SVM Based Maximum Power Consumption Excess Forecast Alert for Large-Scale Power Consumers

Seigo Haruta[1][a], Ken-ichi Tokoro[2][b] and Takashi Onoda[1][c]
[1] *Aoyama Gakuin University School of Science and Engineering, Kanagawa, Japan*
[2] *Central Research Institute of Electric Power Industry, Kanagawa, Japan*

Keywords: Support Vector Machine (SVM), Machine Learning, Discrimination Problem, Excess Forecast Alert, Large-Scale Power Consumers, Maximum Power Consumption, Imbalanced Data.

Abstract: Large-scale power consumers, such as buildings and factories, make high-voltage power contracts with the Japanese electric power companies. The basic fee for high-voltage power contracts is based on the maximum power consumption in the past year. If the power consumption in the present month does not exceed the maximum power consumption in the past year, large-scale power consumers can suppress the basic fee. So, large-scale power consumers need the alert to prevent the maximum power consumption in the present month from exceeding the maximum power consumption in the past year. In this study, excess forecasting was performed considering the characteristics of power consumption in each industry. In addition, we proposed SVM improvements for imbalanced data. We applied this method to power consumption data, which is imbalanced data, to perform excess forecast. As a result, we have improved the accuracy of the excess forecast and contributed to effective alerts to many large-scale power consumers.

1 INTRODUCTION

Consumers with large-scale electricity demand, such as buildings, factories, and hospitals, are called large-scale power consumers. In Japan, electric power companies sold 837.4 billion kWh of electricity in 2021. Of this, 552 billion kWh, or 62.3%, is used by large-scale power consumers. Therefore, large-scale power consumers are important customers for electric power companies. In addition, electricity liberalization is progressing in Japan. Electric power companies must continue to provide services to large-scale power consumers who use a lot of electricity. Large-scale power consumers make the high-voltage power contract with the Japanese electric power companies. The power contract for general households has a low basic fee and a high power usage fee that is calculated in proportion to the amount of power used. The high-voltage power contract has a low power usage fee and a high basic fee. The basic fee accounts for a large proportion of the electricity charges of large-scale power consumers. The basic fee for high-voltage power contracts is based on the maximum power consumption in the past year.

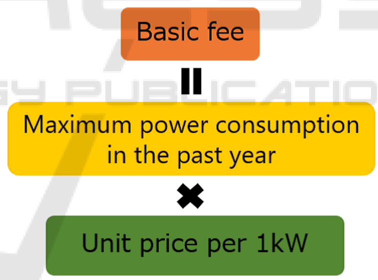

Figure 1: How the basic fee is determined for large-scale power consumers.

Figure 1 shows how the basic fee for high voltage power contracts is determined. Large-scale power consumers want to reduce their electricity charges. If the power consumption in the present month does not exceed the maximum power consumption in the past year, large-scale power consumers can suppress the basic fee(Tokoro et al., 2019). It is most important for large-scale power consumers to prevent exceeding the maximum power consumption for the past year. Figure 2 shows the maximum power consumption that determines the basic fee for August 2022. As shown in figure 3, if the maximum power consumption is exceeded in August 2022, the maximum power consumption that determines the basic fee in

[a] https://orcid.org/0000-0002-1311-7246
[b] https://orcid.org/0000-0001-5074-8250
[c] https://orcid.org/0000-0002-5432-0646

Haruta, S., Tokoro, K. and Onoda, T.
SVM Based Maximum Power Consumption Excess Forecast Alert for Large-Scale Power Consumers.
DOI: 10.5220/0011666100003411
In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2023)*, pages 806-813
ISBN: 978-989-758-626-2; ISSN: 2184-4313

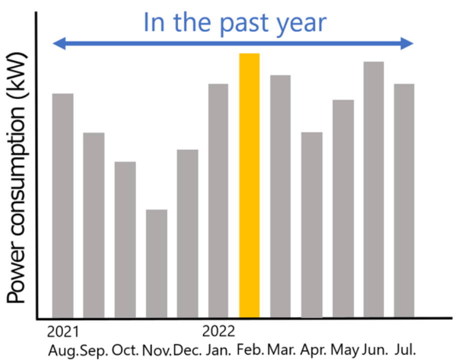

Figure 2: Maximum power consumption to determine basic fee in August 2022.

