regression, Decision Trees, Random Forest, KNN and K-means. Nowadays, Machine Learning is very important to be able to predict sales in businesses, keep up with customer demands, view changing customer trends, improving network security and support development of products amongst more.

2.1.2 Difference between supervised and unsupervised

In Machine Learning, different techniques are used. These techniques include supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. Each of these techniques works by implementing different algorithms which process and learn data in different ways. The most used techniques are supervised and unsupervised learning. The main difference between these two techniques is that supervised training required the programmer to label data beforehand in order for the machine learning algorithm to predict an outcome (focuses more on classifying the data using labels), while unsupervised focuses more on clustering the data, reducing dimensions and identifying sequences by association [4].

Unsupervised learning is mostly used in scenarios of speech processing, object categorisation, audio classification, and automatic labelling. On the other hand, Supervised Learning is commonly used in sentiment analysis, predictive analysis based on regression or categorical classification, natural language processing, detecting email spam and image classification. Unsupervised learning includes algorithms such as K-Means, Hierarchical Clustering, Density-Based Clustering and A-priori. Supervised learning includes algorithm such as Decision Tree, Random Forest, Support Vector Machines and Linear Regression. Algorithms which can be implemented in both supervised and unsupervised learning are Naïve Bayes, Auto Regressive Integrated Moving Average (ARIMA) and Deep Learning. In this research, supervised learning will be used rather than unsupervised since the data needs to be classified and labelled based on different time-periods and it has multiple factors which are affecting it in different ways.

2.2 Forecasting Sales or Demand:

2.2.1 Why forecasting is used

The majority of decisions made in our everyday life requires forecasting. These forecasts can be as simple as deciding what time one needs to wake up in the morning to be at work on time. A student needs to determine which course to study depending on the future job of their choice. Other complex forecasting includes deciding how to allocate money and budgeting depending on your wage and expenses.

Nowadays, forecasting using computerised methods has become a very popular process that is used in different fields to be able to make predictions of future data based on past data. In around 15 years, the field of forecasting has seen amazing growth in both theory and practice. Using today’s advanced technology, one can immediately get a detailed hour-by-hour weather forecast. In this forecast, the temperature, rain, weather conditions, UV index, wind and much more are predicted. Probability forecasts are used when uncertainty is irreducible, for example it can be used during elections to forecast which party will win the election. This does not necessarily mean that the party forecasted to win will surely win, but there is a great possibility that the forecast is correct. Forecasts can also be utilised in companies when they are selling a service or products to predict future sales based on the past sales made [5].

All these forecasts can be estimated based on using multiple past parameters called the predictors. This prediction is made available using the different algorithms found in Machine Learning to be able to not only forecast upcoming data, but also help in decision-making and analysing current status such as the company’s performance. The information based on the forecasting can help businesses to allocate resources, anticipate expenses and plan their budgets wisely. When it comes to predicting sales, one can also have an idea of how the production schedules need to be set [6]. When performing the forecasting, one has to pay attention to the data being used and how it is used in the forecasting, as if the data is not used correctly, the forecast will not output the correct results either. If a company makes their decisions based on the incorrect predictions, the company may suffer many losses in both target audiences and also the sales of products or services.

2.2.2 Time Series Modelling

At present, the most commonly used forecasting method of sales forecasting is time series modelling. Sales predictions need to be analysed based on three different components which are: seasonal trend, irregular, cyclical. …

2.3 Importance of sales forecasting:

2.3.1 Reaching customer demand

In businesses, past sales information can be used to forecast the future sales for the upcoming months. Forecasting in companies is used, so that the business can have a general grasp of the market demand and to help formulate a more suitable marketing strategy. Due to the increase in the use of AI technology, competition in the market is growing at a more rapid pace [100]. The company can also have an indication of what type of services or products a customer is searching for and in what ways are the trends and patterns are changing and evolving. If the company does not satisfy the market demand for a product, customers will have to opt to purchase from competitors, which will then result in loss of sales and possibly even losing the customers’ loyalty. A business can get all the demand information by applying their sales data to any of the desired forecasting models and increasing the chances of improving the satisfaction levels of its customers.

