Journal of Engineering Science and Technology Review 16 (5) (2023) 102 - 107

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Utilizing Ridge Techniques to Combat Food Waste

Hussam Mezher Merdas* **and Ayad Hameed Mousa**

Department of Computer Science, University of Kerbala, Karbala, Iraq.

Received 8 August 2023; Accepted 17 October 2023

Abstract

Daily, the world incurs significant losses due to food wastage. This study proposes the development of a predictive model for the actual food demand within companies or restaurants. The proposed model focuses on utilizing ridge techniques, specifically the three machine learning algorithms: Ridge Regressor, Kernel Ridge, and Bayesian Ridge. The selection of these three algorithms was deliberate, as they have not received as much attention compared to other machine learning and deep learning algorithms. This model utilized a proprietary dataset related to food product sales, comprising several widely recognized features commonly found in most companies and restaurants. This study employed various metrics such as (MSE, RMSE, MAE, and R2) to compare the prediction results obtained from these algorithms, and the results were as follows (Ridge Regressor: 82.87, Kernel Ridge: 70.30, and Bayesian Ridge: 82.98). This research sets a limit to the daily food wastage that incurs significant financial losses worldwide.

Keywords: food wastage; machine learning algorithms; Ridge Regressor; Kernel Ridge; Bayesian Ridge.

1. Introduction

Many food sales companies relied on very primitive methods in estimating the quantities that they would need to sell in the market. The managers of these companies employed many employees with specializations in mathematics and economics to increase the profits of their companies by reducing the losses resulting from damage to goods as a result of the expiration of their storage [1]. In the end, these employees and experts fail to give accurate accounts, which puts their companies in the dilemma of losing products. Add to that, their many numbers which lead cost the companies pay large salaries. In addition, many companies are competing with each other. Each of these companies wants to increase its profits more than its competitors [2]. On the other hand, they desire the least losses in terms of employee salaries and numbers, as well as in terms of reducing product losses as a result of their damage due to their remaining in the companies' warehouses without selling them because the availability of these goods is more than the demand for them [3].

From this standpoint, and to achieve the best profits, companies have developed research and scientific aspects to reduce the burden of product loss [4]. The researchers took from the past sciences as an important pillar and proceeded to use modern technologies to develop efficient forecasting systems to estimate the future need for food sales. This has been of great benefit to companies by storing only the quantities they will need without falling into the problem of stockpiling goods [5].

This study relies on supervised ML. Specifically, the proposed model is based on the regression technique. Where the AI algorithms used are employed on real and continuous data to give forecast results for food sales, then the accuracy obtained from these algorithms is evaluated by comparing the results obtained from them with the actual numbers in the datasets. By reviewing and analyzing the relevant studies, found that they differed in terms of choosing the data on which they depend, depending on the companies that provide this data. Food datasets often need a lot of pre-processing before they can be used for prediction. After the stage of preparing the datasets, the stage of training the algorithms begins and then testing them to see their accuracy in prediction to use them later, whether in future studies or for the benefit of the food-selling companies' sector. Several practical measures were used to measure the accuracy of the proposed model and the results obtained from it. This study takes into account the juice of previous studies by designing a model that uses important and famous AI algorithms. Therefore, this study can be considered a brief and useful summary that is presented to companies without the need to waste time knowing the appropriate food sales forecast models.

2. Related works

Many researchers have worked hard over the years to study and develop food sales forecasting models. Companies have become more flexible by providing the necessary datasets for studies to serve researchers and companies themselves. The most important related works can be summarized in the following paragraphs:

High promotions produce high sales that are difficult to deal with and manage, especially in how these sales are scheduled and their quantities. To mitigate such problems good sales forecasting models are used. On this basis, H. Dai, et al. suggested designing a model based on the random forest (RF) algorithm based on clustering to forecast sales on peak days. A dataset provided by a major grocery store in China was used. The algorithm gave 13.43% better prediction results than the non-clustered RF algorithm, and 22.85% better than

*E-mail address: hussam.m@s.uokerbala.edu.iq
ISSN: 1791-2377
doi:10.25103/jestr.165.13

the non-clustered autoregressive integrated moving average (ARIMA) model [6].

