Application of Neural Networks to Explore Manufacturing Sales Prediction

Po-Hsun Wang [1], Gu-Hong Lin [1] and Yu-Cheng Wang [2],*

[1] Department of Industrial Engineering and Management, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan; Ph.wang@fcs.com.tw (P.-H.W.); ghlin@nkust.edu.tw (G.-H.L.)
[2] Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung 41349, Taiwan
* Correspondence: tony.cobra@msa.hinet.net

check for updates

Abstract: Manufacturing sales prediction is an important measure of national economic development trends. The plastic injection molding machine industry has its own independent R and D energy and mass production technology, with all products sold globally through international brands. However, most previous injection molding machine studies have focused on R and D, production processes, and maintenance, with little consideration of sales activity. With the development and transformation of Industry 4.0 and the impact of the global economy, Taiwan's injection molding machine industry growth rate has gradually flattened or even declined, with company sales and profits falling below expectations. Therefore, this study collected key indicators for Taiwan's export economy from 2008 to 2017 to help understand the impact of economic indicators on injection molding sales. We collected 35 indicators, including net entry rate of employees into manufacturing industries, trend indices, manufacturing industry sales volume indices, and customs export values. We used correlation analysis to select variables affecting plastic injection machine sales and artificial neural networks (ANN) were applied to predict injection molding machine sales at each level. Prediction results were verified against the correlation indicators, and seven key external economic factors were identified to predict accurate changes in company annual sales prediction, which will be helpful for effective resource and risk management.

Keywords: neural networks; prediction; economic variables; plastic injection molding machine

1. Introduction

Production and sales volumes for Taiwan's plastic and rubber machinery in 2017 was $NTD 43.1 B, representing 6.4% growth over 2016. Total export value was $USD 1.163 B, 12.8% increase over 2016, with corresponding import $USD 0.387 B, 8.7% increase over 2016. However, 2018 sales volumes were only $NTD 7.6 B, 8.5% reduction compared with the previous year's $NTD 8.3 B, for 4,528 units. The largest export country was China, then Vietnam ($NTD 1.34 B, 21.4%) and India ($NTD 0.721 B, 11.4%) [1]. Market changes in recent years have been somewhat volatile since plastic and rubber machinery makers are traditional resource-dense industries and manufacturers sought to reduce manufacturing cost by gradually transferred China production bases to Southeast Asia [2]. Industry 4.0 has rapidly increased the demand for high technology precision machinery and automated equipment, which forced transformations onto plastic and rubber manufacturing industries. It was hoped that the large number of plastic and rubber machinery manufacturers in Taiwan (more than 300 in 2018) could be gradually integrated more to help markets become more competitive. Product classification, reconciled functional differences, quality improvements, and market segmentation, helped reduce friction and prices, but a consensus was, and still is, required for manufacturers to avoid using price bargaining as a competitive market strategy [1].

The plastic injection molding machine industry was one of the few Taiwanese industries with recognized international brands. The FCS Group for example, manufactured a wide range of highly customized products to meet global market demand due to geo-correlation. Final manufacturer production and sales strongly impact many supporting industries and groups, including raw material suppliers, mold fixtures, consumer demand research, advertising promotion, etc. The primary motivation for the current study was how to reduce company costs, where we considered the expected major impact aspects of production profit, sales and service costs, and business management, along with various potentially important factors such as accounts receivable, inventory turnover, foreign currency exchange rates, and sales prediction and planning. These all affect revenue, inventory accumulation, capital backlog, financial interest, etc.

Previous studies regarding machinery industries focused mainly on the tool machine industry, with few considering plastic injection machine industries, and these focused almost exclusively on institutional design or injection molding conditions. Although some of these latter studies considered plastic injection molding machine markets, they did not include sales activities. Sales prediction is an important research area, with sales revenues deeply affected by economic changes, and overall economics strongly affecting goods sales. Commodity market supply and demand should be closely considered to improve sales predictions. Recent and rapid development of technology and information has exacerbated unpredictability for many factors, resulting in sometimes severe sales revenue fluctuations. Therefore, to ensure correct predictions, we must carefully consider future resource issue developments; government and financial sector viewpoints on economic policies, important indicator changes for primary and processing industries, national and sector growth rates, with particular focus on economic impact due to emergencies.

Enterprise sales predictions traditionally employ regression analysis or direct artificial expert prediction, based on many analyses and market environment variables [3]. Although many large and complex models have been developed based on sales and sales structures, others remain relatively simple. Recent markets have been largely dominated by price competition. If a given company's prediction model was insufficiently accurate, and predicted higher sales volumes than were actually achieved, inventory could increase greatly, with consequential cost increases. In contrast, when sales were above expectations, production eventually is unable to meet customer demands, which causes loss of business opportunity. Thus, how can plastic injection machine manufacturers understand factors affecting the market economy and verify sales volume through different prediction models? This would help improve sales prediction accuracy, reducing inventory, capital, procurement, and logistics costs while improving production efficiency and order delivery rates with consequential benefits for profitability.

Therefore, this study objectives were to:

1. Establish an accurate model;
2. Understand external economic variables' impacts on prediction models; and
3. Establish a sales prediction framework for practical and feasible plastic injection machine markets to provide assistance for business management decisions.

The remainder of this article is organized as follows. Section 2 reviews the relevant literature, and Section 3 introduces the assessment methodology employed. Section 4 verifies the proposed methodology using an example with practical variables affecting plastic injection machine sales data. Section 5 summarizes and concludes the paper, and discusses some useful directions for future research.

