Modeling Energy Consumption Using Machine Learning

Sai Aravind Sarswatula[1], Tanna Pugh[2] and Vittaldas Prabhu[1,3]*

[1]*Department of Industrial, Pennsylvania State University, University Park, PA, United States,* [2]*Pennsylvania Technical Assistance Program, University Park, PA, United States,* [3]*Marcus Department of Industrial and Manufacturing Engineering, University Park, PA, United States*

Electrical, metal, plastic, and food manufacturing are among the major energy-consuming industries in the U.S. Since 1981, the U.S. Department of Energy Industrial Assessments Centers (IACs) have conducted audits to track and analyze energy data across several industries and provided recommendations for improving energy efficiency. In this article, we used statistical and machine learning techniques to draw insights from this IAC dataset with over 15,000 samples collected from 1981 to 2013. We developed predictive models for energy consumption using machine learning techniques such as Multiple Linear Regression, Random Forest Regressor, Decision Tree Regressor, and Extreme Gradient Boost Regressor. We also developed classifier models using Support Vector Machines, Random Forest, K-Nearest Neighbor (KNN), and deep learning. Results using this data set indicate that Random Forest Regressor is the best prediction technique with an R^2 of 0.869, and the Random Forest classifier is the best technique with precision, recall, F1 score, and accuracy of 0.818, 0.884, 0.844, and 0.883, respectively. Deep learning also performed competitively with an accuracy of about 0.88 in training and testing after 10 epochs. The machine learning models could be useful in benchmarking the energy consumption of factories and identifying opportunities to improve energy efficiency.

Keywords: machine learning, sustainability, energy footprint, industries, statistics

INTRODUCTION

Life Cycle Engineering (LCE) generally aims to enhance sustainability and in this article, we focus specifically on its energy aspects. Energy-aware manufacturing can lead to energy efficiency and can be viewed at different levels including product level, machine level, factory level, and supply chain level as shown in **Figure 1**.

At each of these levels, there are a number of decisions that can support LCE to enhance sustainability. With increasing adoption of smart systems in Industry 4.0, a variety of technologies such as networked sensors, elastic cloud computing, and machine learning are enabling an entirely new class of energy-aware manufacturing systems (Prabhu et al., 2015; Sarkis and Zhu, 2018). Key performance indicators for energy efficiency are also important (May et al., 2013). Numerous studies have focused on the energy consumption of specific products (Sim and Prabhu, 2018) and machines (Calvanese et al., 2013). In discrete manufacturing systems consisting of multiple machines, energy consumption at the factory level can be expected to be influenced by the higher-level production control systems and their associated policies, which can be studied using simulation and analytical models (Prabhu and Taisch, 2012, Prabhu et al., 2012). Energy-aware feedback control for production scheduling and capacity control has been proposed (Lee and Prabhu 2015). Similarly, a green genetic algorithm has been proposed for multi-objective problems that arise at
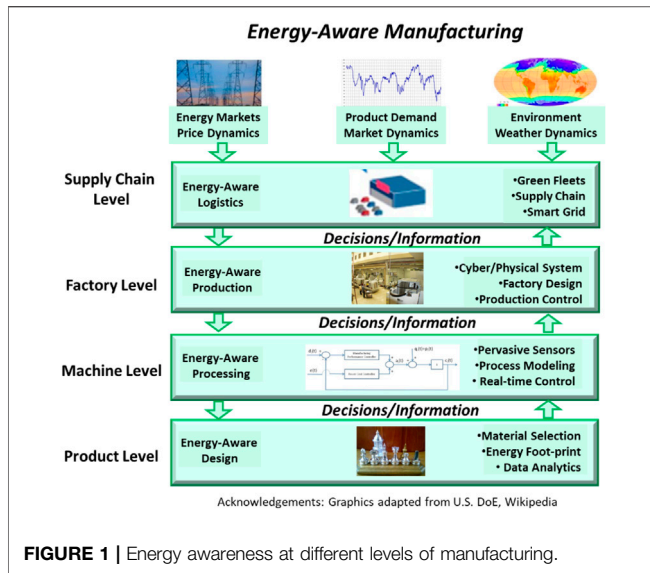
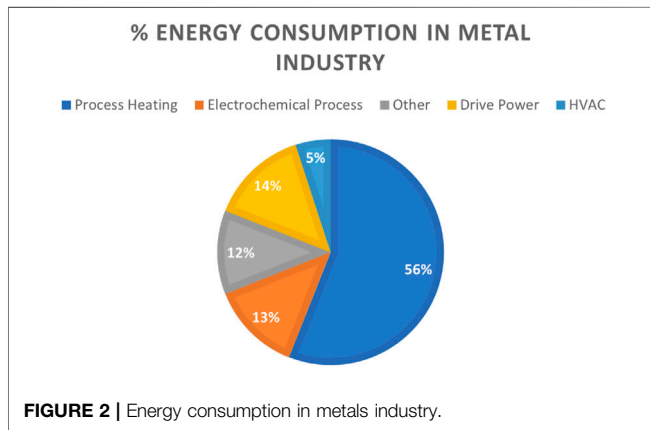FIGURE 1 | Energy awareness at different levels of manufacturing.

FIGURE 2 | Energy consumption in metals industry.