This study was proposed by C.-H. Wang and aims to (1) rely on economic indicators and their associated timelines as two basic elements for forecasting, (2) use both ML algorithms and deep learning algorithms to make the forecasting process, and (3) the dynamic interactions between competing companies from on the one hand, and the sales sector on the other. Preliminary results show that the sales revenues of the retail sectors are affected mainly by the retail employment census, real wages, as well as the consumer price index. As for seasonal factors, they affect hypermarkets only. The results showed that learning deep learning algorithms give the best results [7].

Given the transition to the era of big data, R. Odegua proposed in this study a prediction model using the data of a store called “Chukwudi Supermarkets”, with the application of three machine learning algorithms “K-Nearest Neighbor, Gradient Boosting, and Random Forest”. The obtained results showed that the Random Forest algorithm gives better accuracy than the gradient-boosted and K-Nearest Neighbor algorithms. To improve the accuracy of prediction, the model focused on three main variables: the type of store, the price of the product, and the year in which the store was opened [8].

In this study, A. P. Wellens, et al. proposed a model based on the decision tree. A dataset containing 4,523 products from a leading retail store in Belgium was used and some external events such as national holidays and promotions were taken into account. The results gave an improvement in prediction accuracy over commonly used statistical methods by 9.34% and up to 20.52%. The proposed model relied on external variables, which improved the results obtained [9].

There is no way to accurately predict sales during promotions. Because of this problem, J. Wolters and A. Huchzermeier proposed a model that deals with products during seasonal offers. The model is based on two stages. In the first stage, the seasonal sales cycle is predicted by using the harmonic regression model, and this is done by not using promotional sales data during the holiday period in the second stage, the results of the first stage are combined with a function that uses holidays and promotions data. Then the final model is formed using the ridge regression algorithm. The data provided by the grocery store is used for short periods as well as for long periods. This model gave good accuracy in forecasting sales [10].

To forecast customer requirements in “Supply Chain Management SCM”, N. Vairagade, et al. tried to use machine learning algorithms to make an appropriate prediction. A “representative set of ML-based prediction techniques” has been applied to the used dataset. Then “R^2, Mean Squared Error (MSE), and MAE” scores were used to evaluate the accuracy of the algorithms used. Based on the results, the Random Forest algorithm gives a better prediction result compared to the artificial neural network that was applied to the same data [11].

In this study, V. Adithya Ganesan, et al. propose a model that predicts food sales in advance for a retail store in India. The model is taught through the internet as well as features engineering is adopted, after that, a neural network algorithm is used to make the prediction process. The proposed model gives better accuracy than the results obtained from the traditional time series models, as well as it gives better accuracy than the corporate’s current model by 7.7%. According to researchers working on this model, it leads to saving 170 units of food per day [12].

In this study, L. Zhou, et al. gave an introduction to DL and how neural networks work, as well as how to train the model. In this study, a survey of data related to food sales was conducted, so the survey included calories in foodstuffs, types of foodstuffs, their sources of contamination, and other matters. The specific problems, the existing data sets, the applied neural networks, as well as the obtained accuracy, were studied with a comparison with other research. The result of the research indicates that deep learning is superior to traditional artificial intelligence algorithms [13].

3. Materials and Method

This study used a distinctive dataset with the features it contains, which can be generalized to most companies. Focus was also placed on three algorithms that are underserved in terms of research and study. It was noticed that there was a lack of focus on these important algorithms. The important steps in this study will be detailed later.

3.1 The selected dataset

The first dataset is sourced from “*publicly available Alibaba's Tianchi platform data*” and it consists of 1000 rows and 15 columns as shown in Table 1. This data set contains food sales of different categories on the Alibaba platform in several cities in the year 2019.

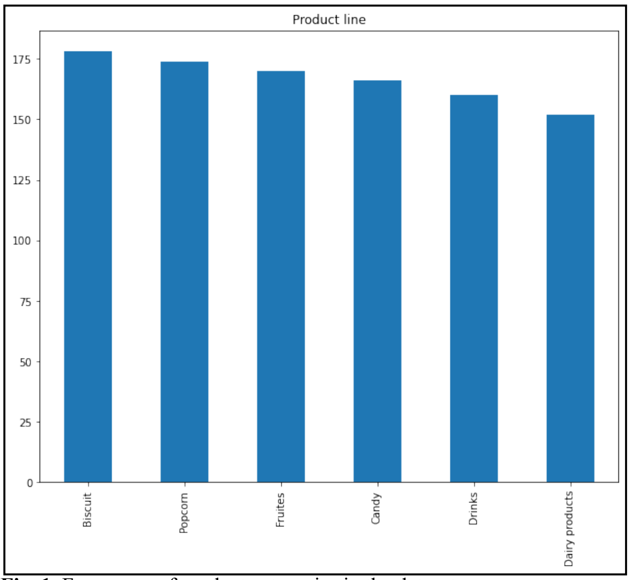

Fig. 1. Frequency of product categories in the dataset.