2. Literature Review

Prediction is an important aspect of data exploration based on various statistical methods to identify useful trends or models from historical data, which can then be used to obtain predictions for the next period or cycle [4]. Thus, predictions anticipate future conditions, focusing on practical applications and integration, and generally incorporate mathematical models based on historical data

to provide subjective or intuitive future expectations, or a synthesis of these, i.e., a mathematical model incorporating expert judgment and adjustment.

Every artificial neural network (ANN, NN) model can be classified in terms of its architecture and learning algorithm. The architecture (or topology) describes the neural connections, and the learning (or training) algorithm provides information on how the ANN adapts its weight for every training vector [5]. Artificial neural networks (ANNs) have become widely used for many commercial applications [6], but only a few previous studies have focused on the overview of published research results in this important and popular realm. Most previous studies have focused on financial distress and bankruptcy issues, stock price prediction, and decision support with particular emphasis on classification tasks [7]. ANNs have been widely used for commercial prediction, with traditional multilayer feed-forward networks with gradient descent back-propagation and various hybrid networks developed to improve standard model performances [8].

2.1. Sales Prediction

Sales predictions have been used in research and practical application for many years, with the various prediction models broadly divided into qualitative and quantitative approaches. Quantitative approaches collect historical data and analyze developing trends to provide useful input for management decision making. Prediction methods include case based reasoning, ANNs incorporating artificial intelligence and regression analysis, moving average methods, and exponential smoothing methods based on statistical logic. Analyzing key product characteristics and combining previous sales history data helps to provide better sales predictions from limited data volumes.

Sales prediction is also important for e-commerce, critically impacting wise decision making, and helping to manage workforce, cash flow and resources; and optimize manufacturer supply chains. Thus, sales prediction is generally somewhat challenging because sales are affected by many factors, including promotions, price changes, user preferences, etc. [9], and yet impacts significantly on important company policies and plans, with accurate predictions helping to reduce costs, optimize inventory management, and improve product competitiveness [10].

2.2. Relevant Factors

Relevant factors meant the correlation between variables in data and the correlation between variables and variables in view of whether any linear correlation. Most statistical methods consider correlations between independent variables and variables, whereas factor analysis investigates the relevance for multiple variables. Potential variables affecting variables behind were found even if these variables were not directly measured. These inferred potential variables were called factors [11,12]. Previous factor studies have covered commercial, e.g., [13,14]; sales, e.g., [15–17]; industry, e.g., [18–20]; and manufacturing, e.g., [21–23] applications amongst many others, and has been shown to effectively extract important factors that can then be employed for prediction using different methods.

Companies use identified sales factors for overall economic, industry, and company sales predictions. Overall economic prediction includes inflation, unemployment rates, interest rates, consumer spending and savings, investment, government spending, net export, and gross national output; whereas industry prediction uses the sales factors and other industrial environmental indicators to predict industry sales, and company sales are generally predicted by including company market share, new product listings, and marketing activities [24].

2.3. Artificial Neural Network Prediction

Artificial neural networks were developed in 1950s, and the original neuron-like perceptron model was proposed in 1969 [25]. However, the original theory was somewhat simple and was not taken seriously until Hopfield defined the modern ANN form in 1982. Many diverse ANNs have been subsequently proposed incorporating new architectures and theories, and the enormous increase of

computing power has allowed modern and powerful ANNs to be widely used, with many practical ANN applications across various industries.

Various market sales prediction models have been proposed based on back-propagation ANNs, incorporating improved methods to compensate for known back-propagation inadequacies, particularly for problems with agnostic rules. Although the PCB industry has significant impact on the Taiwanese economy, serious inventory dumping and material shortages remained. Therefore, Chang et al. [26] established an ANN-based prediction model to solve these inventory and material shortage problems. Au et al. [27] developed an optimized ANN prediction system for clothing sales prediction using two years of clothing sales data. They compared model performance with basic fully connected ANN and traditional predictive models, and showed that the proposed algorithm was superior to traditional prediction methods for products with low demand uncertainty and weak seasonal trends. Thus, the method was suitable for fashion retailers to produce short-term retail predictions.

Kong and Martin [28] used back-propagation networks (BNPs) to predict future food sales for large wholesalers in Victoria. They showed that the proposed BPN method provided superior prediction results compared with traditional methods for trending and market analysis, i.e., simple linear regression models. Hence, the BPN model could provide a useful tool for sales prediction. Thiesing et al. [29] used an ANN trained by BPN to predict future time series values, including weekly demand for supermarket items, considering prices, advertising activity, and holiday impacts.

Convenience stores are an integral part of the retail industry, distributing goods to consumers from suppliers. Some convenience stores were out of stock of customer desired products, where others had excess of the same products. However, customer satisfaction was always significantly impacted where desired goods were out of stock, regardless of how good the service was. Thus, controlling ordering and inventory has become a very important issue for store convenience store management. Therefore, Chen et al. [30] developed an ANN based system to control orders and manage the inventory in convenience stores based on business circle and sales prediction operational characteristics. The proposed system improved order and discard rates to better ensure ordering the right items in the correct quantities.

Vhatkar and Dias [31] considered several sales prediction algorithms and developed an inverse transfer-like ANN model to predict oral care product sales and error rates for several different products. Mo et al. [32] proposed an optimized BPN method to improve supermarket daily average rice sales prediction accuracy. Weron [5] proposed an electricity price forecasting. This review article aims to explain the complexity of available solutions, their strengths and weaknesses, and the opportunities and threats that the forecasting tools offer or may be encountered. Cincotti et al. [33] used an ANN to analyze electricity spot-prices of the Italian Power Exchange.