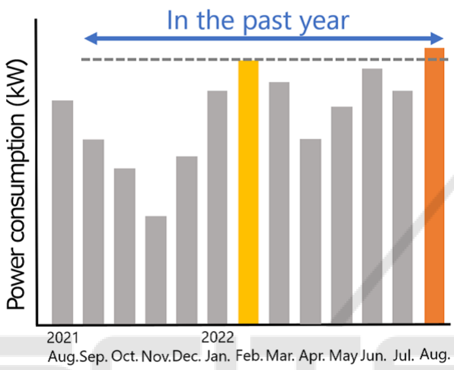

Figure 3: Maximum power consumption to determine basic fee in September 2022.

September 2022 changes. In order to solve such problems, power demand forecasting has been studied in various ways from general households to large-scale power consumers(Suganthi and Samuel, 2012)(Haghi and Toole, 2013)(Motamedi et al., 2012).

2 RELATED RESEARCH

This chapter describes related research on forecast alerts for exceeding maximum power consumption based on support vector machine (SVM)(Cortes and Vapnik, 1995)(Burges et al., 1999)(Vapnik, 1999b)(Vapnik, 1999a) to reduce basic fee for large-scale power consumers(Tokoro et al., 2019).

2.1 Excess Forecast Alert

As one of the various forms of power demand forecasting being studied, an electric power company proposed an excess forecast alert based on a discrimination problem as a service to be provided to large-scale power consumers. Figures 4 and 5 show the flow of excess forecast alerts. At forecast execution time t,

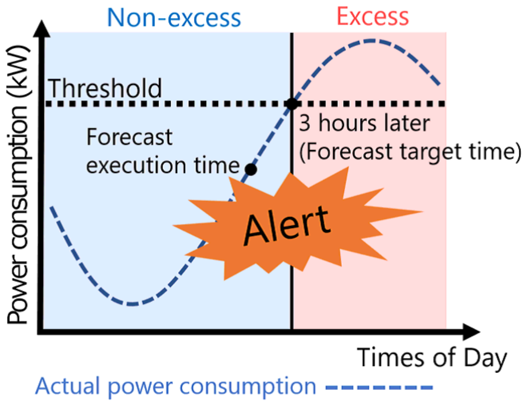

Figure 4: How alerts are sent.

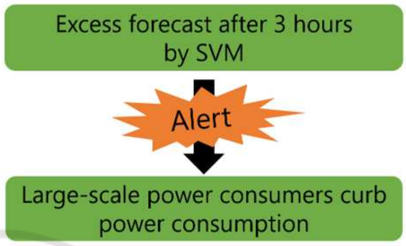

Figure 5: Large-scale power consumers curb power consumption.

SVM is used to forecast whether the power consumption at forecast target time $t+3$ after 3 hours will exceed the threshold. The threshold is 90% of the maximum power consumption in the past year. If the threshold is forecast to be exceeded, an excess forecast alert is sent. Large-scale power consumers who receive alerts can reduce their power consumption so that they do not exceed their maximum power consumption.

2.2 Results and Issues of Related Research

Table 1 shows the prediction accuracy (recall, precision, F-measure)(Ahmad et al., 2018)(Powers, 2020)(Goutte and Gaussier, 2005)of an excess forecast alert in the power consumption data of 5,727 large-scale power consumers in Japan. The issue of related research is that the discriminant function for an excess forecast alert is created under the same conditions for all industries, even though the amount of electricity used differs depending on the industry of large-scale power consumers. Large-scale power consumers are diverse as shown in figure 6. Because the discriminant function is created without considering the characteristics of the industry, the forecast accuracy of an excess forecast alert may be low.

