Technical Report

Improving milk sales quantitative estimation by using POS data

Shin-ichi Shibata (Graduate School of Informatics, Daido University, shibatashinichi0512@gmail.com)
Yuya Kashiwazaki (MCOR Co., Ltd., kashiwazaki.yy@gmail.com)
Toshihiko Shimauchi (Advanced Course, Komatsu College, shimauchi@komatsu-c.ac.jp)
Haruhiko Kimura (Faculty of Production Systems Engineering and Sciences, Komatsu University, haruhiko.kimura@komatsu-u.ac.jp)

Abstract
In this study, experiments on milk sales estimation were conducted in two phases to optimize stock ordering. Sales data of two brands of milk between 2007 and 2008 were obtained through 9 stores in Ishikawa Prefecture. In the first phase, several variable reduction methods were investigated to prevent overfitting of prediction models using methods of the Principal Component Analysis and Decision Tree. In the second phase, the estimation experiments were conducted by using learning data generated through the variable reduction methods. Neural network, k-NN algorithm and RBF network were used for prediction models. The results showed k-NN algorithm with original non-reduced variables and neural network with cumulative contribution rate of 90% yielded higher accuracy in sales estimation.

Key words
sales estimation, PCA, decision tree, neural network, k-NN algorithm

1. Introduction
An information management system called point of sales (POS) system has been widely installed in retail businesses including convenience stores and supermarkets (Namatame, 2007). POS data collected through the POS system include various sales information such as product names, unit price, the number of items sold and date of sales. The information is vital for many businesses since it allows sales forecast (Hokazono et al., 2009) and customer behavior analysis (Kuwabara and Namatame, 2002)

Many stores use the data stored in POS system for stock ordering task. However, the task is not fully automated, hence in many cases, the ordering is conducted based on experience and intuition of the purchasing staff, leading to inaccurate ordering (Matsumura et al., 2016). Some popular items can be out of stock; unpopular items can be overstocked. Losses from missed sales opportunity and inventory disposal can be controlled through optimizing the task.

There are several studies on milk sales quantitative estimation for stock ordering optimization. Milk is perishable items requiring daily order. Terashima and Tsubaki (2006) estimated out-of-stock percentage based on inventory data and sales forecast. They applied a generalized additive model (GAM), which presupposed poison distribution, to sales quantity in estimating the out-of-stock percentage for each item. They pointed out as future directions the necessity of improvement of sales forecast model and refinement of variables selection.

Takahashi and Ishikawa (2000) used neural network for milk sales forecast. They pointed out sales price, day of the week and temperature were statistically significant in affecting milk sales quantity. At the same time, due to limited time-period for the experiment, they were not able to investigate the best way for selecting variables and to compare several time-series models.

Suzuki (2001) also used neural network for milk sales forecast. In this study, multiple correlation coefficients between actual and forecasted sales quantities were used for validating the model accuracy.

This study also used milk POS data. Based on the findings in these related studies which indicated the necessity for variables selection, this study applied a statistical method (principal component analysis) and a machine learning method (decision tree) in order to reduce the number of variables for milk sales estimation for preventing model's overfitting and for improving estimation accuracy. Two experiments were conducted to find the optimum number and composition of the explanatory variables. After the experiments, third experiments were conducted using the results of preceding two experiments to improve sales estimation accuracy by employing following three models:

- Neural network, which usually yield higher estimation accuracy compared to traditional linear regression analysis
- k-nearest neighbor algorithm (k-NN), a memory-based learning method in which the function is approximated locally and all computation is deferred until classification
- radial basis function (RFB) network, which takes into account a position and a distribution of a cluster based on radial basis function.

