**Introduction:**  
Sales prediction is essential for growing a business. Business intelligence is mostly derived from an accurate prediction which is very important in their progress. It helps companies to understand how to manage their workforce, cash flow, and resources effectively. Businesses analyze past and current sales data to estimate future sales and surely is crucial for financial planning. Accurate sales predictions ensure companies to make informed decisions, leading to better supply chain management, increased profits, and improved customer experiences.

For new businesses, sales prediction is vital for managing resources efficiently. Understanding sales trends and consumer behavior through data insights also helps in creating effective marketing strategies and targeting the right audience. Additionally, sales forecasts serve as benchmarks for business performance and assist in planning demand and supply actions. Overall, sales prediction provides a solid foundation for making strategic business decisions and ensuring long-term success.

**Exploratory Data Analysis (EDA)**

**Performance Metrics:**

AI-driven sales forecasting models are evaluated for accuracy using a range of statistical metrics. These measurements provide in-depth, quantitative understanding of how well a model forecasts future sales:

1. Mean Absolute Error (MAE):

The average of the absolute differences between the expected and actual sales values is determined by the Mean Absolute Error (MAE) metric. It indicates the average difference between the model's forecasts and the real sales statistics. A lower MAE suggests that overall, the model is producing more accurate predictions.

1. Mean Squared Error (MSE):

MSE computes the average by squaring the differences between the actual and expected sales figures. Models with occasional but considerable errors can be identified with the help of MSE, a useful metric that emphasizes larger mistakes by squaring the errors. A lower MSE indicates that while the model is cautious in avoiding significant errors, it is generally correct.

1. Root Mean Squared Error (RMSE):

The error is converted back to the original scale of the data using the Root Mean Squared Error (RMSE), which is the square root of the MSE. A reduced root mean square error (RMSE) indicates that the model's predictions are reasonably close to the actual sales data. This statistic gives an idea of how dispersed the errors are. When we want an error statistic that is comparable to the data and easy to read, RMSE is frequently chosen.

1. R-Squared (R2):

This statistic assesses the extent to which the model can account for the variance in the real sales data. To put it another way, it indicates how well the model represents the broad patterns found in the data. A model is good at explaining why sales figures fluctuate if it fits the data well, as shown by an R-squared value closer to 1.

Even though these measures are critical for assessing model performance, the business context must be taken into account when determining which indicator should take precedence. For instance, measures like MAE or RMSE can be more appropriate if it's critical to forecast the precise amount of sales. R-squared, however, might be more helpful if the objective is to comprehend more general sales trends.

**Model Validation and Performance Monitoring**

After a model is trained and tested, the evaluation of AI-driven sales forecasting models continues. Maintaining the model's accuracy in the face of changing business conditions requires continuous assessments. The following techniques are applied to this ongoing evaluation:

K-Fold Cross-Validation: This technique divides the data into k distinct pieces, or "folds". After training on k-1 sections, the model is tested on the remaining portion. To get a better indication of how well the model will perform with additional, unseen data, this process is repeated k times.

Hold-Out Set Validation: Using this technique, part of the data is reserved and not utilized during training. This hold-out set is used to test the model after it has been trained. Anyone can easily predict how well the model will function in actual use.

Monitoring Forecast Errors: Organizations can determine whether the model is becoming less accurate by routinely examining the variations between the model's projections and actual sales data. It could be necessary to retrain the model with fresh data or modify its settings if the predictions start to become less accurate.

AI-driven sales forecasting models are kept accurate and dependable over time by using these validation and monitoring techniques. To enhance sales effectiveness and make data-driven decisions business owners should have this level of accuracy.

Literature Review:

1. Walmart Sales Forecasting Using Different Models (Chenghao Yu) 2024

This study compares three methods Random Forest, Linear Regression, and Lasso Regression—to see which one best predicts Walmart sales. It uses past sales data and factors like time, unemployment rate, CPI, and temperature. After testing, the study found that Random Forest performed better than the other two methods based on MSE and R squared scores. This suggests Random Forest could be a useful tool for predicting sales, managing inventory, and planning at Walmart and other stores.