Table 1: Forecast accuracy for an excess forecast alert.

Recall	Precision	F-measure
62.17%	78.74%	68.28%

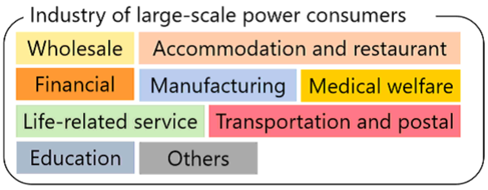

Figure 6: Industry of large-scale power consumers.

2.3 The Purpose of This Research

We perform excess forecasts based on the characteristics of each industry to improve forecast accuracy. We also improve the forecast accuracy by improving the SVM.

3 PROPOSED METHOD

Power consumption data for 5,727 large-scale power consumers in Japan can be divided into nine industries. To improve the accuracy of excess forecast alerts, we analyzed the characteristics of each industry. We also improved the SVM for imbalanced data to improve the accuracy of excess forecast alerts.

3.1 Characteristics of Power Consumption in Each Industry

There is a difference in the power consumption between weekdays and holidays depending on industry of the large-scale consumer. The figure 7 shows a graph of changes in power consumption by wholesalers and retailers in the Kanto region of Japan (Monday, April 1, 2013 to Sunday, April 7, 2013). The figure 8 shows a graph of changes in power consumption by the financial and insurance industries in the Kanto region of Japan (Monday, April 1, 2013 to Sunday, April 7, 2013). The power consumption of wholesalers and retailers is constant regardless of the day of the week.

Also, depending on the type of business of large-scale power consumers, there are cases where there are many excesses only in the summer, and there are cases where there are many excesses in both summer and winter. Figure 9 shows a case of high excesses only in the summer. Figure 9 shows the number of monthly excesses for all wholesaler and retailer customers in the Kanto region of Japan. Figure 10 shows a case of high excesses in both summer and winter.

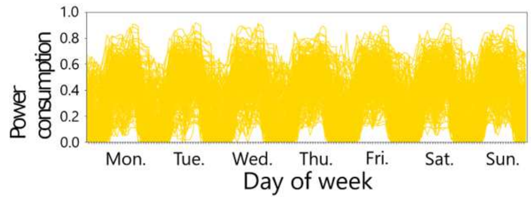

Figure 7: Changes in power consumption by wholesalers and retailers in the Kanto region of Japan (Monday, April 1, 2013 to Sunday, April 7, 2013).

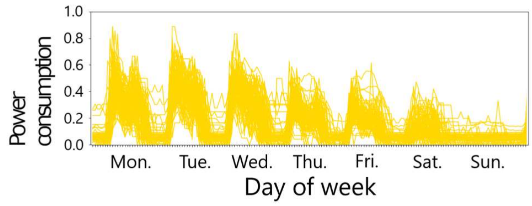

Figure 8: Changes in power consumption by the financial and insurance industries in the Kanto region of Japan (Monday, April 1, 2013 to Sunday, April 7, 2013).

Figure 10 shows the number of monthly excesses for all financial and insurance industry customers in the Kanto region of Japan.

In the related research, the features used for the discriminant function did not take into consideration the industries of large-scale power consumers. In this research, we select features that affect the excess of the maximum power consumption for each industry of large-scale power consumers. The method of selecting features that affect exceeding the maximum power consumption is shown below.