2. Experiments
2.1 Milk POS data and methods used in sales estimation
The sales of milk are strongly influenced by its price. Milk is perishable item and good for less than a week. The older milk does not sell well because consumer wants to buy fresher

Table 1: POS data used in this study

Stores	9 stores from the same supermarket chain in Ishikawa (store 1 to store 9)
Period	From April 2007 till March 2008
Data item	Date, customer code, department ID, item name, JAN code, unit price
Number of records	Total average 450,000 per month Maximum 750,000 per month (store 6) Minimum 250,000 per month (store 8)

milk. Milk POS data used in the experiments are shown in Table 1. POS data of two competing brands (Genki-ni-Nare Milk, hereafter Genki Milk, and Hokkaido Tokachi Milk, hereafter Hokkaido Milk) were used for their high sales number with constant sales throughout the year.

Using these POS data, this study tried to estimate sales by machine learning with following three experiments:

- Experiment 1:
  Variables reduction using PCA (Principal Component Analysis)
- Experiment 2:
  Variables reduction using Decision Tree
- Experiment 3:
  Sales estimation by comparing variables reduction and three learning models (Neural Network, k-nearest neighbor algorithm and Radial Basis Function Network).

2.2 Software

Visual Mining Studio developed by NTT DATA Mathematical Systems was used for the experiments.

2.3 Variables

Sales estimation for the two brand (Genki Milk and Hokkaido Milk) uses the sales quantity for respective target variables and 31 variables for respective explanatory variables. Tables 2 and 3 respectively shows the variables for Genki Milk and Hokkaido Milk. D-2 signifies two day before and D-1 signifies one day before the day on which sales estimation is conducted.

Table 2: Variables for Genki Milk

Variables	Remarks	Number
Day of the week	Sunday to Saturday (dummy)	7
Customers number	D-2, D-1	2
Temperature (average)	D-2, D-1, D (forecasted on D-1)	3
Temperature (high)	D-2, D-1, D (forecasted on D-1)	3
Temperature (low)	D-2, D-1, D (forecasted on D-1)	3
Precipitation	D-2, D-1, D (forecasted on D-1)	3
Unit price (Genki Milk)	D-2, D-1, D (flier information)	3
Units sold (Genki Milk)	D-2, D-1	2
Unit price (Hokkaido)	D-2, D-1, D	3
Units sold (Hokkaido)	D-2, D-1	2
Total number of variables		31

Table 3: Variables for Hokkaido Milk

Variables	Remarks	Number
Day of the week	Sunday to Saturday (dummy)	7
Customers	D-2, D-1	2
Temperature (average)	D-2, D-1, D (forecasted on D-1)	3
Ttemperature (high)	D-2, D-1, D (forecasted on D-1)	3
Ttemperature (low)	D-2, D-1, D (forecasted on D-1)	3
Precipitation	D-2, D-1, D (forecasted on D-1)	3
Unit price (Genki Milk)	D-2, D-1, D (flier information)	3
Units sold (Genki Milk)	D-2, D-1	2
Unit price (Hokkaido)	D-2, D-1, D	3
Units sold (Hokkaido)	D-2, D-1	2
Total number of variables		31

The day of the week is a dummy variable. For example, Sunday is expressed in 1 for Sunday with other days being expressed in 0. Weather variables contain D variables obtained through weather report forecasted on D-1. Unit price variables also contain D variables obtained through fliers from the store distributed beforehand. These 31 explanatory variables were used for sales estimation of the two milk brands.

2.4 Variables reduction methods

Experiments 1 were conducted with PCA. Experiments 2 were conducted with Decision Tree. Both experiments aimed to reduce the number of explanatory variables for preventing overfitting of the models and improving estimation accuracy.

PCA reduces the number of variables through amalgamation, while decision tree selectively narrows down effective variables. Both aims to find a best combination of explanatory variables. In the experiments, we tried to find a parameter with minimum cumulative errors which are obtained from difference between milk sales estimation and actual sales.

Milk sales estimations were conducted with three models: neural network, k-NN algorithm, and RBF network. Leave-one-out cross validation method was employed for the validation. From 9 stores included in the POS data, 3 stores were selected for the experiments: Store 6 (the highest milk sales), Store 7 (average milk sales) and Store 8 (the lowest milk sales).