TABLE 1 | Specific energy consumption in the plastics industry.

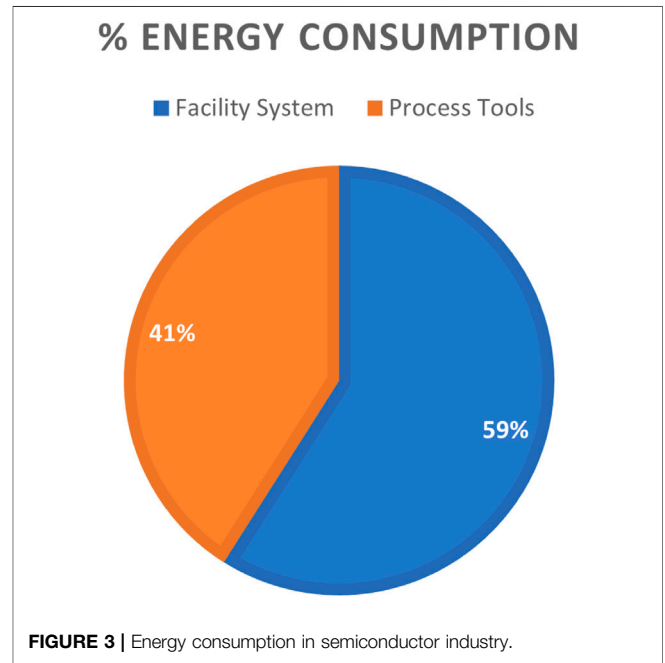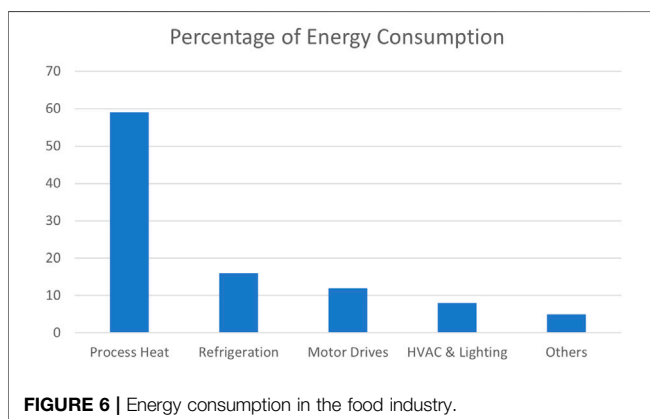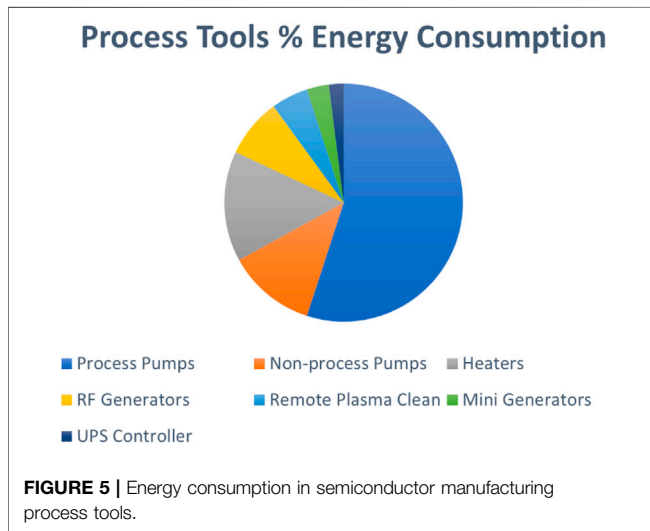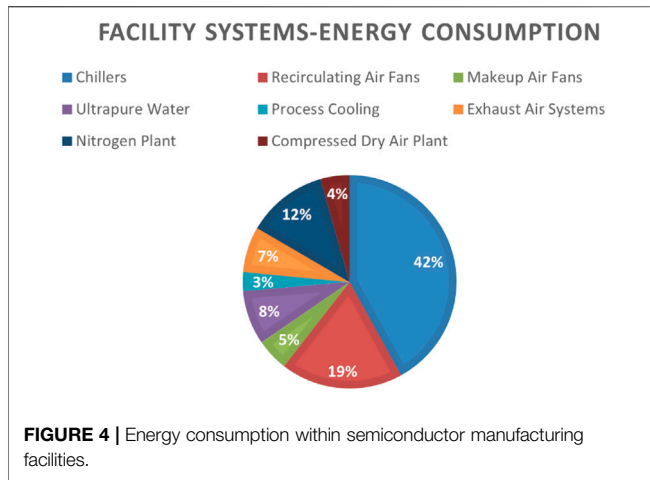| Plastic processing type	Specific energy usage in kwh/kg
Injection molding	3.118
Pipe and profile extrusion	1.506
Film extrusion	1.346
Thermoforming	6.179
Rotational moulding	5.828
Compressional moulding	3.168
Fiber extrusion	0.850
Compounding	0.631

FIGURE 3 | Energy consumption in semiconductor industry.

the confluence of productivity and sustainability in manufacturing operations (May et al., 2015). Recently a survey on sustainability through scheduling of manufacturing operations has outlined some of the key challenges (Giret et al., 2015). Modeling approaches have also been suggested for benchmarking the footprint of manufacturing sectors, which can be used to drive energy efficiency (Jeon et al., 2015; Boyd et al., 2008). From a manufacturing supply chain perspective, there are several opportunities to improve sustainability through energy efficiency in transportation and distribution (Lee et al., 2016; Anand et al., 2014). This can be extended in scope for architecting highly distributed intelligent operation of manufacturing systems consisting of production processes, lighting, and HVAC systems, renewable energy systems, and energy storage systems to cooperatively achieve load control across the smart grid (Prabhu et al., 2012; Schoonenberg and Farid, 2017). Models for the variability of electrical power demand in a variety of manufacturing operations have been studied, which can be helpful for managing peak demand at a factory or grid level (Jeon et al., 2016).