1. Normalize all features from minimum 0 to maximum 1
2. Find the median value of each feature when exceeding and when not exceeding
3. Select features with a median difference of 0.2 or more

Figure 11 is a boxplot of the feature quantity "Power consumption of forecast execution time" at

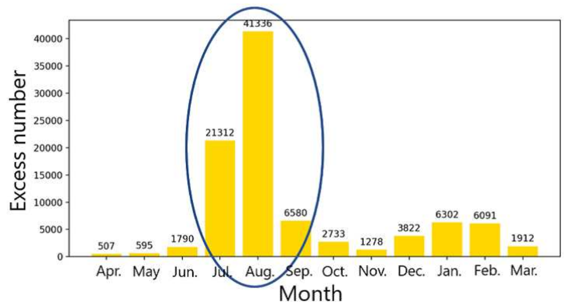

Figure 9: The number of monthly excesses for all wholesaler and retailer customers in the Kanto region of Japan.

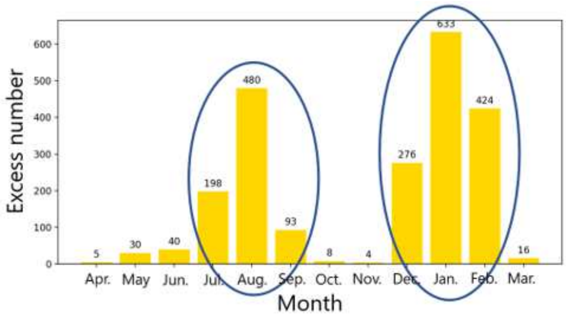

Figure 10: The number of monthly excesses for all financial and insurance industry customers in the Kanto region of Japan.

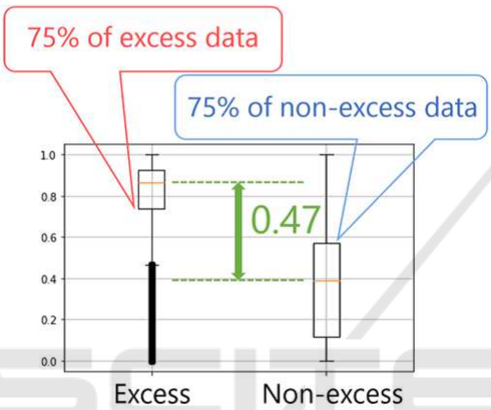

Figure 11: The difference between the median values of the feature for excess and non-exceeding.

the time of excess and non-exceeding in the Kanto region wholesale and retail industry. As shown in figure 11, when the difference between the median values of the feature for excess and non-exceeding is 0.2 or more, 75% of the excess data and 75% of the non-exceeding data have different values.

3.2 SVM Improvements for Imbalanced Data

In related research, the ratio of when consumers power consumption exceeds the maximum power consumption of the past year and when it does not is imbalanced. In a discrimination problem using such imbalanced data, there are cases where the model is tilted by the majority class and the minority class is neglected. Therefore, an excess forecast accuracy of related research was low. Other related research uses methods such as oversampling and undersampling as solutions to the problem(Mohammed et al., 2020)(He and Garcia, 2009)(Sun et al., 2009)(Krawczyk, 2016).

In this research, we propose a method to solve the discrimination problem of imbalanced data by moving the optimal hyperplane of SVM by the ratio of the number of training failures in each class. SVM calculates discriminant function $f(x)$ from training data. Let $f(x) = 0$ be the optimal hyperplane for discrimination. However, in the discrimination problem of imbalanced data as shown in figure 12, there are cases where learning is not successful. We will improve the recall by moving the optimal hyperplane as shown in figure 13. Let the optimal hyperplane after being moved be the modified optimal hyperplane $g(x) = 0$. Equation (1) shows the modified optimal hyperplane $g(x) = 0$ using FP, FN and $f(x) = 0$ in the figure12. By moving the optimal hyperplane, misidentification of class 2 increases, but misidentification of class 1 decreases. Equation (2) shows the recall calculation. Equation (3) shows the precision calculation.