Table 1. Part of the First Dataset

Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Date	Time	Cost	Gross income	Rating	Total
765-26-6951	A	Yangon	Normal	Male	Candy	72.61	6	21.783	1/1/2019	10:39	435.66	21.783	6.9	457.443
530-90-9855	A	Yangon	Member	Male	Drinks	47.59	8	19.036	1/1/2019	14:47	380.72	19.036	5.7	399.756
891-01-7034	B	Mandalay	Normal	Female	Fruites	74.71	6	22.413	1/1/2019	19:07	448.26	22.413	6.7	470.673
493-65-6248	C	Naypyitaw	Member	Female	Candy	36.98	10	18.49	1/1/2019	19:48	369.8	18.49	7	388.29
556-97-7101	C	Naypyitaw	Normal	Female	Fruites	63.22	2	6.322	1/1/2019	15:51	126.44	6.322	8.5	132.762
133-14-7229	C	Naypyitaw	Normal	Male	Dairy products	62.87	2	6.287	1/1/2019	11:43	125.74	6.287	5	132.027
651-88-7328	A	Yangon	Normal	Female	Biscuit	65.74	9	29.583	1/1/2019	13:55	591.66	29.583	7.7	621.243
182-52-7000	A	Yangon	Member	Female	Candy	27.04	4	5.408	1/1/2019	20:26	108.16	5.408	6.9	113.568
416-17-9926	A	Yangon	Member	Female	Fruites	74.22	10	37.11	1/1/2019	14:42	742.2	37.11	4.3	779.31

Fig. 1 shows the frequency of food categories in the dataset. It is observed that there is a convergence in the distribution of numbers in this dataset.

Fig. 2 shows the relationship between the quantity of products and their category, as the average quantity of products in the dataset was taken.

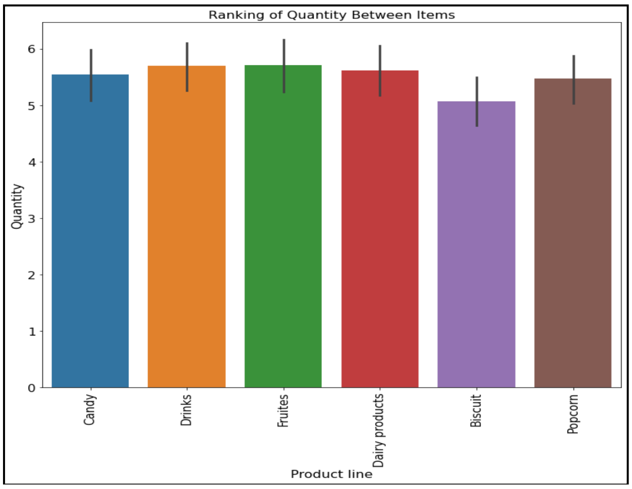

Fig. 2. The Relationship between Product Category and Quantity in the dataset.

Fig. 3 includes a 3D chart showing the relationship between product category, quantity, and total price of the first dataset.

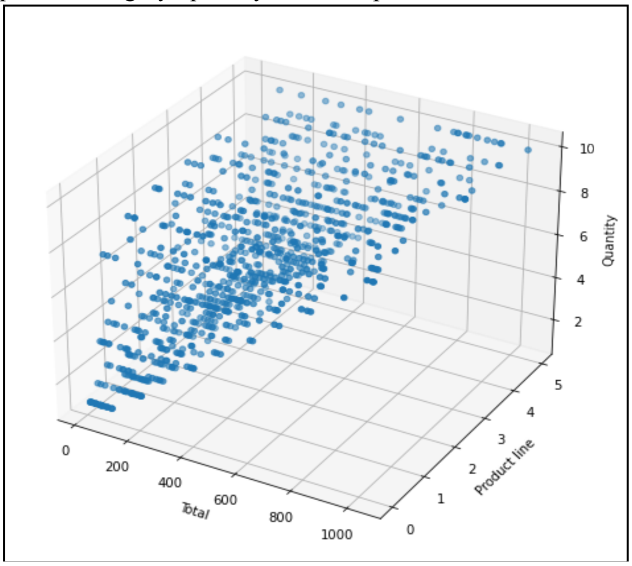

Fig. 3. The Relationship between Product Category, Quantity, and Total price in the dataset.