3. Methodology

3.1. Pearson's Correlation Coefficient

Pearson's correlation, Γ, is a commonly employed measure for the linear correlation between two variables [34]:

$$\Gamma = \frac{\sum_i \sum_j ij n_{ij} - (\sum_i in_{j+})(\sum_j jn_{+j}) n_{++}^{-1}}{SQRT\left\{\left[\sum_i i^2 n_{i+} - (\sum_i in_{i+})^2 n_{++}^{-1}\right]\left[\sum_j j^2 n_{+j} - (\sum_j jn_{+j})^2 n_{++}^{-1}\right]\right\}}, \tag{1}$$

where n_{i+} is the number of observations in the i-th row, n_{+i} is the number of observations in the i-th column, n_{ij} is the number of observations in the i-th row, j-th column cell, and n_{++} is the total number of observations. A perfect linear correlation between the variables would have $|\Gamma| = 1$, with positive result meaning the two variables increase together, and negative meaning one decreases as the other increases.

Hypothesis testing is also commonly employed to test if the data supports the null hypothesis, i.e., the two variates are not statistically different. The metric for this test is the probability of generating a type I error:

$$\rho = \frac{\sum\limits_{i=1}^{n} (X_i - \bar{X})(y_i - \bar{y})}{(n-1)S_X S_y},\tag{2}$$

where $\bar{X}$ and $\bar{y}$ are sample means for the first and second variable, respectively; S_X and S_y are standard deviations for the first and second variable, respectively; and n is the column length.

Thus, the null hypothesis for Pearson's correlation is H0: r = 0, and the commonly apply test is to reject the null hypothesis if $\rho \le 0.05$, i.e., at the 95% confidence level.

3.2. Artificial Neural Network Principles

The BPN is one of the most widely used and representative models in current ANN learning models. Werbos [35] proposed the basic concept and Rumelhart et al. [36] defined the underlying theory and algorithm from the propagation or generalized delta learning rule. ANNs simulate biological nerves, using continuous learning and error correction to achieve accurate output. Modern ANNs commonly incorporate large learning databases, parallel processing, nonlinear outputs, and multiple layers.

The basic ANN component is a processing element (PE), where the output from each processing unit is connected to the processing unit of the next layer. Input for the next processing unit originates from corresponding processing units in the upper layer with the sum of the output values. Processing unit output is generally expressed as a function of the sum of the input and linked weighted product:

$$Y_j = f(net) = f(\sum_i W_{ij}X - \theta_j),\tag{3}$$

where Y_j is the output from a neural network processing unit, X_i is the set of input values for the processing unit, *net* are integration functions, θ_i is the applied threshold, W_{ij} is the influence of intensity of the *i*-th processing unit in the previous layer exerted on the *j*-th processing unit in the latter layer between processing units of neural networks were between the data transfer paths between the processing units, and f is the transfer function of neural network processing units obtained the output value via the processing unit transfer functions by using the sum of the input values input from the other processing units and join weight values.

3.3. Back-Propagation Network Algorithm

Back-propagation networks were widely used in ANN architectures [37], providing supervised learning [36]. BPN architecture uses input layers to represent network input variables, where the number of processing units depends on problems by using a linear transfer function, a number of hidden layers, and a single output layer. Propagation in a neural network can be forward or backward. Through these two processes, output errors can be reduced, and the network can be set to learn how to solve a given problem. Forward propagation starts from the input layer and produces computations at each layer until it reaches the final layer. Reverse propagation starts from the output layer and computes weights and errors for each previous layer until it reaches the input layer. Back-propagation compares the gaps between the targeted output and the actual computed output values. The synapse values are re-adjusted to minimize errors [38]. The propagation rule is the basis of the architecture shown in Figure 1.
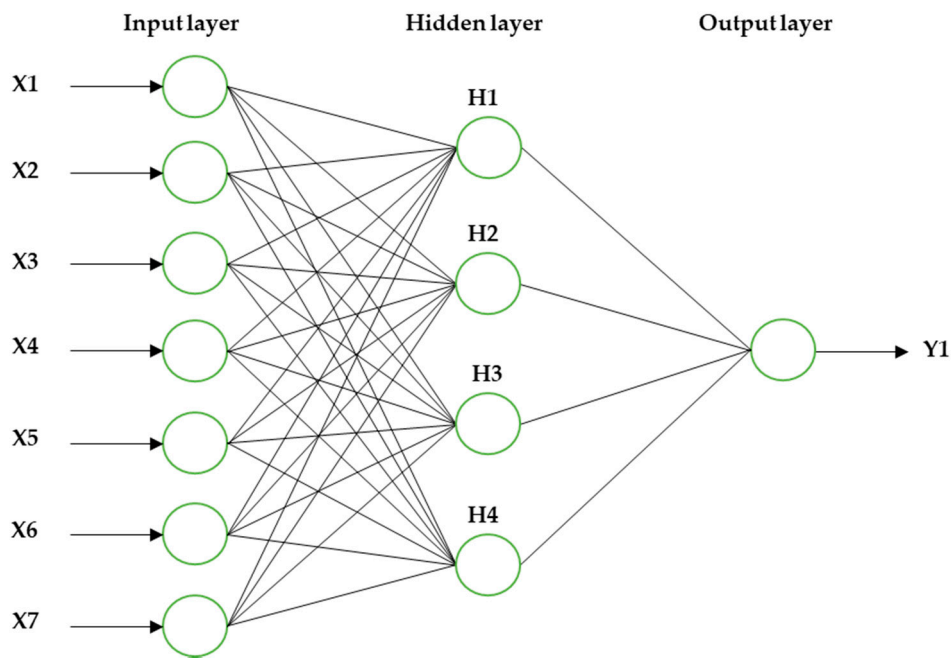
Figure 1. Propagation network architecture.