2.4.1 Variables reduction with PCA

2.4.1.1 Experiment 1

In this experiment, the parameters with minimum cumulative errors between milk sales estimation and actual sales were investigated through variable reduction method using PCA. In PCA, test data employing principal components with cumulative contribution rate of 80 % (hereafter PCA 80 %) and that of 90 % (hereafter PCA 90 %) were used for the analysis.

2.4.1.2 Results of experiment 1

Tables 4 to 6 show the result of the experiment for Store 6, 7, and 8 respectively. The results show parameters using principal components with cumulative contribution rate of 90 % yielded the lowest cumulative errors. These parameters were used as learning data for estimation models in experiment 3.

2.4.2 Variable reduction with Decision Trees

2.4.2.1 Experiment 2

In this experiment, the parameters with minimum cumulative errors between milk sales estimation and actual sales were investigated through decision trees. Following five decision tree models were constructed for the experiment:

- Decision Tree 1 (DT1) uses all the variables without reduction.
- Decision Tree 2 (DT2) uses top 8 significant variables.
- Decision Tree 3 (DT3) uses 80 % of significant variables.
- Decision Tree 4 (DT4) uses principal components with cumulative contribution rate of 80 %.
- Decision Tree 5 (DT5) uses principal components with cumulative contribution rate of 90 %.

2.4.2.2 Results of experiment 2

Tables 7 to 9 show the result of the experiments for Store 6, 7, and 8 respectively. The result for Store 8 shows DT2 yielded the smallest cumulative errors between estimation and actual sales volume. On the other hand, in Stores 6 and 7, DT 5 yielded the smallest errors.

Based on these results, in experiment 3, principal compo-

Table 4: Results of Store 6 (PCA)

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
PCA 80 %	R (multiple correlation coefficient)		0.96	0.96	0.92	0.93	0.91	0.88
	Errors	Mean	22	24	30	21	21	26
		Max	119	101	256	116	149	139
		SD (Standard Deviation)	20	19	30	19	22	25
	Cumulative errors		7,440	7,890	9,896	6,510	6,681	8,261
PCA 90 %	R (multiple correlation coefficient)		0.96	0.96	0.92	0.94	0.92	0.87
	Errors	Mean	22	24	32	19	21	26
		Max	130	99	217	89	141	155
		SD(Standard Deviation)	20	20	29	17	21	26
	Cumulative errors		7,283	7,983	10,735	5,872	6,562	8,289

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

Table 5: Results of Store 7 (PCA)

			Genki			Hokkaido		
			k-NN	RBFN	NN	k-NN	RBFN	NN
PCA 80 %	R (multiple correlation coefficient)		0.92	0.95	0.89	0.83	0.83	0.81
	Errors	Mean	20	17	25	23	23	25
		Max	218	134	195	150	149	150
		SD (Standard Deviation)	27	20	28	21	22	22
	Cumulative errors		7,016	6,107	8,916	7,602	7,445	8,149
PCA 90 %	R (multiple correlation coefficient)		0.93	0.95	0.85	0.77	0.83	0.78
	Errors	Mean	18	17	32	25	23	27
		Max	187	98	177	185	165	146
		SD (Standard Deviation)	24	19	33	27	22	23
	Cumulative errors		6,234	6,103	11,612	8,116	7,445	8,723

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

Table 6: Results of Store 8 (PCA)

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
PCA 80 %	R (multiple correlation coefficient)		0.91	0.89	0.83	0.98	0.94	0.88
	Errors	Mean	10	12	14	4	8	11
		Max	69	55	56	26	51	50
		SD (Standard Deviation)	9	10	12	4	7	9
	Cumulative errors		3,297	3,832	4,678	1,260	2,378	3,531
PCA 90 %	R (multiple correlation coefficient)		0.90	0.90	0.84	0.98	0.94	0.89
	Errors	Mean	10	11	15	4	8	11
		Max	88	53	52	55	53	47
		SD (Standard Deviation)	11	9	12	5	7	9
	Cumulative errors		3,301	3,709	4,744	1,259	2,490	3,317

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

nents with cumulative contribution rate of 90 % were used for learning data in building decision tree models.