3.2 Algorithms used

To highlight an important technique, the Ridge technique, this study was proposed. Three artificial intelligence algorithms trace their origins to the Ridge concept, namely: Ridge Regression, Bayesian Ridge, and Kernel Ridge. These algorithms differ significantly in their core principles and also exhibit variations in other essential aspects. Below is a summary of the functionality of each of them individually:

3.2.1 Ridge Regression

Ridge Regression is an ML algorithm used to estimate the coefficients of multiple regression models in cases where the independent variables in the data are highly correlated [14]. This method is used to analyze data that suffer from multicollinearity. The multicollinearity problem causes the variances to be large, resulting in predicted values that are far from the actual values. As mentioned earlier, LASSO uses L1 regularization while Ridge Regression uses L2 regularization. Ridge Regression uses a cost function similar to the LASSO algorithm, except that Ridge adds a "square expression" to the coefficient as a penalty term [15]:

$$\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{1}$$

Here, the lambda value is the critical value. It should not be equal to zero, and it should not be too large in order not to add too much weight, which leads to under-fitting. This algorithm works efficiently to avoid overfitting.

The first action that this algorithm takes is standardizing the variables. Where it works to standardize the independent and dependent variables by subtracting their averages and dividing the result by the standard deviations [16]. To distinguish whether the variables are standardized or not, an annotation is placed in a formula. Therefore, to avoid notations indicating whether individual variables are standardized or not, this algorithm always uses standardized variables. Also, in the end, coefficients can be returned to their original scales.

Ridge Regression uses the ridge estimator as a shrinkage tool [17]. When there is a multicollinearity in the data, the ridge estimator is used to shrink the least squares estimator to obtain a better estimate. **Fig. 4** shows the accuracy of the prediction results obtained from this algorithm compared to the actual values.

Fig. 4. Results obtained from the Ridge Regressor algorithm.

3.2.2 Bayesian Ridge

Bayesian is a technique used to define and estimate statistical models [18]. The benefit of Bayesian regression appears when the data is not well distributed or when the data is not sufficient at all. This algorithm is based on the probability distribution to get the prediction [19]. The objective 'y' is obtained from a normal distribution (where the mean and variance are normalized). Bayesian Regression aims to find the "posterior" distribution of the model's parameters. Posterior is the probability of an event occurring as a result of another event that has already occurred [20]. This is equivalent to Bayes' theorem which states [21]:

$$P(A \mid B) = \frac{P(B|A)\,P(A)}{P(B)} \quad (2)$$

Here, the value of P(A) represents the probability that event A will occur, and P(A | B) is the probability that event A will occur because event B has already occurred. Based on the above formula, noticed that there is a subsequent distribution of the model parameters. This distribution is proportional to the probability of multiplying the data by the previous probability of the parameters. It is worth noting that increasing the number of data points leads to a significant increase in the probability value compared to the previous value. When this algorithm works to cover more data points, the built model becomes less erroneous. So, Bayesian Ridge needs a large amount of training data to make the model more accurate. For a fully probabilistic model with Bayesian regression, a Gaussian distribution around X_w is used to obtain the output "y" as shown in the equation below [22]:

$$p(y|X,w,\alpha) = N(y|Xw,\alpha) \quad (3)$$

Where alpha is treated as a random variable estimated from the data. Bayesian Ridge estimation is a probabilistic model of the regression problem. As indicated in the following equation [22]:

Where α and λ are chosen to be the gamma distributions.
The model used for this equation is called Bayesian Ridge Regression, and it is efficient for low-dimensional data. It is similar to the classic Ridge model.
Fig. 5 shows the accuracy of the prediction results obtained from this algorithm compared to the actual values.