Hidden layers represent interactions between input units. There is no specified method to decide upon the appropriate or correct number of processing units or hidden layers, hence most studies used accumulated experience to determine these. Generally, the number of hidden layer processing units increases with increasing problem complexity. However, too many processing units causes memorization of the network and poor inductive power; and too few means the network will be unable to obtain the correct mapping relationship between the input and output.

The number of outer layer processing units also depends on problem complexity. If the output is a single real value, the number of output layer processing units = 1, whereas the number of processing units for classification problems depends on the number of classifications. To refrain the processing of a big sample set of data, an ANN tool including a particular training-validation-test procedure for small datasets has been developed some years ago and recently refined in order to obtain realistic and reliable regression laws [39–41].

The basic BPN is to minimize the error function using the gradient steepest descent. For a given training paradigm, the network slightly adjusts link weighting values proportional to the error function's sensitivity to the weighting values. Thus, error functions are proportional to the magnitude of the partial differential values of weightings [12]. Typically, the error (or energy) function representing learning quality has the form:

$$\Delta W = -\eta \frac{\partial E}{\partial W}, \tag{4}$$

where η is the learning rate, which controls the magnitude for every change of weighting value; η is learning rate; and E is the error, hence larger E implies poorer network quality:

$$E - \frac{1}{2}\sum(T_j - Y_j)^2 \tag{5}$$

where T_j is the target output value for the j output unit in the output layer, and T_j is the inference output value of the j output unit in an output layer.

Substituting Equation (5) into Equation (4) and applying the chain rule:

$$\frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial Y_j}\frac{\partial Y_j}{\partial net_j}\frac{\partial net_j}{\partial W_{ij}} = -(T_j - Y_j) \cdot f'(net_j), \tag{6}$$

If we define δ_j as the error of the j output processing unit in an output layer, then:

$$\delta_j = (T_J - Y_J) \cdot f'(net_j),\tag{7}$$

Thus, the correction for W_{ij} is:

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} = \eta \cdot (T_j - Y_j) \cdot f'(net_j) \cdot X_i = \eta \cdot \delta_j \cdot X_i,\tag{8}$$

and the correction to the threshold for the output unit is:

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial \theta_j} = -\eta \cdot \delta_j.\tag{9}$$

Backward correction, i.e., Equations (7)–(9), are repeated layer by layer to obtain the weighting value between each layer and threshold for each neuron, network output values and target values were gradually reduced. The error made the establishment of network models reachable.

Generally, the momentum coefficient or inertia factor, α, is added to Equations (8) and (9) to correct the previous weighting value, reducing oscillation during convergence and making training smoother, i.e.:

$$\Delta W_{ij}(n) = \eta \delta_j X_i + \alpha \Delta W_{ij}(n-1),\tag{10}$$

and:

$$\Delta \theta_{ij}(n) = -\eta \delta_j + \alpha \Delta \theta_j(n-1).\tag{11}$$

There are many data splitting methods reported and used in literatures. These methods can be roughly categorized into three different types [42–44]: (1) Cross-validation, (2) randomly selecting a proportion of samples and retaining these as a validation set and then using the remaining samples for training, and (3) the Kennard–Stone algorithm. In this study, we choose the random method for the ease of application, and are applied on 7/8 portion of training data. The remaining 1/8 portion is the test data.

The ANN is trained by selecting a training set of samples to repeat the learning steps multiple times until errors converge. However, training loops should not be too long to avoid over-training the ANN, ensuring new input samples produce correct outputs.

3.4. Performance Indicators

Predictive metrics compare predicted and actual values to confirm model feasibility. This study used the root mean squared error (RMSE) [16]:

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (T_i - A_i)^2},\tag{12}$$

and mean absolute error (MAE) [17]:

$$MAE = \frac{1}{N} \sum_{j-1}^{N} b_j^2,\tag{13}$$

to measure differences between actual and predicted values. Smaller RMSE and MAE shows predicted values were closer to the actual values [45], and hence the model provided better predictive power.

4. Results and Discussion

The collected dataset comprised 120 entries from 2008 to 2017 for key external economic indicators affecting predicted sales, including net employee entry rate into manufacturing and service industries,

trend indices, manufacturing industry sales volume indices, and customs export indices, as shown in Table 1.

Table 1. The collected dataset.