In the U.S. electrical, metal, plastic, and food manufacturing are some of the major energy-consuming industries. Metal industries consist of metal casting, sheet metal, metal welding, etc. Among different kinds of metals, steel industries account for the majority. Since World War II, energy consumption in the steel industry has reduced significantly by improving practices such as pelletizing, sintering, coke making, etc (Stubbles, 2000). In **Figure 2**, we can see that process heating consumes the most while HVAC consumes the least in primary metal industries (Business Energy Advisor, 2020).

The plastics industry includes major operations such as injection molding, CNC machining, and blow molding. Installing large storage tanks and improving water cooling systems are among the top two recommendations made by DOE to save energy costs for plastic industries (The Society of the Plastics Industry and Inc and the U.S Department of Energy, 2005).

Table 1 shows different types of plastic processing and specific energy usage. Thermoforming processing has the highest specific energy usage while compounding processing has the lowest

FIGURE 4 | Energy consumption within semiconductor manufacturing facilities.

FIGURE 5 | Energy consumption in semiconductor manufacturing process tools.

FIGURE 6 | Energy consumption in the food industry.

specific energy usage (Maximpact, 2020). In the electrical manufacturing industry which largely consists of semiconductor and electronic components manufacturing, major energy consumption occurs in HVAC, chillers, and

process equipment. Improving these can lead to efficiency in energy consumption (Chen and Gautam, 2013).

In **Figure 3**, we can see that energy consumption in semiconductor industry can be divided into facility and process tools. Within facilities, chillers and recirculating air fans consume the most. In process tools, process pumps consume the most energy, as shown in **Figures 4**, **5** (Plepys, 2004). In food processing industries, wet corn milling and beet sugar are the most energy-intensive industries (Drescher et al., 1997). Nearly 50% of the energy is consumed during the process of heating and cooling in food industries (Bracco, 2015; as shown in **Figure 6**).

Since 1981 the U.S. ; Industrial Assessments Centers (IACs) have conducted audits to track and analyze energy data across several industries and provided recommendations for improving energy efficiency. This article aims to draw insights from this IAC dataset with over 15,000 samples collected from 1981 to 2013 as a basis for modeling energy consumption using machine learning. Such machine learning models could be useful in benchmarking the energy consumption of factories and identifying opportunities to improve energy efficiency. In Section 2 we review extant literature at the intersection of machine learning and energy modeling. In Section 3 we present an exploratory data analysis of the IAC data set. In Section 4 we develop predictive models for energy consumption using machine learning techniques such as Multiple Linear Regression, Random Forest Regressor, Decision Tree Regressor, and Extreme Gradient Boost Regressor. We also developed classifier models using Support Vector Machines, Random Forest, K-Nearest Neighbor (KNN), and deep learning. Section 4 and 5 present results and discussion, respectively. Section 6 presents conclusions and possible directions for future work.

2 LITERATURE REVIEW

A literature survey of machine learning models for energy systems has been published (Mosavi et al., 2019) (Mosavi et al., 201,9). Similarly, a survey of studies covering predicting and forecasting electrical energy consumption in the manufacturing industry has also been published recently (Walther and Weigold, 2021). Morariu et al. propose an approach for predicting energy consumption patterns using Long Short-term Memory neural networks and deep learning in real-time for batch cost optimization by reassigning resources and detecting anomalies (Morariu et al., 2020) (Morariu et al., 2020). Deep learning has also been proposed for demand-side energy consumption forecasting, and the Gated Recurrent Unit is found to be superior to other machine learning and manual technique in the case of a Brazilian thermoplastic resin manufacturing plant (Ribeiro et al., 2020). Similarly, for a chemical plant, machine learning techniques are proposed for predicting machine-specific load profiles using energy disaggregation, which is in turn used to predict the machine's activity state and the respective production capacities (Tan et al., 2021) (Tan et al., 2021). Predictive intelligence of the smart manufacturing approach has been
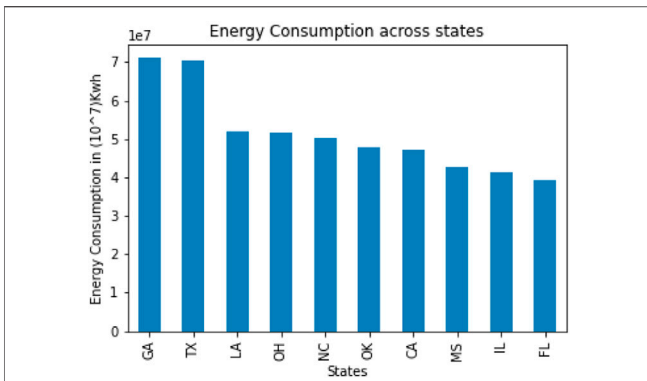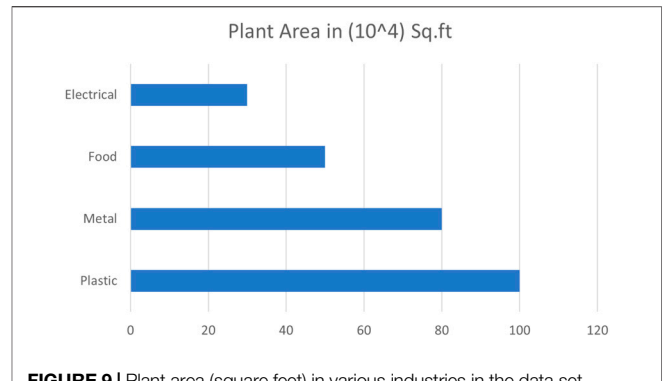
FIGURE 7 | Energy consumption across states.