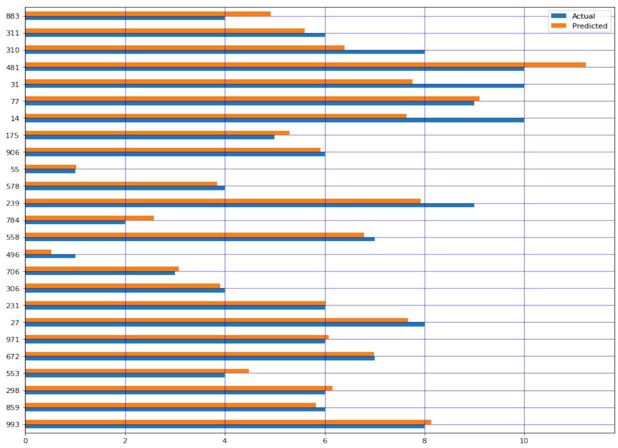

Fig. 5. Results obtained from Bayesian Ridge algorithm.

3.2.3 Kernel Ridge

The Kernel ridge regression (KRR) algorithm works by combining both the functions of classical ridge regression and classification (linear least squares with L2-norm regularization) on the one hand and the kernel trick on the other [23]. This algorithm learns a linear function in the resulting space from the kernel and the respective data.

The shape of the model learned by the KRR is similar to the shape of the model produced by the Support Vector Regression (SVR) algorithm [24]. However, SVR differs from KRR by the loss functions used, since KRR works with squared error loss while SVR uses loss Ɛ-insensitive, and both operate with L2 regularization. Kernel Ridge fit is performed in a closed form, unlike SVR, in which KRR is faster on medium-sized datasets [25]. In contrast, the model learned from the KRR is slower than SVR because it is not sparse [26].

The kernel trick works well with linear data provided the number of features is large and the number of inputs is small [27]. What distinguishes the kernel ridge regression algorithm is the existence of formulas that enable the computation of the mean squared error of the leave using the results obtained from a single individual training over the entire training set, i.e. without performing the leave-one-out.

Thus, efficient hyper-parameters (ridge and kernel parameters) can be obtained. In addition, if a single value decomposition operation is performed, solutions matching many values can be calculated using a single training edge [28]. Applying this mechanism makes ridge improvement highly effective. Fig. 6 shows the accuracy of the prediction results obtained from this algorithm compared to the actual values.

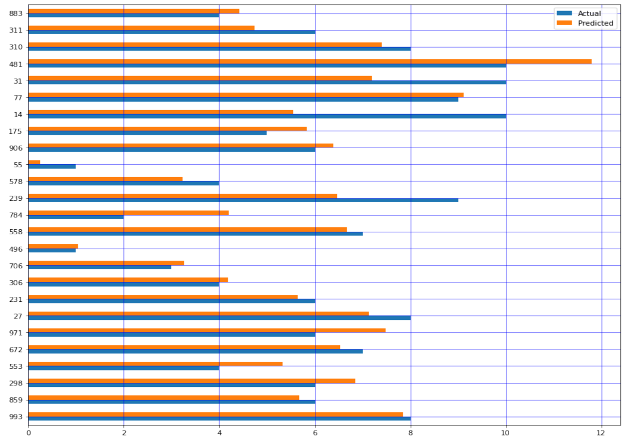

Fig. 6. Results obtained from Kernel Ridge algorithm.

4. Results and Discussion

This study presents a model based on three convergent and important artificial intelligence algorithms. The aim is to find satisfactory results to predict the amount of food needed by each company, each restaurant, or the like, to reduce food spoilage, thus saving these companies from losing money. To compare the results of the algorithms and determine the accuracy of the prediction obtained, several metrics were used, which will be detailed later.

4.1 R-Squared

R-squared (R^2) is a statistical equation that determines the percentage of variance in the dependent variable that can be predicted by the independent variable(s) [29]. This metric gives the proportion of the fit of the data to the regression model. Its value is between (−∞-1) and was multiplied by 100 in this study to know the percentage of accuracy obtained [30]. R^2 can be explained mathematically as follows [31]:

$$R^2 = \frac{Variance explained by model}{Total variance} \quad (5)$$

4.2 Mean absolute error (MAE)

The mean absolute error (MAE) is a metric used with regression models. Where it results in the average absolute difference between the expectation obtained from the model and the target value. MAE can be explained mathematically as follows [32]:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2 \quad (6)$$

where n is the observation number, y_i is the observed value and $\hat{y}_i$ is the predicted value.