	X6 Net Entry Rate of Employees in Manufacturing Industries and Service Industries (%)	X8 Coincident Indicators Excluding Trend Index	X11 Manufacturing Industry Sales Volume Index	X12 Customs Export Values ($NTD Billion)	X27 Business Turnover	X29 General Price Index	X31 The Total Value of Trade Imports ($NTD)	Sales Amount ($NTD)
1	0.12	110.72	97.3	723.28	1,160,273	98.99	679,550	67,953,351
2	−0.34	111.24	81.84	569.69	984,024	98.43	522,532	24,804,652
3	0.26	111.39	97.86	757.57	1,101,315	98.23	763,095	82,065,179
4	0.08	111.14	96.78	690.24	1,091,997	99.62	665,578	84,636,977
5	0.12	110.27	96.81	727.09	1,110,420	102.36	661,892	59,689,586
6	0.25	108.59	94.47	746.27	1,090,907	104.24	704,088	57,378,259
7	0.71	105.93	94.55	702.83	1,107,411	105.87	719,137	111,436,433
8	0.07	102.36	92.05	784.49	1,104,864	104.68	785,341	52,641,164
9	−0.27	97.68	88.47	701.1	1,082,218	102.73	672,977	50,619,993
10	−0.37	92.07	86.96	676.33	1,073,243	98.56	580,759	25,432,050
11	−0.7	86.33	70.03	560.84	958,906	92.59	510,760	82,767,645
12	−1.36	81.66	68.5	459.51	944,137	88.82	399,343	36,808,567
13	−1.41	78.87	62.32	413.65	938,376	88.25	301,650	18,262,761
14	−0.85	78.13	68.72	431.96	896,766	89.24	373,591	20,653,199
15	−0.32	78.95	78.09	545.31	969,395	89.08	432,850	40,703,997
16	−0.21	80.97	81.52	504.36	986,153	88.52	436,381	39,451,103
17	−0.13	83.71	80.9	543.17	976,558	88.54	439,880	29,829,289
18	0.25	86.86	88	560.73	1,040,586	89.96	505,506	21,569,185
19	0.42	89.89	91.24	574.13	1,078,566	91.06	515,403	68,449,065
20	0.38	92.58	86.99	629.34	1,082,739	93.15	568,094	45,444,464
				~				
100	0.01	96.48	97.19	720.25	1,129,508	84.26	565,035	50,847,793
101	0.07	97.29	101.78	763.61	1,150,372	85.07	649,604	83,251,179
102	0.08	98.2	101.96	743.41	1,178,038	85.33	627,467	59,205,118
103	0.48	99.13	101.2	776	1,186,582	84.81	657,570	94,401,114
104	0.1	100	105.38	780.01	1,198,745	83.89	653,520	56,923,071
105	−0.06	100.78	98.41	712.91	1,192,628	84.03	574,549	72,765,826
106	0.1	101.44	101.75	840.86	1,221,370	84.72	701,613	54,305,574
107	0.18	101.81	104.94	801.18	1,236,081	85.2	664,895	78,899,175
108	0.12	101.81	105.54	820.74	1,261,626	86.51	665,533	92,900,118
109	0.06	101.39	97.23	757.13	1,224,804	87.14	645,628	58,716,581
110	−0.09	100.72	89.06	702.8	1,037,538	86.63	599,181	93,628,196
111	0.21	100.07	107.95	792.48	1,195,296	85.88	670,891	48,821,852
112	0.11	99.63	95.37	738.98	1,149,976	85.09	654,942	55,235,186
113	0.12	99.54	101.97	769.31	1,187,775	83.98	665,412	43,872,675
114	0.19	99.87	105.53	777.74	1,217,628	83.84	602,507	54,494,222
115	0.62	100.43	102.42	823.94	1,217,436	84.25	660,763	67,840,954
116	0.14	101.1	109.34	840.25	1,264,341	84.87	667,526	77,709,509
117	−0.03	101.66	105.08	869.47	1,263,093	85.64	668,974	71,531,514
118	0.12	101.97	103.39	832.34	1,266,194	86.12	674,863	57,053,528
119	0.21	102.13	105.42	869.41	1,278,629	86.54	691,787	55,927,346
120	0.12	102.1	104.98	884.84	1,275,362	86.72	701,060	94,772,146

4.1. Pearson's Correlation

To improve BPN prediction accuracy the data was pre-processing used Minitab V16 to identify significant correlations, as shown in Table 2.

Table 2. Pearson correlations between collected indicators.

Item	Indicator	Pearson Correlation	Hypothesis Test
X1	Leading indicator composite index	0.161	0.08
X2	Leading indicator excluding trend index	0.273	0.003
X3	Export order trend index	−0.174	0.058
X4	Total currency count M1B ($NTD Billion)	0.033	0.718
X5	Stock index	0.123	0.184
X6	Net entry rate of employees in manufacturing industries and service industries (%)	0.399	**0.000**
X7	Coincident indicator composite index	0.217	0.018
X8	Coincident indicators excluding trend index	0.339	**0.000**
X9	Industrial production index	0.264	**0.004**
X10	Enterprise total electricity consumption (10^9 kWh)	0.207	0.024

Table 2. *Cont.*

Item	Indicator	Pearson Correlation	Hypothesis Test
X11	Manufacturing industry sales volume index	0.381	**0.000**
X12	Customs export values ($NTD Billion)	0.339	**0.000**
X13	Import value of machinery and electrical equipment ($NTD Billion)	0.273	**0.003**
X14	Lagging indicator composite index	0.116	0.21
X15	Leading indicator excluding trend index	0.076	0.411
X16	Unemployment rate (%)	−0.115	0.214
X17	Manufacturing unit output labor cost index	−0.236	0.010
X18	Financial industry overnight interest rate	−0.031	0.741
X19	Loan and investment for all financial institutions ($NTD Billion)	0.009	0.925
X20	Manufacturing inventory (10^3)	0.243	0.008
X21	Consumer confidence index	0.225	0.014
X22	National income	0.011	0.902
X23	Consumer price index	0.042	0.652
X24	Industrial production index	0.306	**0.001**
X25	Manufacturing industry production index	0.297	**0.001**
X26	Plastic products manufactured per worker per month	0.288	**0.002**
X27	Business turnover	0.368	**0.000**
X28	Labor productivity index for plastic products manufacturing	0.151	0.102
X29	General price index	0.323	**0.000**
X30	Export orders (machinery)	0.278	0.002
X31	The total value of trade imports ($NTD)	0.441	0.000
X32	Commercial container unloading capacity (equivalent to 20 containers)	−0.075	0.419
X33	Freight tonnage (metric ton)	0.254	0.005
X34	$USD spot exchange rate—exchange rate between bank and customer—buy ($NTD)	−0.430	0.000
X35	$USD spot exchange rate—exchange rate between bank and customer—sell ($NTD)	−0.430	0.000

Pearson's correlation was considered significant (i.e., $\rho < 0.05$) for 12 items: X6, X8, X9, X11, X12, X13, X24, X25, X26, X27, X29, and X30. To improve prediction accuracy, it was recommended to use the correlation coefficient with moderate correlation and low correlation indicator items, with *p* values selected from the significance below 0.001 as an indicator item. The values with negative correlations were not included. The scenario studied contains 120 entries of actual data released by the government [46].