The number of explanatory variables after reduction experiments are shown in Table10.

2.5 Sales estimation accuracy

2.5.1 Experiment 3

For explanatory variables, three patterns (original data, reduction by PCA, and reduction by DT) were used. For the sales estimation, three methods (neural network, k-NN algorithm, and RBF network) were used. The combination of these patterns and methods produced 9 models. For each model, correlation, errors, cumulative errors and percentage of cumulative errors were calculated.

2.5.2 Result of experiment 3

Tables 11 to 13 summarize the results of the experiment for Store 6, 7, and 8 respectively.

For the brand, Genki Milk showed lowest cumulative errors by using original information without reducing variables and predicting with k-NN algorithm. On the other hand, Hokkaido Milk showed more complicated results according to the sales volume. Specifically, in Store 6 with the largest sales, the combination of variable reduction by decision tree 5 with estimation by neural network yielded the lowest cumulative errors. In Store 7 with average sales, the combination of variable reduction by PCA 90 % and estimation by k-NN algorithm showed the lowest cumulative errors.

In Store 8 with the lowest sales, the combination of original information without variable reduction with neural network

Table 7: The result of decision tree method for Store 6

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
DT1	R (multiple correlation coefficient)		0.96	0.96	0.94	0.83	0.83	0.52
	Errors	Mean	23	23	28	30	28	30
		Max	147	130	129	203	194	203
		SD (Standard Deviation)	22	20	25	32	32	32
	Cumulative errors		7,771	7,574	9,309	9,308	8,714	9,308
DT2	R (multiple correlation coefficient)		0.73	0.73	0.62	0.68	0.70	0.58
	Errors	Mean	36	37	47	29	28	34
		Max	243	220	217	177	160	145
		SD (Standard Deviation)	46	43	43	31	27	29
	Cumulative errors		12,092	12,542	15,741	9,562	9,473	11,470
DT3	R (multiple correlation coefficient)		0.94	0.93	0.92	0.94	0.94	0.86
	Errors	Mean	28	28	32	30	30	50
		Max	274	269	177	280	260	227
		SD (Standard Deviation)	27	28	28	32	34	45
	Cumulative errors		9,299	9,411	10,736	9,528	9,292	15,736
DT4	R (multiple correlation coefficient)		0.88	0.93	0.89	0.85	0.87	0.78
	Errors	Mean	37	28	35	28	25	33
		Max	188	237	227	186	215	214
		SD (Standard Deviation)	38	30	37	27	27	34
	Cumulative errors		12,414	9,292	11,655	8,864	7,828	10,404
DT5	R (multiple correlation coefficient)		0.96	0.95	0.92	0.98	0.95	0.90
	Errors	Mean	22	24	30	13	17	23
		Max	108	201	210	90	109	172
		SD (Standard Deviation)	20	23	32	11	18	23
	Cumulative errors		7,534	7,993	10,013	3,935	5,438	7,145

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

estimation showed the lowest cumulative errors

3. Results

The experiments on variable reductions showed following results:

- In PCA reduction method, the cumulative errors were lower in the experiment using cumulative contribution rate of 90 % than those of 80 %.
- In decision tree method, Decision Tree 5 (DT5) using principal components with cumulative contribution rate of 90 % showed lower cumulative errors than the other models.
- In both experiments, the two brands did not show any difference in the cumulative errors among different combinations of variable reduction methods

For the brand, Genki Milk showed lowest cumulative errors by using original information without reducing variables and predicting with k-NN algorithm. On the other hand, Hokkaido Milk showed more complicated results according to the sales volume. Overall accuracy was highest in the estimation using PCA 80 % combined with neural network.

For Genki Milk, the estimation accuracy decreased in proportion to the sales. However, for Hokkaido Milk, the estimation for Store 8 with the least sales quantity was the most accurate.

For the cumulative errors, Genki Milk showed a decrease in the errors according to the sales volume. However, Hokkaido Milk showed otherwise. For both brands, cumulative errors were the smallest in Store 8 with the lowest sales quantity.