4.3 Mean squared error

Mean squared error (MSE) is a metric used with regression models. MSE represents how close the regression line is to the data points. where it is considered a risk function. The lower its value, the more accurate the model. MSE can be explained mathematically as follows [33]:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2 \quad (7)$$

where n is the observation number, Y_i is the observed value and Y'_i is the predicted value.

4.4 Root mean squared error

Root mean squared error (RMSE) is obtained through the root of the MSE. It is a common measure used with regression models. The following equation shows it mathematically [34]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (8)$$

Using the above standards, the picture becomes clearer, and the results of these algorithms can easily be compared and the accuracy obtained from them can be easily determined.

When applying the above metrics to the three algorithms used, the metrics (MSE, RMSE, MAE) yielded results close to zero in all the algorithms, indicating a low margin of error in prediction. When using the R2 metric, the results obtained, as referred to in Table 2 below, were multiplied by 100. Whereas as the accuracy approaches 100, this means that the algorithm provided a better prediction accuracy.

Table 2. The Results of the Algorithms Using the Selected Dataset.

Algorithm	RMSE	MSE	MAE	Accuracy (R2)
Ridge Regressor	1.1684	1.3652	0.8384	82.87
Bayesian Ridge	1.1646	1.3565	0.8316	82.98
Kernel Ridge	1.5384	2.3668	1.1497	70.30

From observing the results in Table 2, it was found that the best prediction algorithm among the three algorithms used was (Bayesian Ridge), which gave an accuracy of (82.98). In general, the three algorithms gave good and important results, through which they can predict the actual need for food products to curb food waste.

5. Conclusion

Restaurants and companies are making every effort to mitigate food waste and, consequently, increase profits. This study proposed a model based on comparing the predictive results of three machine learning algorithms that haven't been compared in previous studies: Ridge Regressor, Kernel Ridge, and Bayesian Ridge. A multi-feature dataset was used to make the processing more effective. Several accuracy metrics were used to measure the performance of the proposed algorithms. The results revealed that the Bayesian Ridge algorithm performed the best with an accuracy of 82.98. The proposed system can be further improved by updating and expanding the dataset to include more stores and restaurants. Additionally, other algorithms can be compared to the ones proposed in this study.

References

[1] M. Santalova, I. Soklakova, D. Balakhanova, E. Lesnikova, and E. Chudakova, "Target organizational structure and human potential," *SHS Web Conf.*, vol. 101, p. 02009, Apr. 2021, https://doi.org/10.1051/shsconf/202110102009

[2] T. Chakraborty, S. S. Chauhan, and M. Ouhimmou, "Cost-sharing mechanism for product quality improvement in a supply chain under competition," *Int. J. Prod. Econ.,* vol. 208, pp. 566-587, Feb., 2019, https://doi.org/10.1016/j.ijpe.2018.12.015

[3] M. F. Ibrahim, Y. S. Mardhiyyah, A. Rusdiansyah, M. K. Boer, and D. M. Utama, "A three-phased perishable inventory simulation model with quality decrease consideration," *J. Ilmiah Teknik Ind.,* vol. 19, no. 2, pp. 198-211, Dec., 2020, https://doi.org/10.23917/jiti.v19i2.11769

[4] B. Kuswandi, M. Moradi, and P. Ezati, "Food sensors: Off-package and on-package approaches," *Packaging Tech. and Sci.*, vol. 35, no. 12, pp. 847-862, Aug. 2022, https://doi.org/10.1002/pts.2683

[5] W. Wenji, "Study on inventions of fresh food in commercial aspects using e-commerce over internet," *Acta Agri. Scandinavica, Section B—Soil & Plant Sci.*, vol. 71, no. 4, pp. 303-310, Jan. 2021, https://doi.org/10.1080/09064710.2021.1880625

[6] H. Dai, H. Yu, Q. Xiao, and W. Zhou, "A Clustering-based Sales Forecast Method for Big Promotion Days in O2O On-Demand Retailing," in *2019 IEEE Int. Conf. on Ind. Eng. and Eng. Manag. (IEEM)*, Macao, Macao:IEEE, Dec., 2019, pp. 1079-1083. DOI: 10.1109/IEEM44572.2019.8978778