FIGURE 8 | Electricity usage trend in the data set.

FIGURE 9 | Plant area (square feet) in various industries in the data set.
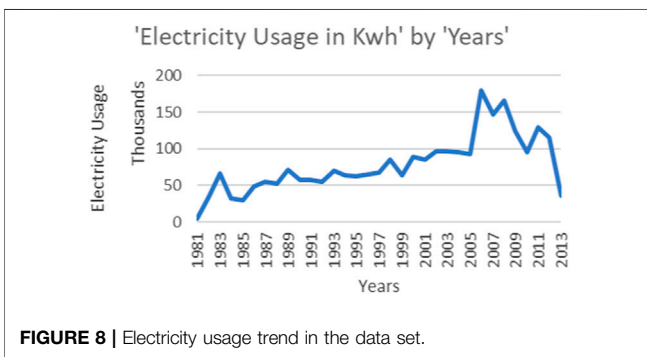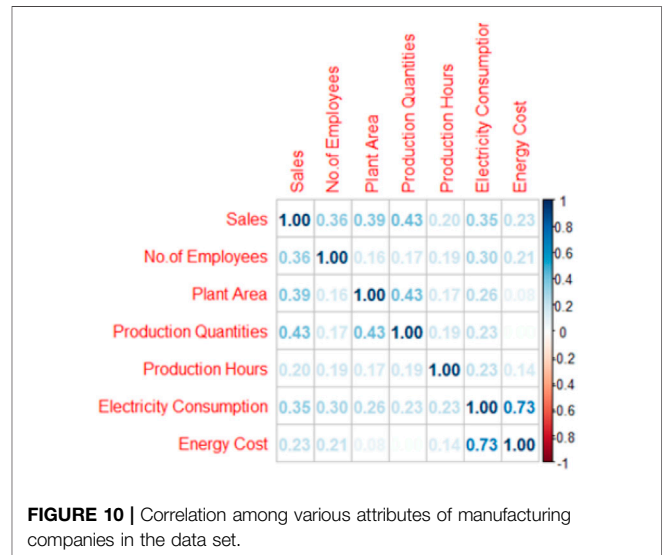

FIGURE 10 | Correlation among various attributes of manufacturing companies in the data set.

extended to predict energy consumption based on historical data (Essien et al., 2020; Kumar et al., 2021). Machine learning models were used to determine energy consumption in the metal forming process and are applied to the radial-axial ring rolling process (Mirandola et al., 2021). For complex additive manufacturing systems, a hybrid machine learning approach that integrates extreme gradient boosting (XGBoost) decision tree and density-based spatial clustering of applications with noise (DBSCAN) technique are applied (Li et al., 2021). Predictive models for machine energy efficiency and optimization tools to minimize energy consumption were developed for the application in the semiconductor industry (Kuo-Hao Chang et al., 2021) (Chang et al., 2022). Classification machine learning models like K-Nearest Neighbors, Neural Networks, Random Forest, and Support Vector Machines are used to validate the quality of production processes and their parameters in the food processing industry (Milczarski et al., 2020). In plastic-processing SME's, machine learning is applied to optimize energy consumption and reduce incorrectly produced plastic parts (Willenbacher et al., 2021).

3 EXPLORATORY DATA ANALYSIS

From the DOE IAC data set, data cleaning and exploratory data analysis are performed first.

Figure 7 shows the top 10 states in the U.S. with the highest energy consumption, with Georgia and Texas being the top two.

Figure 8 shows electricity usage according to years. There is a steady increase till 2007, after which there is a significant increase and a sudden drop during 2011–2013. This decline can be attributed to fewer data collected from 2011–2013 rather than any substantive change in the industries.

Figure 9 shows the sum of the plant area, in square feet, for various industries within the data set. We can see that plastic and metals have a considerably larger area within this data set. Plant area is an important attribute because it indicates the lighting and heating/cooling space, which in turn consume energy.

In **Figure 10**, which is a correlation plot, we can see that electricity consumption is highly correlated with energy cost, which is as expected. Electricity consumption is moderately correlated with sales, the number of employees, plant area, and production quantities in the data set.