Thus, we included X6, X8, X11, X12, X27, X29, and X31 (7 items), as shown in Table 3.

Table 3. The Pearson correlation results.

Item	X6	X8	X11	X12	X27	X29	X31
Pearson correlation	0.399	0.339	0.381	0.339	0.368	0.323	0.441
p-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

4.2. Artificial Neural Network Prediction

Study parameters were established from previous ANN studies. Zhang et al. [47] showed that a single hidden layer was sufficient for most nonlinear problems, and Lee and Chen [48] showed there was no standard for setting the number of hidden layer nodes, but recommended to use n, i.e., the column length or number of variates, as the number of nodes in the input layer to test the number of hidden layer nodes. Previous studies considering generally considered setting the learning rate for predictive and estimated problems between 1–10, with subtypes were set between 0.1–1.0 Kuo [49] indicated there was no absolute rule to determine network parameters, and parameter settings providing improved results could only be found through trial and error.

Therefore, we set BPN within Neuro Solutions 7. The network has one hidden layer, with its number neurons as twice the input layer, whereas one neuron in the outer layer is used. The conversion function used is sigmoid, with learning rate, training iterations and error targets as 0.01–1.0, 2000 epochs, and MSE ≤ 10.5, respectively.

The BPN operation steps are as follows:

1. Set the number of input, hidden, and output layers; inertia coefficient, learning rate, learning cycles and other relevant parameters.
2. Randomly determine initial BPN weights.
3. Transfer to the BPN to start calculation.
4. Correct join and bias weights according to errors.
5. If not yet converged, repeat steps 3–5.

After selecting the relevant factors, they were input into the BPN learning using 7/8 of the available data as the training set with the remaining 1/8 served as the test set. The BPN training algorithm generally uses a steep gradient algorithm, but this has slow convergence and can easily fail. Therefore, this study used the Levenberg–Marquardt (LM) algorithm for training, with maximum iterations = 2000. Figure 2 shows the average MSE and standard deviation from three runs.
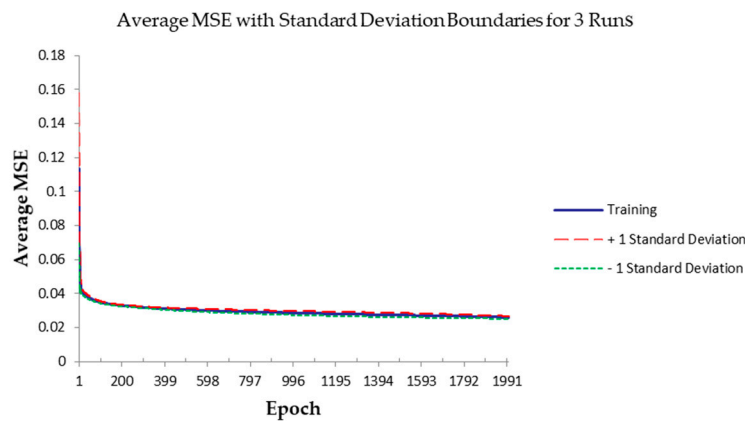
Figure 2. Prediction accuracy (average over three runs).

The training data achieved convergence after 3 runs. All MSEs were below 0.1 with targeted effect reachable shown as Figure 3.
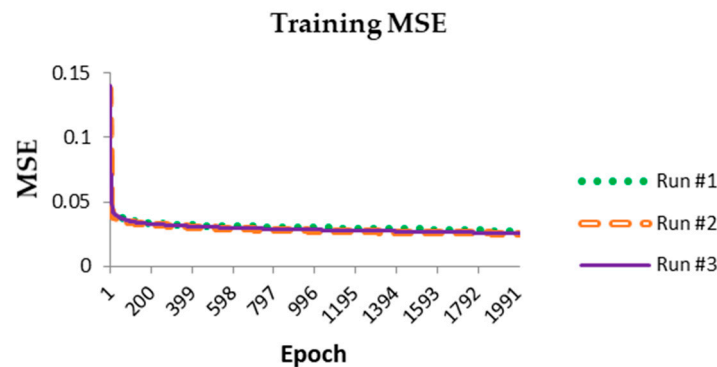
Figure 3. Average MSE for 3 Runs.

The training parameters were entered into the BPN and the test dataset predicted, as shown as Figure 4. The training set had MSE = 0.025882721 ± 0.000802464., as shown Table 4.

Table 4. Training error.

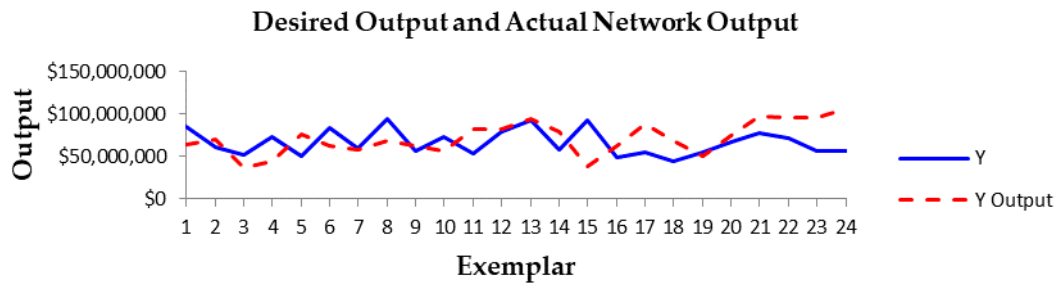| All Runs	Training Minimum	Training Standard Deviation
Average of Minimum MSEs	0.025882721	0.000802464
Average of Final MSEs	0.025882721	0.000802464

Figure 4. Test results from the back-propagation network.