[7] C.-H. Wang, "Considering economic indicators and dynamic channel interactions to conduct sales forecasting for retail sectors," *Comp. & Ind. Eng.*, vol. 165, p. 107965, Mar. 2022, https://doi.org/10.1016/j.cie.2022.107965

[8] R. Odegua, "Applied Machine Learning for Supermarket Sales Prediction," Project: Predictive Machine Learning in Industry, Jan., 2020. Availabe: https://www.researchgate.net/publication/338681895

[9] A. P. Wellens, R. N. Boute, and M. Udenio, "A Tree-based Framework to Democratize Large-scale Retail Sales Forecasting with Big Data," *SRN Journal*, 2022, Sep., 2022, Doi: 10.2139/ssrn.4213618.

[10] J. Wolters and A. Huchzermeier, "Joint in-season and out-of-season promotion demand forecasting in a retail environment," *J. Retailing*, vol. 97, no. 4, pp. 726-745, Dec., 2021. https://doi.org/10.1016/j.jretai.2021.01.003

[11] N. Vairagade, D. Logofatu, F. Leon, and F. Muharemi, "Demand Forecasting Using Random Forest and Artificial Neural Network for Supply Chain Management," in *Computat. Collect. Intellig.*, vol. 11683, N. T. Nguyen, R. Chbeir, E. Exposito, P. Aniorté, and B. Trawiński, Eds., in Lecture Notes in Computer Science, vol. 11683. , Cham: Springer International Publishing, 2019, pp. 328–339. doi: 10.1007/978-3-030-28377-3_27.

[12] V. Adithya Ganesan, S. Divi, N. B. Moudhgalya, U. Sriharsha, and V. Vijayaraghavan, "Forecasting Food Sales in a Multiplex Using Dynamic Artificial Neural Networks," in *Adv.Comp. Vis.*, vol. 944, K. Arai and S. Kapoor, Eds., in Advances in Intelligent Systems and Computing, vol. 944. , Cham: Springer International Publishing, 2020, pp. 69–80. doi: 10.1007/978-3-030-17798-0_8.

[13] L. Zhou, C. Zhang, F. Liu, Z. Qiu, and Y. He, "Application of deep learning in food: a review," *Compreh. reviews in food sci. food safety*, vol. 18, no. 6, pp. 1793-1811, Sep,.2019. https://doi.org/10.1111/1541-4337.12492

[14] M. Suhail, S. Chand, and B. G. Kibria, "Quantile based estimation of biasing parameters in ridge regression model," *Comm. Stat.-Simul. Comp.*, vol. 49, no. 10, pp. 2732-2744, Jan., 2020. https://doi.org/10.1080/03610918.2018.1530782

[15] P. Bharambe, B. Bagul, S. Dandekar, and P. Ingle, "Used Car Price Prediction using Different Machine Learning Algorithms," *Int. J. Res. Appl. Sci. Eng. Tech.,* vol. 10, pp. 773-778, Apr., 2022. https://doi.org/10.22214/ijraset.2022.41300

[16] S. Çankaya, S. Eker, and S. H. Abacı, "Comparison of Least Squares, Ridge Regression and Principal Component approaches in the presence of multicollinearity in regression analysis," *Turkish J. Agri.-Food Sci. Tech.*, vol. 7, no. 8, pp. 1166-1172, Aug., 2019,https://doi.org/10.24925/turjaf.v7i8.1166-1172.2515

[17] R. L. Obenchain, "Efficient generalized ridge regression," *Open Stat.*, vol. 3, no. 1, pp. 1-18, Nov.,2022, https://doi.org/10.3390/math11112522

[18] R. McElreath, Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC, Nov., 2020, https://doi.org/10.1111/rssa.12755

[19] P. Arora, D. Boyne, J. J. Slater, A. Gupta, D. R. Brenner, and M. J. Druzdzel, "Bayesian networks for risk prediction using real-world data: a tool for precision medicine," *Value in Health,* vol. 22, no. 4, pp. 439-445, Apr., 2019.https://doi.org/10.1016/j.jval.2019.01.006