2.3.2 Seasonality

Seasonality stands for regular patterns that are determined by different seasons over a given number of periods. Seasonality in terms of sales refers to the oscillations in total amount of sales that occur throughout one year and then repeats in the following years. Seasonality is not determined by the volume of sales of the whole year but in volumes during specific periods. This periodic seasonality can be based on short volumes such as weeks or months [101]. In an industry, seasonality in sales is highly influenced by the different seasons of the year, holiday periods such as summer breaks or Christmas holidays, and other notable dates such as Mother’s Day, Father’s Day or Valentine’s Day. When a forecast is predicted based on seasonality, the accuracy must be very high because if the forecast is inaccurate, this may cause major issues in marketing, production, investment and expenses. If the marketing and advertisement is not scheduled in the correct seasonality, the targets will not be achieved and the money for campaigns would be spent for nothing. Over production of products cause by incorrect seasonality can also be an issue as certain products have to be sold in a specific period of time. If these products are not sold in this period of time, they may have to be either thrown away or sold at a very cheap price that does not cover expenses.

2.3.3 Shelf life

Once an aesthetic product is produced, the average shelf life of an unopened product is of around 2 years. This varies depending on the chemicals which are added to products. The more chemicals used, the longer the shelf life will be. Products such as dermal fillers cannot contain certain chemicals, as they need to be injected directly through the skin. For the aesthetic industries, shelf life is a very important feature to guarantee the quality and safety of the product. When a customer purchases an aesthetic product, they would not expect the product to expire in a couple of days or months.

2.3.4 Aesthetic Documentation

Cosmeceuticals are the products which fit the niche between drugs and cosmetics. This term is used in the professional skin care industry to describe a product that has measurable biological effect in the skin, like a drug, but is regulated as a cosmetic since it claims to affect appearance. These professional skin care products come in the form of sunscreen, antiaging creams, foundation, facemasks, derma fillers and more. Cosmeceuticals are the fastest growing segment of the personal care industry and their formulations have expanded from skin to body to hair and a number of tropical cosmeceutical treatments for conditions. Studies focusing on the cosmeceutical products highlight that there will be strong growth perspective for this industry in the coming years [102].

2.4 Extraordinary events affecting sales:

2.4.1 Covid

2.4.2 Brexit

2.4.3 Change in Government

# Bibliography

|  |  |
| --- | --- |
| [1] | J. Li, "A Feature Engineering Approach for Tree-based Machine Learning Sales Forecast, Optimized by a Genetic Algorithm Based Sales Feature Framework," IEEE Xplore, 1 May 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9820532. [Accessed 10 January 2023]. |
| [2] | G. Bonaccorso, Machine Learning Algorithms, Birmingham: Packt Publishing Ltd, 2017. |
| [3] | P. Popovski, G. Veljanovski, M. Kostov and M. Atanasovski, "Optimizing Short Term Load Forecast: A study on Machine Learning Model Accuracy and Predictor Selection," IEEE Xplore, 1 June 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9828783. [Accessed 10 January 2023]. |
| [4] | K. R. Dalal, "Analysing the Role of Supervised and Unsupervised Machine Learning in IoT," *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC),* vol. 4, no. 4, 2020. |
| [5] | F. Petropoulos, D. Apiletti, V. Assimakopoulos and M. Z. Babai, "Forecasting: theory and practice," *International Journal of Forecasting,* vol. 38, no. 3, 2022. |
| [6] | D. J. Dalrymple, "Sales forecasting methods and accuracy," *Business Horizons,* vol. 18, no. 6, pp. 69-73, 1975. |

[100] H. Wei and Q. Zeng, “Research on sales Forecast based on XGBoost-LSTM algorithm Model,” Journal of Physics: Conference Series, vol. 1754, no. 012191, 2021, doi: 10.1088/1742-6596/1754/1/012191.

[101] J. de Mesquita and H. Martins, “Retail industry: seasonality in sales, and financial results,” Brazilian Business Review, vol. 8, no. 3, pp. 64–82, Jun. 2011, doi: 10.15728/bbr.2011.8.3.4.

[102 ]A. Lohani, A. Verma, H. Joshi, N. Yadav, and N. Karki, “Nanotechnology-Based Cosmeceuticals,” ISRN Dermatology, vol. 2014, no. 843687, pp. 1–14, May 2014, doi: 10.1155/2014/843687.