4. Conclusions and Future directions

In this study, milk sales estimations were conducted. In the first phase, two experiments were run using original data and

Table 8: The result of decision tree method for Store 7

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
DT1	R (multiple correlation coefficient)		0.65	0.75	0.61	0.56	0.66	0.31
	Errors	Mean	44	36	47	36	31	44
		Max	245	206	212	199	154	150
		SD (Standard Deviation)	52	41	46	37	28	30
	Cumulative errors		15,241	12,584	16,283	11,843	10,241	14,321
DT2	R (multiple correlation coefficient)		0.73	0.73	0.62	0.84	0.87	0.77
	Errors	Mean	36	37	47	30	27	41
		Max	236	220	217	276	262	226
		SD (Standard Deviation)	47	43	44	40	36	42
	Cumulative errors		12,561	12,923	16,478	9,862	8,818	13,378
DT3	R (multiple correlation coefficient)		0.77	0.81	0.72	0.67	0.72	0.58
	Errors	Mean	32	31	40	30	28	35
		Max	242	206	203	180	161	142
		SD (Standard Deviation)	43	36	41	31	26	29
	Cumulative errors		11,311	10,717	14,137	9,765	9,346	11,352
DT4	R (multiple correlation coefficient)		0.77	0.75	0.70	0.62	0.66	0.62
	Errors	Mean	35	35	42	33	31	36
		Max	218	217	218	183	168	145
		SD (Standard Deviation)	40	42	43	29	28	28
	Cumulative errors		12,319	12,375	14,614	10,875	10,301	11,698
DT5	R (multiple correlation coefficient)		0.84	0.77	0.64	0.75	0.68	0.68
	Errors	Mean	28	34	46	27	31	31
		Max	249	202	213	194	153	153
		SD (Standard Deviation)	34	40	45	25	27	27
	Cumulative errors		9,887	11,960	16,069	8,926	10,062	10,062

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

data with variable reductions to identify better combination of learning data. Validations were done by comparing cumulative errors of each model.

In the experiments 1 and 2, variable reduction methods were conducted. PCA and decision tree models were used to identify the variable reduction methods for improving estimation accuracy. The results showed, in both models, better sales estimation accuracy can be obtained by variables with 90% contribution rate as learning data.

In experiment 3, milk sales estimations were conducted using original data and data applied variable reductions in experiments 1 and 2. Three models used were neural network, k-NN algorithm and RBF network. The results showed k-NN algorithm with original non-reduced variables and neural network with cumulative contribution rate of 90% yielded higher accuracy in sales estimation. This suggests the better method of preprocessing of data depend on the estimation model into which the data are used.

In every experiment, multiple correlation coefficient is 0.8 or higher and cumulative error ratios are approximately between 10 % to 40 %. These results show effectiveness of our proposed method in accurately estimating milk sales.

In this study, the experiments were conducted to the stores with unique sales features such as higher or lower than the average. In the future, the experiments will be conducted to other stores to evaluate the effectiveness of variable reduction and estimation methods proposed in this paper.

Another limitation is that the POS data used in this study contained no ID number, which precludes combining purchasing behavior and consumer data. As a result, sales quantitative estimation could not incorporate purchasing history and customer information. To build a more accurate estimation model for milk sales, consumer data such as sex, age, and family composition will be necessary.