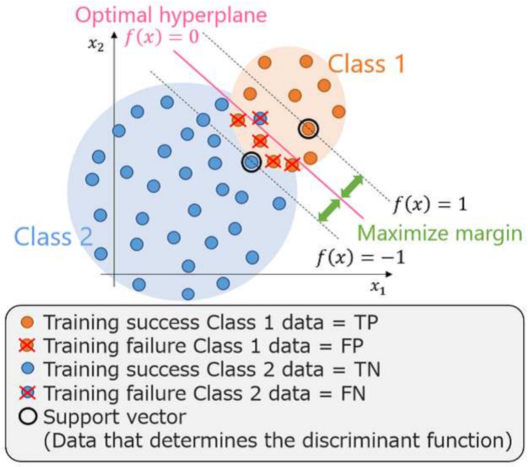

Figure 12: Discrimination problem of imbalanced data.

$$g(x) = f(x) - \frac{-1 \times FP + 1 \times FN}{FP + FN} \tag{1}$$

$$Recall = \frac{TP}{TP + FP} \tag{2}$$

$$Precision = \frac{TP}{TP + FN} \tag{3}$$

As shown in figure 13, moving the optimal hyperplane decreases FP and increases TP. Instead, FN also increases. Equation (4) shows the calculation of the recall$'$ after the improvement in figure 13, and equation (5) shows the calculation of the precision$'$ after the improvement in figure 13. Equation (4) shows that the recall$'$ increases. Equation (5) shows that TP increases even if FN increases. Therefore, the precision$'$ does not drop significantly.

$$Recall' = \frac{TP'}{TP' + FP'} = \frac{TP + 4}{TP + FP} \tag{4}$$

$$Precision' = \frac{TP'}{TP' + FN'} = \frac{TP + 4}{TP + 4 + FN + 1} \tag{5}$$

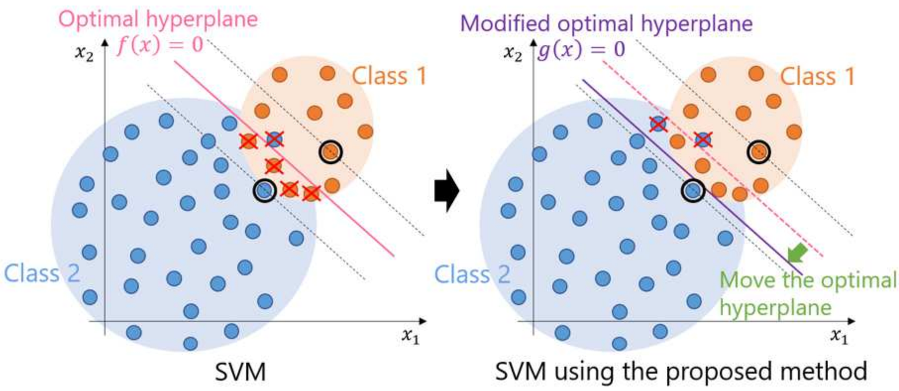

Figure 13: Move the optimal hyperplane.

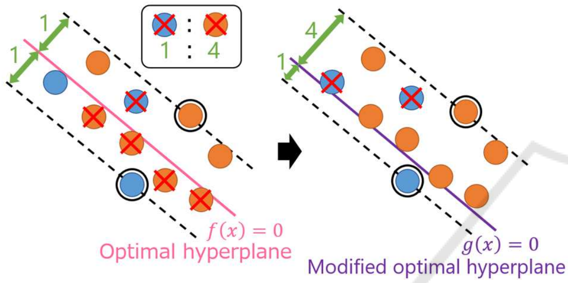

Figure 14: How to move the optimal hyperplane.

Figure 14 shows how to move the optimal hyperplane. Move the optimal hyperplane according to the proportion of the number of training failures within the margin. In the example in figure 14, class 1 has 4 learning failure data and class 2 has 1 learning failure data. In that case, move the optimal hyperplane so that the ratio within the margin is 4:1. However, this proposed method assumes a hard-margin.