To assess if there is a significant difference in energy consumption between food, metal, plastic, and electric industries, we use analysis of variance (ANOVA). From

TABLE 2 | Results of ANOVA.

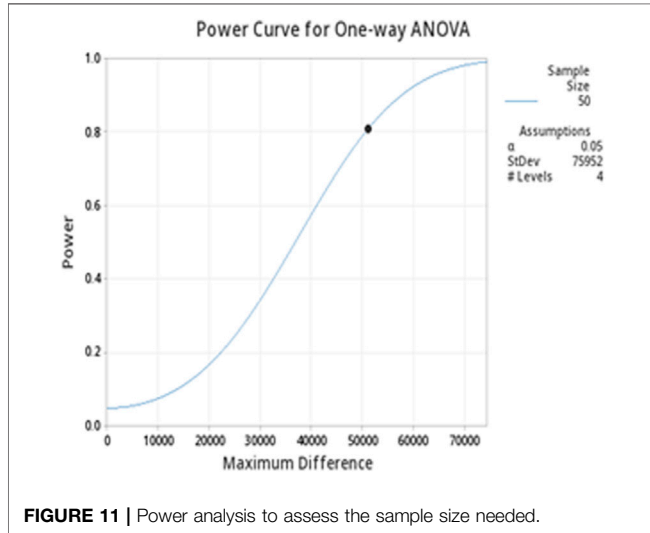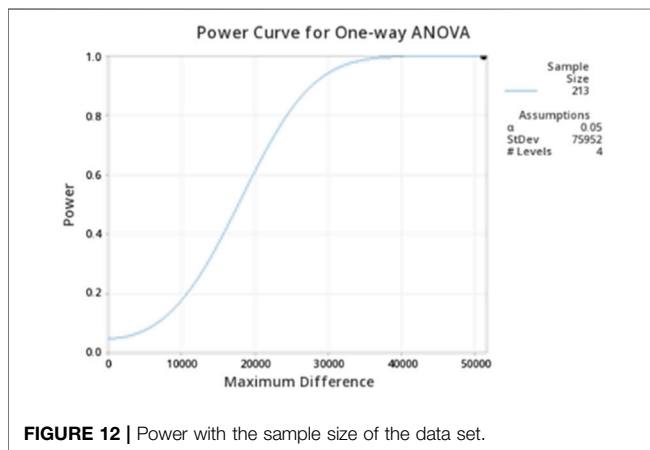| Source	DF	Adj SS	Adj MS	F-value	p-value
Type	3	1.15415E+12	3.84716E+11	55.74	0.00
Error	2085	1.43908E+13	6902085506		
Total	2088	1.55450E+13			

FIGURE 11 | Power analysis to assess the sample size needed.

FIGURE 12 | Power with the sample size of the data set.

Table 2, we can see that p value is less than 0.05, and we reject the null hypothesis. We are 95% confident that the energy consumption of these four industries is significantly different.

Null hypothesis: Means of Energy Consumption of the four industries are equal.

Alternative hypothesis: At least one industry has different mean energy consumption than others.

In order to assess if the sample size of the data set is adequate, we use power analysis. For getting a power of 80, we need a sample size of 50 companies in each of the four industries, as shown in **Figure 11**. The data set has a minimum of 213 samples in every industry group, which is adequate. **Figure 12** shows that with the given sample size of the data set the power is 0.99.

4 Machine Learning for Modeling Energy Consumption

In this section, we will apply a variety of machine learning techniques prediction and classification for modeling energy consumption. We provide a brief description of the techniques for completeness; mode detailed descriptions can be found in published literature (Predictive analytics, 2021). Software for all of these is developed using Python and open source libraries.

4.1.1 Prediction Approaches

For prediction, we use widely studied techniques of Multiple Linear Regression, Random Forest Regressor, Decision Tree Regressor, and Extreme Gradient Boost Regressor for predicting energy consumption (Towards Data Science, 2019; Machine Learning Mastery, 2020a; Predictive analytics, 2021).

I. Multiple Linear Regression: This models the relationship between the response variable, energy consumption in this article and other explanatory variable (Industry Type, Production Quantities, Number of employees, Plant area, etc.) by fitting a linear relationship to the data.

$$\mu(y) = \beta0 + \beta1 * x1 + \beta2 * x2 \ldots \ldots ..$$

μ indicates the mean of the response variable y, and x1, x2, x3…. indicate explanatory variables along with their corresponding coefficients β.

II. Decision Tree Regressor: This works by the greedy search approach. It follows a tree structure with the root nodes, interior nodes, and leaf nodes. It uses a set of binary rules to calculate a target value. Root nodes represent the entire sample, interior nodes represent explanatory variables, branches represent decision rules, and leaf nodes represent the outcome.

III. Random Forest Regressor: This is an extension of Decision Tree Regression. Here decision trees are created randomly, so it merges the output of multiple decision trees to generate the output.

IV. Extreme Gradient Boost Regressor: Extreme Gradient Boosting constructs ensembles from decision tree models. Prediction errors made by the prior models will be corrected by adding trees one at a time using ensemble and fit to minimize the loss of gradient.