1. Sales Prediction Using Linear and KNN Regression (Shreya Kohli) 2021

This paper uses two types of regression models: linear regression and KNN regression. The dataset used in this paper comes from the Rossmann Store Sales on Kaggle, which includes data from 3,466 stores of the German drugstore chain Rossmann. Most products were sold on Mondays and Fridays, and the number of customers peaked at 256 million in 2013 but dropped to 108–148 million by 2015. Sales and customer visits were higher during store promotions, with more buyers in the fourth quarter.

The study began by collecting and cleaning the data, handling missing values, and selecting important features. Then, predictive analysis was done using classifiers, with k-fold cross-validation (k=10) to improve accuracy. The model was evaluated using statistical methods like RMSE and MAPE. The results showed that linear regression slightly outperformed KNN regression.

1. Time Series Forecasting Model for Supermarket Sales using FB-Prophet (Bineet Kumar Jha) 2021

This research used the FB Prophet tool to predict supermarket sales. The dataset, taken from Kaggle, includes 9,994 records of sales data related to products with 20 attributes (mainly furniture) sold in a supermarket between 2014 and 2017. The study compared different forecasting models, including the additive model, ARIMA model, and FB Prophet. The results showed that FB Prophet was the best model for accurate predictions with low error. However, the study suggests that performance could be improved by combining FB Prophet with other techniques. The research also noted that scalability could be a challenge when working with large datasets, and using a transfer learning approach could help handle this issue. Accurate real-time predictions depend on the model used for training and validation.

1. Leveraging Artificial Intelligence for Enhanced Sales Forecasting Accuracy: A Review of AI-Driven Techniques and Practical Applications in Customer Relationship Management Systems (Srinivasan Venkataramanan) 2024

Experienced software engineers explored the growing use of AI in sales forecasting within CRM platforms. AI techniques, including machine learning, natural language processing (NLP), and deep learning models were reviewed to assess their effectiveness in predicting sales. The study used data from sources like historical sales, customer interactions, social media sentiment, market trends, and economic indicators. They examined the quality of the data, how well the models could be understood, and their ability to find hidden patterns and non-linear relationships. The research focused on practical uses of AI in sales management, such as improving pipeline optimization by predicting the chances of closing deals and helping sales teams prioritize high-value prospects. AI-driven forecasting also aids in resource allocation, ensuring sales teams have what they need to target the best opportunities, leading to better sales efforts and higher returns on investment.

1. Time-series forecasting of seasonal items sales using machine learning – A comparative analysis (Yasaman Ensaf) 2022

The author used the same dataset as in a previous study [3], which includes real sales data for furniture in a retail store. The authors analyzed the furniture sales data and performed a unique exploratory data analysis to uncover more insights about the dataset. Unlike previous studies that only used a few forecasting methods, they applied several traditional time-series forecasting techniques like SARIMA to predict furniture sales. They also used advanced forecasting methods based on Artificial Neural Networks, including Prophet, LSTM, and CNN. The accuracy of the results was compared using metrics like RMSE and MAPE. The findings showed that the Stacked LSTM method performed the best among all methods. The Prophet and CNN models also performed well, and including a holiday factor improved the performance of the Prophet model.

1. Machine learning model for sales forecasting by using XGBoost (Zhang Shilong) 2021

The author created a sales forecasting model using machine learning focusing on both efficiency and accuracy. They started by extracting important features from historical sales data and used eXtreme Gradient Boosting (XGBoost) to predict future sales. XGBoost was chosen for its efficiency and scalability than other machine learning methods. They tested their model on Walmart sales data from the past 1913 days across three states, aiming to predict sales for the next 28 days. Since XGBoost is sensitive to outliers, they cleaned the data first and optimized it to save memory. The model's performance was measured using the RMSSE score, and it was found to be 16.3% better than Linear Regression and 15.4% better than Ridge Regression.

1. Retail sales forecasting with meta-learning (Shaohui Ma) 2020

The authors developed a meta-learning framework using deep convolutional neural networks that can automatically learn features from raw sales data. These features are then used to combine various forecasting methods into one model. Their experiments with weekly sales data showed that this approach outperforms many advanced forecasting methods, although the accuracy improvement over some other sophisticated methods is modest, and the learned features are hard to interpret. They recommend using a combination of different forecasting methods rather than relying on just one when forecasting retail sales.