Table 9: The result of decision tree method for Store 8

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
DT1	R (multiple correlation coefficient)		0.81	0.83	0.62	0.98	0.98	0.95
	Errors	Mean	14	14	20	2	7	10
		Max	105	88	98	77	69	62
		SD (Standard Deviation)	15	13	18	5	7	9
	Cumulative errors		4,620	4,541	6,618	730	2,086	2,975
DT2	R (multiple correlation coefficient)		0.85	0.82	0.83	0.98	0.96	0.94
	Errors	Mean	13	14	15	2	6	8
		Max	87	88	88	78	69	61
		SD (Standard Deviation)	13	13	13	5	6	7
	Cumulative errors		4,057	4,673	4,987	697	1,794	2,437
DT3	R (multiple correlation coefficient)		0.83	0.83	0.71	0.98	0.96	0.94
	Errors	Mean	14	14	20	2	6	8
		Max	84	88	74	78	69	54
		SD (Standard Deviation)	14	13	15	5	6	7
	Cumulative errors		4,445	4,509	6,422	716	1,966	2,484
DT4	R (multiple correlation coefficient)		0.78	0.73	0.64	0.96	0.94	0.90
	Errors	Mean	16	17	19	6	8	10
		Max	91	100	111	55	50	60
		SD (Standard Deviation)	15	16	18	6	8	9
	Cumulative errors		5,053	5,498	6,198	1,738	2,466	3,101
DT5	R (multiple correlation coefficient)		0.80	0.76	0.67	0.98	0.87	0.54
	Errors	Mean	15	16	19	4	8	11
		Max	88	93	91	50	56	82
		SD (Standard Deviation)	14	15	17	4	9	11
	Cumulative errors		4,793	5,291	6,123	1,319	2,597	3,302

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

Table 10: Number of variables after different reduction methods

		Store 6		Store 7		Store 8	
		Genki	Hokkaido	Genki	Hokkaido	Genki	Hokkaido
No reduction		31	31	31	31	31	31
PCA	80 %	8th	8th	10th	10th	8th	8th
	90 %	11th	11th	12th	14th	12th	11th
Decision Tree	1	14	15	17	13	14	25
	2	8	8	8	8	8	8
	3	11	12	14	10	13	10
	4	5th	4th	7th	6th	6th	6th
	5	6th	6th	9th	8th	8th	7th

Note: Ordinal numbers in the table signifies Nth principal component.

Table 11: The result for Store 6

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
No reduction	R (multiple correlation coefficient)		0.94	0.97	0.93	0.87	0.92	0.88
	Errors	Mean	27	22	30	28	21	27
		Max	168	88	192	151	136	138
		SD (Standard Deviation)	25	18	29	26	22	25
	Cumulative errors		9,114	7,236	10,165	8,662	6,599	8,378
	Cumulative error ratio		18.0 %	14.3 %	20.1 %	24.3 %	18.5 %	23.5 %
PCA 90 %	R (multiple correlation coefficient)		0.96	0.96	0.92	0.94	0.92	0.87
	Errors	Mean	22	24	32	19	21	26
		Max	130	99	217	89	141	155
		SD (Standard Deviation)	20	20	29	17	21	26
	Cumulative errors		7,283	7,983	10,735	5,872	6,562	8,289
	Cumulative error ratio		14.4 %	15.8 %	21.3 %	16.5 %	18.4 %	23.3 %
DT and PCA 90 %	R (multiple correlation coefficient)		0.96	0.95	0.92	0.98	0.95	0.90
	Errors	Mean	22	24	30	13	17	23
		Max	108	201	210	90	109	172
		SD (Standard Deviation)	20	23	32	11	18	23
	Cumulative errors		7,534	7,993	10,013	3,935	5,438	7,145
	Cumulative error ratio		14.9 %	15.8 %	19.8 %	11.1 %	15.3 %	20.1 %