4.1.2 Classification Approaches

We have built classifier models using Support Vector Machines, Random Forest, K-Nearest Neighbor (KNN), and deep learning (Machine Learning Mastery, 2020a).

I. Support Vector Machine (SVM): Each data sample is represented as a point in the n-dimensional sample space. SVM creates hyperplanes which differentiates the categories.

II. KNN Classifier: It works by considering K-Nearest Neighbors to predict the class of new datapoint. In our case, we have considered five neighbors.

III. Deep learning: We have built a sequential model with a dense layer with 100 input nodes with 2 hidden layers and used "Relu" as the activation function for input and hidden layers and "Softmax" as the activation function for the output layer. We used "Adam" as an optimizer and ran 10 epochs (Machine Learning Mastery, 2020b).

TABLE 3 | Prediction approach results.

Model	RMSE	R- square
Multiple linear regression	76787.2	0.742
Decision tree	75966.2	0.748
Random forest regressor	54761.4	0.869
XG Boost	62565.36	0.829

TABLE 4 | Classification approach results.

Model	Precision	Recall	F1 score	Accuracy
Support vector Machine	0.773	0.792	0.782	0.87
Random forest classifier	0.818	0.884	0.844	0.883
KNN	0.798	0.854	0.819	0.854

FIGURE 13 | Accuracy of deep learning with increasing epochs.

5 RESULTS

5.1 Prediction Approach Results

The results of prediction approaches are compared using metrics of the square root of the mean of errors (RMSE) and coefficient of determination (R^2). RMSE are residuals which indicate differences between actual and predicted values (Towards Data Science, 2019). R^2 measures the proportion of variance of a dependent variable that is explained by variables in the model. These results are summarized in **Table 3**.
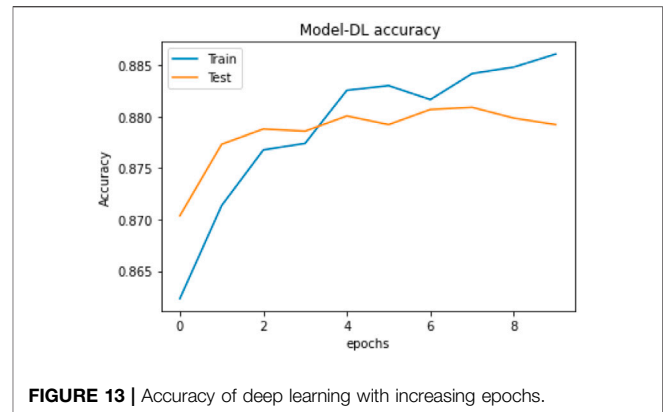
5.2 Classification Approach Results

The results of classification approaches are compared using metrics of accuracy, precision, recall, and F1 Score (Towards Data Science, 2019). Accuracy is the percentage of correct predictions whereas precision is the fraction of true positives with true positives and false positives. The recall is the fraction of true positives with true positives and false negatives. F1 Score is the harmonic mean of Precision and Recall. These results are summarized in **Table 4**.

5.2.1 Deep Learning Results

The results of interest in deep learning is how accuracy improves with increasing epochs for both training and testing for the data set. After 10 epochs, the accuracy of the training model is 0.89, while the accuracy of the test model is 0.87.

6 DISCUSSION

Ideally, a predictive model should have a high R^2 value and low RMSE. Based on this, for the data set used in this article the random forest regressor model outperforms other models with its high R^2 value and low RMSE, which would make it a good candidate for predicting energy consumption in new companies. Additionally, the random forest regressor method has the advantage that it reduces overfitting formed by the decision tree and automates missing values in the data without requiring normalization, which can be beneficial in practical industrial applications.

Among the machine learning classification models, Random Forest Classifier performed better than other models in all measures. Accuracy for SVM is higher than KNN, however, KNN is better than SVM for precision, recall, and F1 score. From **Figure 13**, we can see that the deep learning model has performed well with no case of overfitting or underfitting. However, this behavior might change as we increase the epochs. So, we recommend the Random Forest algorithm for classification tasks as well.

7 CONCLUSION

This article aims to provide applications of statistics and machine learning to achieve sustainability at the industry level by leveraging data collected by the US DOE IAC program in metal, electrical, food and plastic manufacturing. First, we used exploratory data analysis to clean data, visualize trends and draw insights. Next, we used ANOVA to verify that energy consumption in the four industries is statistically different. Power analysis is used to verify that the sample size in the data set is adequate for analysis. We developed predictive models for energy consumption using the following machine learning techniques: Multiple Linear Regression, Random Forest Regressor, Decision Tree Regressor, and Extreme Gradient Boost Regressor. Results indicate that Random Forest Regressor is the best technique with an R^2 of 0.869. We developed classifier models using Support Vector Machines, Random Forest, K-Nearest Neighbor (KNN), and deep learning. Random Forest classifier the other two techniques with precision, recall, F1 score, and accuracy of 0.818, 0.884, 0.844, and 0.883, respectively. Deep learning also performed competitively with an accuracy of about 0.88 in training and testing after 10 epochs.