4 EXPERIMENTAL CONDITIONS

4.1 Data

The forecast accuracy was verified using the power consumption data of 5,727 large-scale power consumers nationwide published by the Japanese Sustainable open Innovation Initiative (SII) and the data of the Japan Meteorological Agency published on the website of the Japan Meteorological Agency. As shown in figure 10, for an industry that has many excesses in both summer and winter, it is assumed that the features that affect excesses will change between summer (May to October) and winter (November to April). Discriminant functions for such industries were created by dividing the year into summer (May to October) and winter (November to April).

4.2 Features

Feature candidates used for excess forecast alerts are described below. In addition to the feature used in the related research, we added "average temperature for the week up to the forecast execution time t" and "power consumption m hours before the forecast execution time ($m = 1, 2, 3$)" was added. In addition, as shown in the figure 8, we added "holiday information" with 1 for Saturdays, Sundays, and holidays and 0 for weekdays, taking into consideration industries where there is a difference in power consumption between weekdays and holidays.

1. Power consumption at the forecast execution time t
2. Power consumption at the forecast execution time t on the previous day
3. Power consumption at the forecast target time $t+3$ on the previous day
4. Temperature at the forecast execution time t
5. Sunshine hours from time $t-1$ to prediction execution time t
6. Cooling index at the forecast execution time t (How many degrees above 20 degrees Celsius)
7. Heating index at the forecast execution time t (How many degrees below 18 degrees Celsius)
8. Average power consumption for the past m days from the day before the forecast target time $t+3$ ($m = 7, 14, 21, 28$)
9. Average power consumption for 4 weeks from the previous week on the same day of the week at the forecast target time $t+3$
10. Average power consumption for x hours up to the forecast execution time t ($x = 2, 4, 6$)
11. Difference between the amount of power consumption at the forecast execution time t and

Table 2: Number of features used for each industry (Describe the number of features used from Section 2 of Chapter 4).

Industry	Used features number
Wholesale industry	1,2,3,4,5,6,7,8,9,14,15
Accommodation and restaurant industry	1,2,3,4,5,6,7,8,9,14,15,16
Financial industry (summer)	1,2,3,4,5,6,8,9,11, 14,15,16
Financial industry (winter)	3,4,7,8,12,14,16
Manufacturing industry (summer)	1,2,3,4,5,6,8,9,11,14,15,16
Manufacturing industry (winter)	1,2,3,4,5,7,8,9,11,14,15,16
Medical welfare industry (summer)	1,2,3,4,5,6,8,9,11,14,15,16
Medical welfare industry (winter)	1,2,3,4,5,7,8,9,11,14,15,16
Life-related service industry	1,2,3,4,5,6,7,8,9,14,15
Transportation and postal industry	1,2,3,4,5,6,7,8,9,14,15,16
Education industry (summer)	1,2,3,4,5,6,8,9,11,14,15,16
Education industry (winter)	1,2,3,4,5,7,8,9,11,14,15,16
Other industry	1,2,3,4,5,6,7,8,9,14,15,16

the average amount of power consumption for the past m days from the previous day ($m = 7, 14, 21, 28$)

12. Difference between power consumption at forecast execution time t and average power consumption for 4 weeks from the previous week on the same day of the week
13. Day of the week information (dummy variable for each day of the week) and national holiday information (dummy variable for whether it is a holiday)
14. Average temperature for the week up to the forecast target time t
15. Power consumption m hours before the prediction execution time ($m = 1, 2, 3$)
16. Holiday information

Based on Chapter 3, the characteristics of each industry were extracted from the above characteristics. The features used in each industry are shown in table 2. For the industries divided into summer and winter, the feature is shown for each.