4.3. Performance Assessment

Table 5 summarizes the final BPN performance (for the test dataset). After analysis by relevant factors, the variables with higher homogeneity were clustered together, but the data with significant differences were separated. Therefore, compared with the traditional statistical regression model, the prediction accuracy was very high relatively.

Table 5. Back-propagation artificial neural network prediction performance.

Performance	Y
RMSE	24858562.25
NRMSE	0.169958497
MAE	20464935.83
NMAE	0.139919183
Min Abs Error	1192904.849
Max Abs Error	55346814.96
r	−0.041973132
Score	44.05494133

5. Conclusions

The Taiwanese plastic injection machine industry has not previously analyzed correlations between predicted industrial sales accuracy and key economic indicators. Therefore, this study used the propagation neural prediction method to improve prediction accuracy for plastic injection machines sales in Taiwan for 2018. We first identified influential factors using correlation analysis and integrated a BPN. Overall prediction accuracy achieved was RMSE = 24,858,562.25 for the test set.

Overall outcomes confirmed it was feasible to predict plastic injection machine sales using factor analysis and back-propagation neural networks, and the proposed prediction system clarified the importance of prediction accuracy (MAE and RMSE) could be clearly seen, and was consistent with predicted results.

However, the proposed system has some several limitations that should be addressed:

- The system should be extended to consider changes occurring for the different factors.
- Implementation required for more time than other prediction methods.
- Current accuracy is acceptable, but could be improved.

Industrial sales are driven by many complicated factors. Although the proposed system provided acceptable prediction accuracy, future research should evaluate alternative prediction methods for sales, including classification and fuzzy. Improved accuracy with help further reduce production inventories, improve sales target predictions, and achieve operational goals.

References

1. Taiwan Machinery Industry Association. *2018 Taiwan Plastic Rubber Machinery Industry Status*; TAMI: Taipei, Taiwan, 2018; Available online: http://www.tami.org.tw/market/week3_20180727 (accessed on 5 November 2019).

2. US Commercial Service. *Plastics Materials and Machinery Export Guide. A Reference for U.S. Exporters in the Plastics Industry*; US Commercial Service: Washington, DC, USA, 2018. Available online: https://www.trade.gov/industry/materials/Plastics%20Export%20Guide%202018_final.pdf (accessed on 5 November 2019).

3. Bahrammirzaee, A. A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems. *Neural Comput. Appl.* **2010**, *19*, 1165–1195. [CrossRef]

4. Clemen, R.T. Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.* **1989**, *5*, 559–583. [CrossRef]

5. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* **2014**, *30*, 1030–1081. [CrossRef]

6. Boussabaine, A.H. The use of artificial neural networks in construction management: A review. *Constr. Manag. Econ.* **1996**, *14*, 427–436. [CrossRef]

7. Shin, K.S.; Lee, Y.J. A genetic algorithm application in bankruptcy prediction modeling. *Expert Syst. Appl.* **2002**, *23*, 321–328. [CrossRef]

8. Tkáč, M.; Verner, R. Artificial neural networks in business: Two decades of research. *Appl. Soft Comput.* **2016**, *38*, 788–804. [CrossRef]

9. Zhao, K.; Wang, C. Sales forecast in e-commerce using convolutional neural network. *arXiv* **2017**, arXiv:1708.07946.

10. Croda, R.M.C.; Romero, D.E.G.; Morales, S.-O.C. Sales prediction through neural networks for a small dataset. *Int. J. Interact. Multimed. Artif. Intell.* **2019**, *5*, 35–41.

11. Lawrence, I.; Lin, K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268.

12. Pituch, K.A.; Stevens, J.P. *Applied Multivariate Statistics for the Social Sciences: Analyses with SAS and IBM's SPSS*; Routledge: Abingdon-on-Thames, UK, 2015.

13. Gupta, K.R. *Business Statistics*; Atlantic Publishers & Distributors: Delhi, India, 2017.

14. Krivic, S.J.; Loh, A. Factors relating to brand loyalty of a fitness health club franchise business in Vienna, Austria. *Int. Res. E-J. Bus. Econ.* **2018**, *2*, 56.

15. Schober, P.; Boer, C.; Schwarte, L.A. Correlation coefficients: appropriate use and interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. [CrossRef]

16. Akerejola, W.O.; Okpara, E.U.; Ohikhena, P.; Emenike, P.O. Availability of infrastructure and adoption of point of sales of selected small and medium enterprises (SMEs) in Lagos State, Nigeria. *Int. J. Acad. Res. Bus. Soc. Sci.* **2019**, *9*, 137–150. [CrossRef]

17. Hai, C.; Yan-Yan, C.; Wei, L.; Jun, W.; Shuai, Y.; Ju, G.; Yi, H.; Xiao-jing, D. Early warning analysis of electricity sales based on multi-factor correlation analysis. *E3S Web Conf. EDP Sci.* **2018**, *53*, 02007. [CrossRef]

18. Gao, Y.; Chang, D.; Fang, T.; Luo, T. The correlation between logistics industry and other industries: An evaluation of the empirical evidence from China. *Asian J. Shipp. Logist.* **2018**, *34*, 27–32. [CrossRef]

19. Kong, Y.; Xie, C.; Zheng, S.; Jiang, P.; Guan, M.; Wang, F. Dynamic early warning method for major hazard installation systems in chemical industrial park. *Complexity* **2019**, *2019*, 6250483. [CrossRef]