Future work should focus on translating these results into an easy-to-use software application in which industry practitioners can quickly benchmark a factory by entering variables such as Annual Production Quantities, Annual Sales, Number of Employees, and Annual Operating Hours. If the predicted energy consumption is significantly lower than what is observed in practice then it is likely an indication that there are one or more opportunities to improve energy efficiency in the factory. Recommendations from prior projects relevant to the specific manufacturing industry can then

be used to focus energy efficiency improvement efforts. Based on several such industrial use-cases, the underlying models should be validated and disseminated for wider application to improve the sustainability of our societies.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: We do not own the data. Requests to access these datasets should be directed to https://www.energy.gov/eere/amo/industrial-assessment-centers-iacs.

AUTHOR CONTRIBUTIONS

SS: part of MS research, data analysis, software development; initial draft TP: editing draft; review of technical content and application VP: concept development; project management; final editing.

REFERENCES

Anand, V., Lee, S., and Prabhu, V. V. (2014). "Energy-aware Models for Warehousing Operations," in *IFIP International Conference on* Advances In Production Management Systems (Berlin, Heidelberg: Springer), 390–397. doi:10.1007/978-3-662-44736-9_48

Boyd, G., Dutrow, E., and Tunnessen, W. (2008). The Evolution of the ENERGY STAR Energy Performance Indicator for Benchmarking Industrial Plant Manufacturing Energy Use. *J. Clean. Prod.* 16 (6), 709–715. doi:10.1016/j.jclepro.2007.02.024

Bracco, S. (2015). *Opportunities for Agri-Food Chains to Become Energy-Smart*".

Business Energy Advisor (2020). source.bizenergyadvisor. Available at: https://esource.bizenergyadvisor.com/article/manufacturing-facilities.

Calvanese, M. L., Albertelli, P., Matta, A., and Taisch, M. (2013). "Analysis of Energy Consumption in CNC Machining Centers and Determination of Optimal Cutting Conditions," in *Re-engineering Manufacturing for Sustainability* (Singapore: Springer), 227–232. doi:10.1007/978-981-4451-48-2_37

Chang, K.-H., Sun, Y.-J., Lai, C.-A., Chen, L.-D., Wang, C.-H., Chen, C.-J., et al. (2022). Big Data Analytics Energy-Saving Strategies for Air Compressors in the Semiconductor Industry - an Empirical Study. *Int. J. Prod. Res.* 60 (6), 1782–1794. doi:10.1080/00207543.2020.1870015

Chen, S., and Gautam, A. (2013). *Bringing Energy Efficiency to the Fab*".

Drescher, S., Rao, N., Kozak, J., and Martin, O. (1997). *A Review of Energy Use in the Food Industry*. USA: American Council for an Energy-Efficient Economy-Summer Study on Energy Efficiency in Industry.

Essien, A., and Giannetti, C. (2020). A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders. *IEEE Trans. Ind. Inf.* 16 (9), 6069–6078. doi:10.1109/tii.2020.2967556

Giret, A., Trentesaux, D., and Prabhu, V. (2015). Sustainability in Manufacturing Operations Scheduling: A State of the Art Review. *J. Manuf. Syst.* 37, 126–140. doi:10.1016/j.jmsy.2015.08.002

Jeon, H. W., Taisch, M., and Prabhu, V. (2016). Measuring Variability on Electrical Power Demands in Manufacturing Operations. *J. Clean. Prod.* 137, 1628–1646. doi:10.1016/j.jclepro.2016.03.102

Jeon, H. W., Taisch, M., and Prabhu, V. V. (2015). Modelling and Analysis of Energy Footprint of Manufacturing Systems. *Int. J. Prod. Res.* 53 (23), 7049–7059. doi:10.1080/00207543.2014.961208

Kumar, M., Shenbagaraman, V. M., Shaw, R. N., and Ghosh, A. (2021). "Predictive Data Analysis for Energy Management of a Smart Factory Leading to Sustainability," in *Innovations in Electrical and Electronic Engineering* (Singapore: Springer), 765–773. doi:10.1007/978-981-15-4692-1_58

Lee, S., Kang, Y., and Prabhu, V. V. (2016). Smart Logistics: Distributed Control of Green Crowdsourced Parcel Services. *Int. J. Prod. Res.* 54 (23), 6956–6968. doi:10.1080/00207543.2015.1132856

Li, Y., Hu, F., Qin, J., Ryan, M., Wang, R., and Liu, Y. (2021). A Hybrid Machine Learning Approach for Energy Consumption Prediction in Additive Manufacturing. *ICPR Int. Work. Challenges, Lecture Notes Comput. Sci.*, 622–636. doi:10.1007/978-3-030-68799-1_45

Machine Learning Mastery (2020). Jason Brownlee. Available at: https://machinelearningmastery.com/types-of-classification-in-machine-learning/.

Machine Learning Mastery (2020). Jason Brownlee. Available at: https://machinelearningmastery.com/multi-label-classification-with-deep-learning/.

Maximpact (2020). Maximpact. Available at: https://www.maximpact.com/energy-efficiency-plastic/.