1. Export sales forecasting using artificial intelligence (Vahid Sohrabpour) 2020

Traditional time series forecasting methods are commonly used but have some limitations. Unlike these, causal forecasting methods predict future sales by looking at the relationships between different factors, not just past trends. This research introduces a new approach for forecasting export sales using Genetic Programming (GP), an AI technique inspired by biological evolution. They applied this method to a Middle Eastern export company facing fluctuating sales. The researchers created a model to forecast sales for six weeks and compared the predictions with actual sales data. They also conducted a sensitivity analysis to see how different factors affect the predictions. The study suggests that comparing traditional forecasting methods with AI-based ones can reveal the strengths and weaknesses of each approach.

1. Application of facebook's prophet algorithm for successful sales forecasting based on real-world data (Emir Žunić) 2020

The proposed framework, useful for any retail company, is based on Facebook's Prophet algorithm and a backtesting strategy. The framework aims to generate monthly and quarterly sales forecasts for products without using extra data. It was tested using real sales data from one of the largest retail companies in Bosnia and Herzegovina.

To assess how accurate the forecasts are, the researchers used an expanding window backtesting strategy, repeating the process 12 times with a one-month interval. Each time, they fitted the Prophet model to historical sales data and compared the forecasts for the next three months with the actual sales. They calculated the percentage error (PE) for each forecast and used the mean absolute percentage error (MAPE) to measure overall forecasting reliability.

While the framework shows promise, there are some limitations. These could be addressed by adding automated tuning of model settings, considering the impact of price changes and promotions, and integrating other forecasting tools alongside Prophet.

| No. | Title | Year | Literature Review |
| --- | --- | --- | --- |
| 1. | Development of a Time Series E-Commerce Sales Prediction Method for Short-Shelf-Life Products Using GRU-LightGBM | 2024 | * A stacking method for prediction is developed based on the **integration of GRU** and **LightGBM**. * Inherits the ability of the GRU model to **capture timing features** accurately. * Acquires the ability of LightGBM to solve **multivariable problems**. |
| 2. | Sales Prediction Based on ARIMA Time Series and Multifactorial Linear Model | 2023 | * Multiple factorial **linear regression** models and the **ARIMA** model in time series analysis will be adopted based on **R** to build prediction models to forecast sales volume, respectively. |
| 3. | Modern Centaurs: How Humans And AI Systems Interact In Sales Forecasting | 2023 | * Used **LightGBM** to build a competitive forecasting system. |
| 4. | A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data-Driven Networks | 2023 | * Constitutes an extensive review of the published scientific literature regarding the comparison of **ARIMA** and **machine learning algorithms** applied to time series forecasting problems, as well as the combination of these two approaches in hybrid statistical-AI models in a wide variety of data applications (finance, health, weather, utilities, and network traffic prediction). |
| 5. | A machine learning-based framework for forecasting sales of new products with short life cycles using deep neural networks | 2023 | * Apply one traditional statistical (**ARIMAX**) and three machine learning methods based on **deep neural networks** (DNNs) – **long short-term memory**, **gated recurrent units**, and **convolutional neural networks**. |
| 6. | A hybrid attention and time series network for enterprise sales forecasting under digital management and edge computing | 2023 | * Proposes a combination of enterprise sales forecasting from the perspective of digital management and neural networks, and proposes a network HATT-CNN-BiLSTM model for enterprise sales forecasting. * This work combines **multi-scale CNN** (MSCNN) with an **improved BiLSTM** (IBiLSTM) model. * The MSCNN is used to extract spatial features with different scales. * The processing of time series data is the strength of the LSTM network. |
|  |  |  |  |
| 7. | BMSP-ML: big mart sales prediction using different machine learning techniques | 2022 | * The **random forest predictor** has outperformed **ridge regression**, **linear regression**, and **decision tree models** among the four machine learning techniques implemented in this study. |
| 8. | Building a Lucy hybrid model for grocery sales forecasting based on time series | 2022 | * Build a new **hybrid** model called Lucy Hybrid. * Support the **trend** and **forecast function** for time series data. * Linear Regression, Elastic Net, Lasso, Ridge and Extra Trees Regressor, Random Forest Regressor, K-Neighbors Regressor, MLP Regressor, and XGB Regressor to experiment and create 20 Lucy hybrid sample models. |
| 9. |  |  |  |
|  | Time-series forecasting of seasonal items sales using machine learning – A  comparative analysis | 2021 | * Classical time-series forecasting techniques such as **Seasonal Autoregressive Integrated Moving Average (SARIMA)** and **Triple Exponential Smoothing** are utilized. * More advanced methods such as **Prophet**, **Long Short-Term Memory (LSTM)**, and **Convolutional Neural Network (CNN)** are applied. * The results show the superiority of the **Stacked LSTM** method over the other methods. In addition, the results indicate the good performances of the **Prophet and CNN models**. |
|  |  |  |  |