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

Table 12: The result for Store 7

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
No reduction	R (multiple correlation coefficient)		0.91	0.95	0.91	0.73	0.83	0.80
	Errors	Mean	22	17	23	29	23	25
		Max	161	117	134	191	142	153
		SD (Standard Deviation)	27	19	26	28	22	22
	Cumulative errors		7,824	6,023	8,084	9,613	7,527	8,208
	Cumulative error ratio		33.9 %	26.1 %	35.0 %	41.5 %	32.5 %	38.2 %
PCA 90 %	R (multiple correlation coefficient)		0.92	0.95	0.89	0.83	0.83	0.81
	Errors	Mean	20	17	25	23	23	25
		Max	218	134	195	150	149	150
		SD (Standard Deviation)	27	20	28	21	22	22
	Cumulative errors		7,016	6,107	8,916	7,602	7,445	8,149
	Cumulative error ratio		30.4 %	26.5 %	38.6 %	32.8 %	32.1 %	35.2 %
DT and PCA 90 %	R (multiple correlation coefficient)		0.84	0.77	0.64	0.75	0.68	0.68
	Errors	Mean	28	34	46	27	31	31
		Max	249	202	213	194	153	153
		SD (Standard Deviation)	34	40	45	25	27	27
	Cumulative errors		9,887	11,960	16,069	8,926	10,062	10,062
	Cumulative error ratio		42.8 %	51.8 %	69.6 %	38.5 %	43.4 %	43.4 %

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

Table 13: The result for Store 8

			Genki			Hokkaido		
			NN	k-NN	RBFN	NN	k-NN	RBFN
No reduction	R (multiple correlation coefficient)		0.85	0.90	0.84	0.98	0.94	0.84
	Errors	Mean	13	11	14	2	7	12
		Max	88	59	58	75	80	52
		SD (Standard Deviation)	13	10	12	5	7	10
	Cumulative errors		4,357	3,632	4,669	557	2,327	3,654
	Cumulative error ratio		56.9 %	47.5 %	61.0 %	3.9 %	16.2 %	25.5 %
PCA 90 %	R (multiple correlation coefficient)		0.90	0.90	0.84	0.98	0.94	0.89
	Errors	Mean	10	11	15	4	8	11
		Max	88	53	52	55	53	47
		SD (Standard Deviation)	11	9	12	5	7	9
	Cumulative errors		3,301	3,709	4,744	1,259	2,490	3,317
	Cumulative error ratio		43.1 %	48.5 %	62.0 %	8.8 %	17.4 %	23.1 %
DT and PCA 90 %	R (multiple correlation coefficient)		0.80	0.76	0.67	0.98	0.87	0.54
	Errors	Mean	15	16	19	4	8	11
		Max	88	93	91	50	56	82
		SD (Standard Deviation)	14	15	17	4	9	11
	Cumulative errors		4,793	5,291	6,123	1,319	2,597	3,302
	Cumulative error ratio		62.6 %	69.1 %	80.0 %	9.2%	18.1 %	23.0 %

Notes: PCA 80 % refers to test data employing principal components with cumulative contribution rate of 80 %. PCA 90 % refers to test data employing principal components with cumulative contribution rate of 90 %. NN refers to Neural Network, K-NN refers to k-nearest neighbor algorithm, and RBF refers to Radial Basis Function network.

References

Hokazono, S., Kano, H., and Maeda, K. (2009). A firm-level demand analysis of milk by using scanner data with bargain sale information. *Journal of Food System Research*, Vol. 16, No. 3, 15-23.

Kuwabara, S. and Namatame, T. (2002). Purchasing behavior analysis considering external factors. *Abstracts of the Fall National Conference of Operations Research Society of Japan*, 162-163.

Matsumura, N., Izumi, K., and Yamada, K. (2016). A marketing simulation of a retail store with the consumer reactions to out-of-shelf based on a POS data. *The Journal of The Japanese Society for Artificial Intelligence*, Vol. 31, No. 2, F-F13_1-8.

Namatame, T. (2007). Progress of data analysis in marketing. *Proceedings of Second Conference of Transdisciplinary Federation of Science and Technology*, 53-53.

Suzuki, T. (2001). Prediction of milk sales by neural network. *Statistical Quality Control*, Vol. 52, No. 3, 249-254.

Takahashi, K. and Ishikawa, Y. (2000). A neural network model for milk sales forecasting. *Proceedings of 19th Japan SAS User Annual Conference*, 235-242.

Terashima, M. and Tsubaki, H. (2006). Estimation of stockout-rates using inventory data and forecasted sales. *Journal of Japan Logistics Society*, Vol. 2006, No. 14, 93-100.

(Received: December 19, 2019; Accepted: January 20, 2020)