May, G., Stahl, B., Taisch, M., and Prabhu, V. (2015). Multi-objective Genetic Algorithm for Energy-Efficient Job Shop Scheduling. *Int. J. Prod. Res.* 53 (23), 7071–7089. doi:10.1080/00207543.2015.1005248

May, G., Taisch, M., Prabhu, V. V., and Barletta, I. (2013). "Energy Related Key Performance Indicators - State of the Art, Gaps and Industrial Needs," in *IFIP International Conference on* Advances In Production Management Systems (Berlin, Heidelberg: Springer), 257–267. doi:10.1007/978-3-642-41266-0_32

Milczarski, P., Zieliński, B., Stawska, Z., Hłobaż, A., Maślanka, P., and Kosiński, P. (2020). Machine Learning Application in Energy Consumption Calculation and Assessment in Food Processing Industry. *Artif. Intell. Soft Computing, Lecture Notes Comput. Sci.*, 369–379. doi:10.1007/978-3-030-61534-5_33

Mirandola, I., Berti, G. A., Caracciolo, R., Lee, S., Kim, N., and Quagliato, L. (2021). Machine Learning-Based Models for the Estimation of the Energy Consumption in Metal Forming Processes. *Metals* 11, 833. doi:10.3390/met11050833

Morariu, C., Morariu, O., Răileanu, S., and Borangiu, T. (2020). Machine Learning for Predictive Scheduling and Resource Allocation in Large Scale Manufacturing Systems. *Comput. Industry* 120, 103244. doi:10.1016/j.compind.2020.103244

Mosavi, A., Salimi, M., Faizollahzadeh Ardabili, S., Rabczuk, T., Shamshirband, S., and Varkonyi-Koczy, A. (2019). State of the Art of Machine Learning Models in Energy Systems, a Systematic Review. *Energies* 12 (7), 1301. doi:10.3390/en12071301

Plepys, A. (2004). *Environmental Implications of Product Servicising.-The Case of Outsourced Computing Utilities*. Lund University, Lund, Sweden

Prabhu, V. V., Jeon, H. W., and Taisch, M. (2012). "Modeling Green Factory Physics—An Analytical Approach," in *2012 IEEE International Confe*rence on Automation Science And Engineering (CASE) (IEEE), 46–51.

Prabhu, V. V. (2012). "Services for Competitive and Sustainable Manufacturing in the Smart Grid," in *Service Orientation in Holonic and Multi-Agent Manufacturing Control* (Berlin, Heidelberg: Springer), 227–240. doi:10.1007/978-3-642-27449-7_17

Prabhu, V. V., and Taisch, M. (2012). Simulation Modeling of Energy Dynamics in Discrete Manufacturing Systems. *IFAC Proc. Vol.* 45 (6), 740–745. doi:10.3182/20120523-3-ro-2023.00422

Prabhu, V. V., Trentesaux, D., and Taisch, M. (2015). Energy-aware Manufacturing Operations. *Int. J. Prod. Res.* 53 (23), 6994–7004. doi:10.1080/00207543.2015.1100766

Predictive analytics (2021). Predictive Analytics. (2021, November 8). In Wikipedia. Available at: https://en.wikipedia.org/wiki/Predictive_analytics.

Ribeiro, A. M. N. C., do Carmo, P. R. X., Rodrigues, I. R., Sadok, D., Lynn, T., and Endo, P. T. (2020). Short-term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: a Comparison of Machine Learning and Deep Learning Models. *Algorithms* 13 (11), 274. doi:10.3390/a13110274

Sarkis, J., and Zhu, Q. (2018). Environmental Sustainability and Production: Taking the Road Less Travelled. *Int. J. Prod. Res.* 56 (1-2), 743–759. doi:10.1080/00207543.2017.1365182

Schoonenberg, W. C. H., and Farid, A. M. (2017). A Dynamic Model for the Energy Management of Microgrid-Enabled Production Systems. *J. Clean. Prod.* 164, 816–830. doi:10.1016/j.jclepro.2017.06.119

Sim, J., and Prabhu, V. (2018). The Life Cycle Assessment of Energy and Carbon Emissions on Wool and Nylon Carpets in the United States. *J. Clean. Prod.* 170, 1231–1243. doi:10.1016/j.jclepro.2017.09.203

Stubbles, J. (2000). Energy Use in the U.S. Steel Industry: a Historical Perspective and Future Opportunities. *U. S.* doi:10.2172/769469

Tan, D., Suvarna, M., Shee Tan, Y., Li, J., and Wang, X. (2021). A Three-step Machine Learning Framework for Energy Profiling, Activity State Prediction and Production Estimation in Smart Process Manufacturing. *Appl. Energy* 291, 116808. doi:10.1016/j.apenergy.2021.116808

The Society of the Plastics Industry, Inc and the U.S Department of Energy (2005). *Improving Energy Efficiency at U.S. Plastics Manufacturing Plants*.

Towards Data Science (2019). Shervin Minaee. Available at: https://towardsdatascience.com/20-popular-machine-learning-metrics-part-1-classification-regression-evaluation-metrics-1ca3e282a2ce.

Walther, J., and Weigold, M. (2021). A Systematic Review on Predicting and Forecasting the Electrical Energy Consumption in the Manufacturing Industry. *Energies* 14 (4), 968. doi:10.3390/en14040968

Willenbacher, M., Scholten, J., and Wohlgemuth, V. (2021). Machine Learning for Optimization of Energy and Plastic Consumption in the Production of Thermoplastic Parts in SME. *Sustainability* 13, 6800. doi:10.3390/su13126800