| No. | Title | Year | Contribution |
| --- | --- | --- | --- |
| 1. | Machine-Learning Models for Sales Time  Series Forecasting | 2019 | The effect of **machine-learning generalization** and a **stacking** approach for building a **regression ensemble** of single models has been studied.  In the case study of machine-learning generalization, they used the following additional features:   1. Mean sales value for a specified time period of historical data: This feature involves calculating the average sales value over a particular time period (e.g., weekly, monthly) based on historical data. It provides insights into the typical sales performance of a store during that time period, which can be useful for forecasting future sales. 2. State and school holiday flags: These binary flags indicate whether a particular day corresponds to a state holiday or a school holiday. Holidays can significantly impact consumer behavior and purchasing patterns, so including these flags as features allows the model to capture any seasonal or holiday-related trends in the data. 3. Distance from store to competitor’s store: This feature represents the geographical distance between the store under consideration and its nearest competitor's store. Understanding the proximity to competitors can help predict competitive pressures and their potential impact on sales. 4. Store assortment type: This categorical feature describes the type of assortment (e.g., basic, extended, extra) offered by the store. Different assortment types may cater to different customer segments and have varying sales patterns, so including this feature can capture the diversity in product offerings across stores.   In the stacking approach, the results of multiple model predictions (**ExtraTree**, **ARIMA**, **Random Forest**, **Lasso**, **Neural Network**) on the validation set are treated as input regressors for the next-level models. As the next level model, Lasso regression is used. Only **three** models from the first level (ExtraTree, Lasso, Neural Network) have non-zero coefficients for their results. |
| 2. | Food Sales Prediction “If we only knew what we know” | 2008 | The creation of an ensemble of predictive models using various learning algorithms and different time windows, and then combining them using dynamic integration.   * Ensemble of 24 models: A total of 24 individual models were created. These models were generated by combining different learning algorithms and time windows. * Learning algorithms: Eight different learning algorithms were used to create the models. These algorithms are:  1. 2 decision tree learners 2. 2 rule learners 3. 2 lazy learners 4. 1 support vector machine (SVM) 5. 1 logistic regression model  * Time windows: The data were divided into different time windows to capture temporal patterns. Three different time windows were used:  1. 13-time points 2. 26-time points 3. 52-time points  * Dynamic integration: After creating individual models using different algorithms and time windows, they were combined using dynamic integration. Dynamic integration is a technique for combining the predictions of multiple models in an adaptive manner. Instead of simply averaging the predictions or giving equal weight to each model, dynamic integration adjusts the contribution of each model based on its performance on the current data. This can lead to better predictive accuracy by giving more weight to models that are performing well on the specific task or dataset at hand.   Overall, this approach leverages the diversity of multiple learning algorithms and time windows to create a more robust ensemble model, and dynamic integration ensures that the ensemble adapts to the data and optimally combines the predictions of individual models. |
| 3. | Building a Lucy hybrid model for grocery sales forecasting based on time series | 2022 | Build a new **hybrid** model called Lucy Hybrid. Support the **trend** and **forecast function** for time series data. Experiment with a large dataset of more than 3,000,000 records from a large Ecuadorian-based grocery retailer.  Linear Regression, Elastic Net, Lasso, Ridge and Extra Trees Regressor, Random Forest Regressor, K-Neighbors Regressor, MLP Regressor, and XGB Regressor to experiment and create 20 Lucy hybrid sample models. |