20. Christou, E. Branding social media in the travel industry. *Procedia Soc. Behav. Sci.* **2015**, *175*, 607–614. [CrossRef]

21. Kung'u, J. Effect of liquidity management practices on profitability of manufacturing industry in Kenya. *IOSR J. Econ. Financ.* **2017**, *8*, 84–89. [CrossRef]

22. Van Wassenhoven, M.; Goyens, M.; Henry, M.; Capieaux, E.; Devos, P. Nuclear magnetic resonance characterization of traditional homeopathically manufactured copper (*Cuprum metallicum*) and plant (*Gelsemium sempervirens*) medicines and controls. *Homeopathy* **2017**, *106*, 223–239. [CrossRef]

23. Anh, T.; Thi, L.; Quang, H.; Thi, T. Factors influencing the effectiveness of internal control in cement manufacturing companies. *Manag. Sci. Lett.* **2020**, *10*, 133–142. [CrossRef]

24. Kotler, P.; Scheff, J. *Standing Room Only: Strategies for Marketing the Performing Arts*; Harvard Business School Press: Brighton, MA, USA, 1997.

25. Minsky, M.; Papert, S. *Perceptron: An Introduction to Computational Geometry*; The MIT Press: Cambridge, MA, USA, 1969; p. 2.

26. Chang, P.-C.; Wang, Y.-W.; Tsai, C.-Y. Evolving neural network for printed circuit board sales forecasting. *Expert Syst. Appl.* **2005**, *29*, 83–92. [CrossRef]

27. Au, K.-F.; Choi, T.-M.; Yu, Y. Fashion retail forecasting by evolutionary neural networks. *Int. J. Prod. Econ.* **2008**, *114*, 615–630. [CrossRef]

28. Kong, J.; Martin, G. A backpropagation neural network for sales forecasting. In Proceedings of the ICNN' 95—International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 2, pp. 1007–1011.

29. Thiesing, F.M.; Vornberger, O. Sales forecasting using neural networks. In Proceedings of the International Conference on Neural Networks (ICNN'97), Houston, TX, USA, 12 June 1997; pp. 2125–2128.

30. Chen, C.-Y.; Lee, W.-I.; Kuo, H.-M.; Chen, C.-W.; Chen, K.-H. The study of a forecasting sales model for fresh food. *Expert Syst. Appl.* **2010**, *37*, 7696–7702. [CrossRef]

31. Vhatkar, S.; Dias, J. Oral-care goods sales forecasting using artificial neural network model. *Procedia Comput. Sci.* **2016**, *79*, 238–243. [CrossRef]

32. Mo, M.; Zhao, L.; Gong, Y.; Wu, Y. Research and application of BP neural network based on genetic algorithm optimization. *Mod. Electron. Tech.* **2018**, *41*, 41–44.

33. Cincotti, S.; Gallo, G.; Ponta, L.; Raberto, M. Modeling and forecasting of electricity spot-prices: Computational intelligence vs classical econometrics. *AI Commun.* **2014**, *27*, 301–314.

34. Ahlgren, P.; Jarneving, B.; Rousseau, R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *J. Am. Soc. Inf. Sci. Technol.* **2003**, *54*, 550–560. [CrossRef]

35. Werbos, P.J. *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*; John Wiley & Sons: Hoboken, NJ, USA, 1994.

36. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Learning Internal Representations by Error Propagation*; Technical Report No. Mar–Sep 1985; University of California San Diego: La Jolla, CA, USA, 1985.

37. Horn, J.F.; Calise, A.J.; Prasad, J. Flight envelope cueing on a tilt-rotor aircraft using neural network limit prediction. *J. Am. Helicopter Soc.* **2001**, *46*, 23–31. [CrossRef]

38. Wang, Y.C. Prediction of engine failure time using principal component analysis, categorical regression tree, and back propagation network. *J. Ambient Intell. Humaniz. Comput.* **2018**, 1–9. [CrossRef]

39. Pasini, A. Artificial neural networks for small dataset analysis. *J. Thorac. Dis.* **2015**, *7*, 953.

40. Pasini, A.; Potestà, S. Short-range visibility forecast by means of neural-network modelling: A case-study. *Il Nuovo Cimento C* **1995**, *18*, 505–516. [CrossRef]

41. Pasini, A.; Pelino, V.; Potestà, S. A neural network model for visibility nowcasting from surface observations: Results and sensitivity to physical input variables. *J. Geophys. Res. Atmos.* **2001**, *106*, 14951–14959. [CrossRef]

42. Xu, Y.; Goodacre, R. On splitting training and validation set: A comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *J. Anal. Test.* **2018**, *2*, 249–262. [CrossRef] [PubMed]

43. Kohavi, R.A. study of cross-validation and bootstrap for accuracy estimation and model selection. *Int. J. Conf. Articial Intell.* **1995**, *14*, 1137–1145.

44. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. [CrossRef]

45. Lewis, E. Control of body segment differentiation in Drosophila by the bithorax gene complex. In *Genes, Development and Cancer: The Life and Work of Edward B. Lewis*; Lipshitz, H.D., Ed.; Springer Science+Business: New York, NY, USA, 1982; pp. 239–253.

46. TAITRA Home Page. Available online: https://www.taitraesource.com/default.asp (accessed on 5 November 2019).

47. Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62. [CrossRef]

48. Lee, T.-S.; Chen, N.-J. Investigating the information content of non-cash-trading index futures using neural networks. *Expert Syst. Appl.* **2002**, *22*, 225–234. [CrossRef]

49. Kuo, J.-T.; Hsieh, M.-H.; Lung, W.-S.; She, N. Using artificial neural network for reservoir eutrophication prediction. *Ecol. Model.* **2007**, *200*, 171–177. [CrossRef]