[20] Y. Jin, W. Fu, J. Kang, J. Guo, and J. Guo, "Bayesian symbolic regression," arXiv preprint arXiv:1910.08892, Jan., 2019. https://doi.org/10.48550/arXiv.1910.08892

[21] R. van de Schoot et al., "Bayesian statistics and modelling," *Nat. Reviews Meth. Primers*, vol. 1, no. 1, p. 1, Jan., 2021. https://doi.org/10.1038/s43586-020-00001-2

[22] Y. YANG and Y. YANG, "Hybrid Prediction Method for Wind Speed Combining Ensemble Empirical Mode Decomposition and Bayesian Ridge Regression." *IEEE Access,* vol. 8 ,pp. 71206 - 71218, Mar., 2020, DOI: 10.1109/ACCESS.2020.2984020

[23] I. Mihaylov, M. Nisheva, and D. Vassilev, "Application of machine learning models for survival prognosis in breast cancer studies," *Information*, vol. 10, no. 3, p. 93, Mar., 2019, https://doi.org/10.3390/info10030093

[24] A. Shahsavar, M. Jamei, and M. Karbasi, "Experimental evaluation and development of predictive models for rheological behavior of aqueous Fe3O4 ferrofluid in the presence of an external magnetic field by introducing a novel grid optimization based-Kernel ridge regression supported by sensitivity analysis," *Powder Tech.*, vol. 393, pp. 1-11, Nov., 2021, https://doi.org/10.1016/j.powtec.2021.07.037

[25] P. N. Suganthan, "On non-iterative learning algorithms with closed-form solution," *Applied Soft Comp*., vol. 70, pp. 1078-1082, Sep., 2018. https://doi.org/10.1016/j.asoc.2018.07.013

[26] H. Mai, T. C. Le, D. Chen, D. A. Winkler, and R. A. Caruso, "Machine learning for electrocatalyst and photocatalyst design and discovery," *Chem. Reviews,* vol. 122, no. 16, pp. 13478-13515,Jul., 2022, https://doi.org/10.1021/acs.chemrev.2c00061

[27] T. Kusumoto, K. Mitarai, K. Fujii, M. Kitagawa, and M. Negoro, "Experimental quantum kernel trick with nuclear spins in a solid," npj *Quantum Info*., vol. 7, no. 1, p. 94, Jun., 2021. https://doi.org/10.1038/s41534-021-00423-0

[28] M. Kapralov, N. Nouri, I. Razenshteyn, A. Velingker, and A. Zandieh, "Scaling up Kernel Ridge Regression via Locality Sensitive Hashing," arXiv preprint arXiv:2003.09756, Mar., 2020. https://doi.org/10.48550/arXiv.2003.09756

[29] R. Valbuena et al., "Evaluating observed versus predicted forest biomass: R-squared, index of agreement or maximal information coefficient?," *European J. Remote Sens*., vol. 52, no. 1, pp. 345-358, May, 2019, https://doi.org/10.1080/22797254.2019.1605624

[30] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comp. Sci.*, vol. 7, p. e623, Jun., 2021, https://doi.org/10.7717/peerj-cs.623

[31] F. Rustam et al., "COVID-19 future forecasting using supervised machine learning models," *IEEE Access,* vol. 8, pp. 101489-101499, May, 2020, DOI: 10.1109/ACCESS.2020.2997311

[32] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. dos Santos Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil," *Chaos, Solitons & Fractals*, vol. 135, p. 109853, Jun., 2020. https://doi.org/10.1016/j.chaos.2020.109853

[33] Y. Amellas, A. Djebli, and A. Echchelh, "Levenberg-Marquardt Training Function using on MLP, RNN and Elman Neural Network to Optimize Hourly Forecasting in Tetouan City (Northern Morocco)," *J. Eng. Sci. and Tech. Rev.*, vol. 13, no. 1, pp. 67-71, Jan., 2020, doi:10.25103/jestr.131.09

[34] Q. Wei, L. Zhang, W. Duan, and Z. Zhen, "Global and geographically and temporally weighted regression models for modeling PM2. 5 in Heilongjiang, China from 2015 to 2018," *Int. J. Env. Res. Publ. Health*, vol. 16, no. 24, p. 5107, Dec., 2019, https://doi.org/10.3390/ijerph16245107