4.3 An Excess Alert Based on SVM

Support vector machine (SVM) is used to create a discriminant function that forecasts whether the power consumption after 3 hours will exceed the threshold based on the input data. Soft-margin SVM, which can handle overlapping class distributions, was used to create discriminant functions. Moreover, RBF (Radial Basis Function Kernel) was used as the SVM kernel. The cost parameter C ranges from 100 to 1000, and the RBF kernel parameter γ ranges from 0.01 to 100. Grid search (Syarif et al., 2016) is used to find the most appropriate hyperparameters. Furthermore, a discriminant function was created for each consumer for each hour. However, in our experiments with the proposed improved SVM method for imbalanced data, we used a hard-margin SVM. This is because the improvement proposal method assumes a hard-margin. In addition, we determined how best to move the hyperplane for each industry when using the improved SVM for imbalanced data. Table 3 shows the ratio for each industry.

Table 3: The ratio to move the optimal hyperplane, which is determined for each industry.

Industry	Ratio
Wholesale industry	1.00:2.23
Accommodation and restaurant	1.00:3.00
Financial industry	1.00:4.56
Manufacturing industry	1.00:3.47
Medical welfare industry	1.00:2.77
Life-related service industry	1.00:3.17
Transportation and postal industry	1.00:3.08
Education industry	1.00:3.88
Other industry	1.00:3.88

5 EXPERIMENTAL RESULTS

Section 1 of Chapter 5 presents the results of the proposed method in Section 1 of Chapter 3. Section 2 of Chapter 5 presents the results of the proposed method in Section 2 of Chapter 3.

5.1 Excess Forecast Alert Considering Industry Characteristics

Table 4 shows the accuracy results of the excess forecast alerts in this research, in which discriminant functions were created based on the characteristics

Table 4: An excess alert results by industry using the proposed method (Section 1 of Chapter 3).

Industry	Recall	Precision	F-measure
Wholesale industry	73.44%	83.86%	77.87%
Accommodation and restaurant industry	66.04%	80.95%	72.29%
Financial industry	57.78%	76.61%	65.27%
Manufacturing industry	59.61%	77.31%	66.55%
Medical welfare industry	67.98%	82.22%	74.01%
Life-related service industry	63.96%	80.19%	70.34%
Transportation and postal industry	64.65%	81.06%	71.17%
Education industry	61.77%	79.11%	68.90%
Other industry	63.49%	81.21%	70.43%
Average of all consumers	67.36%	81.33%	73.11%

Table 5: Comparison of the results of related research and this research using the proposed method (Section 1 of Chapter 3).

	Recall	Precision	F-measure
Related research	62.17%	78.74%	68.28%
Proposed method(Section 1 of Chapter 3)(Soft-margin)	67.36%	81.33%	**73.11%**

of each industry of consumers. Even when discriminant functions are used for summer and winter respectively, the forecast accuracy is summed for each industry. Table 5 shows a comparison of accuracy results for excess forecast alerts in related research and excess forecast alert accuracy results in this study, in which a discriminant function was created based on the characteristics of each industry of consumers. We were able to improve the F-measure by about 4.8 points.

Since the amount of power consumption by each consumer industry has its own characteristics, we changed the conditions for creating the discriminant function for each consumer industry to improve the accuracy of the excess forecast alert. As a result of creating a discriminant function according to the characteristics of each consumer industry, we were able to improve the forecast accuracy (recall, precision, F-measure) of the average excess forecast alert for all consumers. As a result, we can send effective alerts to consumers in various industries. However, the recall is 67.36%, which is low compared to the precision of 81.33%. Lower recall means more oversights. We want to aim for an excess forecast alert with a high F-measure and recall.

5.2 Excess Forecast Alerts Based on Improved SVM for Imbalance Data

Table 6 shows the accuracy results of the excess forecast alert in this study, which created the discriminant function based on the improved SVM for the imbalanced data. Table 7 compares the average results for all consumers in Table 6 and the accuracy results for the excess forecast alerts in this study, in which discriminant functions were created based on the characteristics of each industry of consumers using the proposed method(Section 1 of Chapter 3). We were able to improve the recall by about 6.8 points and the F-measure by about 1.3 points.

The excess forecast alert result based on improved SVM for imbalanced data could obtain high recall while improving F-measure. A high recall means a reduction in the number of missed oversights.

6 CONCLUSION

In this research, since the amount of electricity used is characteristic for each industry of consumers, we changed the conditions for creating the discriminant function for each industry of consumers to improve the accuracy of excess forecast alerts. As a result of creating a discriminant function based on the characteristics of each consumer's industry, it was possible to improve the recall, precision, and F-measure of the average excess forecast alert for all consumers. In addition, excess forecast alerts using the improved SVM for imbalanced data were able to produce results with high F-measures and recall. This makes it possible to contribute to the transmission of effective alerts to consumers in a wide variety of industries. However, in order to have a wide range of consumers use it in the future, further improvement in accuracy is required. The accuracy of excess forecast alerts will be improved in our future research.

Table 6: An excess alert results by industry using the proposed method(Section 2 of Chapter 3).

Industry	Recall	Precision	F-measure
Wholesale industry	75.64%	74.55%	74.94%
Accommodation and restaurant industry	72.11%	66.87%	69.28%
Financial industry	68.50%	61.66%	64.66%
Manufacturing industry	67.02%	64.84%	65.71%
Medical welfare industry	72.54%	70.27%	71.27%
Life-related service industry	68.33%	68.65%	68.21%
Transportation and postal industry	68.06%	69.46%	68.49%
Education industry	72.34%	62.24%	66.79%
Other industry	71.44%	66.26%	68.61%
Average of all consumers	72.33%	69.55%	70.72%

Table 7: Comparison of the results of this research using the proposed method (Section 1 of Chapter 3) and this research using the proposed method (Section 2 of Chapter 3).

	Recall	Precision	F-measure
Proposed method (Section 1 of Chapter 3) (Hard-margin)	65.41%	74.75%	69.41%
Proposed method (Section 2 of Chapter 3) (Hard-margin)	**72.33%**	69.55%	**70.72%**

REFERENCES

Ahmad, M., Aftab, S., Bashir, M. S., Hameed, N., Ali, I., and Nawaz, Z. (2018). Svm optimization for sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 9(4).

Burges, C. J., Smola, A. J., and Scholkopf, B. (1999). Advances in kernel methods. *Support Vector Learning*, 53.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Goutte, C. and Gaussier, E. (2005). A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European conference on information retrieval*, pages 345–359. Springer.

Haghi, A. and Toole, O. (2013). The use of smart meter data to forecast electricity demand. *CS229 course paper*.

He, H. and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9):1263–1284.

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, 5(4):221–232.

Mohammed, R., Rawashdeh, J., and Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: overview study and experimental results. In *2020 11th international conference on information and communication systems (ICICS)*, pages 243–248. IEEE.

Motamedi, A., Zareipour, H., and Rosehart, W. D. (2012). Electricity price and demand forecasting in smart grids. *IEEE Transactions on Smart Grid*, 3(2):664–674.

Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

Suganthi, L. and Samuel, A. A. (2012). Energy models for demand forecasting - a review. *Renewable and sustainable energy reviews*, 16(2):1223–1240.

Sun, Y., Wong, A. K., and Kamel, M. S. (2009). Classification of imbalanced data: A review. *International journal of pattern recognition and artificial intelligence*, 23(04):687–719.

Syarif, I., Prugel-Bennett, A., and Wills, G. (2016). Svm parameter optimization using grid search and genetic algorithm to improve classification performance. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 14(4):1502–1509.

Tokoro, K., Tsurumi, T., and Higo, T. (2019). Excessive power consumption alert using svm (in japanese). In *The Institute of Electrical Engineers of Japan National conference*.

Vapnik, V. (1999a). *The nature of statistical learning theory*. Springer science & business media.

Vapnik, V. N. (1999b). An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5):988–999.