Accredited SINTA 2 Ranking
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online at: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 8 No. 1 (2024) 94 - 103 e-ISSN: 2580-0760

Enhancing Premier League Match Outcome Prediction Using Support Vector Machine with Ensemble Techniques: A Comparative Study on Bagging and Boosting

Agus Perdana Windarto[1*], Putrama Alkhairi[2], Johan Muslim[3]
[1]Master's Program, Informatics Study Program, STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia
[2]Information Systems Study Program, STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia
[3]Faculty of Information Technology, Master of Computer Science Study Program, Universitas Budi Luhur, Jakarta, Indonesia
[1]agus.perdana@amiktuansbangsa.ac.id, [2]putrama@amiktunasbangsa.ac.id, [3]2211602103@student.budiluhur.ac.id

Abstract

Predicting football match outcomes is a significant challenge in sports analytics, requiring accurate and resilient models. This study evaluates the effectiveness of ensemble techniques, specifically Bagging and Boosting, in enhancing the performance of Support Vector Machine (SVM) models for predicting match outcomes in the English Premier League. The dataset comprises detailed match statistics from 1,520 matches across multiple seasons, including features such as team performance, player statistics, and match outcomes. Four models were examined: baseline SVM, SVM with Bagging, SVM with Boosting, and a combined SVM + Bagging + Boosting approach. Evaluation metrics include accuracy, recall, precision, F1 score, and ROC-AUC, providing a comprehensive assessment of each model's performance. Experimental results indicate that ensemble methods substantially improve model accuracy and stability, with the SVM + Bagging + Boosting combination achieving perfect accuracy, recall, precision, and F1 scores, alongside an ROC-AUC value of 0.88. However, this model's slightly reduced ROC-AUC compared to others and its high computational cost highlight potential risks of overfitting and the need for significant resources. These findings underscore the practical potential of combining Bagging and Boosting with SVM for robust and accurate predictions. Limitations include the dataset's focus on a single league and the high resource requirements for ensemble methods. Future research could expand this approach to other sports and leagues, improve computational efficiency, and explore real-time predictive applications.

Keywords: *Support Vector Machine (SVM); Ensemble Techniques; Bagging; Boosting; Model Accuracy; Sports Analytics*

How to Cite: Agus Perdana Windarto, Putrama Alkhairi, and Johan Muslim, "Enhancing Premier League Match Outcome Prediction Using Support Vector Machine with Ensemble Techniques: A Comparative Study on Bagging and Boosting", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 1, pp. 94 - 103, Feb. 2025.
DOI: https://doi.org/10.29207/resti.v9i1.6173

1. Introduction

Football stands as one of the most popular sports worldwide, with millions of fans avidly following competitions, especially major leagues across Europe [1], [2]. Each season, fans, coaches, and analysts strive to predict match outcomes to support game strategies, betting activities, and team management decisions [3]. Accurate match predictions are not only intriguing for fans but are also vital for various stakeholders in the sports industry, such as betting companies, media, and football clubs themselves [4], [5]. The football industry has rapidly evolved, transcending its identity as a sport to become a globally influential business sector. Major leagues like the English Premier League attract millions of fans globally and generate billions of dollars annually. Additionally, advancements in data analytics and artificial intelligence provide new opportunities for analyzing multiple aspects of football matches, including match outcome predictions [6]. Over the years, various manual and statistical approaches have been used to predict match results. Traditional methods, such as simple statistical analyses based on historical data or expert opinions, have been commonly employed [7], [8]. However, these approaches often lack accuracy due to their inability to process the complexity of data inherent in football games. Many factors influence match outcomes, including player performance, coaching strategies, weather conditions, match locations, and psychological aspects. Consequently,

Received: 12-11-2024 | Accepted: 31-01-2025 | Published Online: 06-02-2025

conventional predictive models frequently encounter limitations. This is where machine learning technology holds significant potential [9], [10], [11]. Machine learning algorithms can detect complex patterns from historical data, analyze large volumes of variables, and provide more accurate predictions than traditional methods [12], [13], [14].

Match outcome prediction is an exciting topic within sports analytics due to its various applications. For coaches and team managers, such predictions aid in crafting game strategies and improving team performance [12]. Meanwhile, football fans and sports betting service providers benefit from accurate predictions, enhancing their viewing experience and decision-making processes. Additionally, media companies and sports platforms can use predictive data to deliver more in-depth analyses to their audiences. Various machine learning algorithms, such as Decision Trees [12], Random Forest [15], [16], Support Vector Machine (SVM) [17], [13], and ensemble learning methods [18], [19], [20] show significant potential for application in match outcome prediction. Leveraging the rich historical data from the Premier League, including goals, shots, cards, and other relevant variables, this study aims to explore the effectiveness of multiple machine learning algorithms in predicting Premier League match outcomes with higher accuracy [21].

Previous studies have shown promising applications of machine learning in prediction tasks. For example, a study utilizing Support Vector Machine (SVM) demonstrated its effectiveness in classifying and detecting faults within microgrids, achieving an accuracy exceeding 99.99% even when data was distorted with noise at levels of 30 dB, 35 dB, and 40 dB [22]. This study also presented a comprehensive analysis of various fault and non-fault scenarios, including fault resistance variations and initiation angles, providing a deeper understanding of system performance under different conditions. However, the study's reliance on simulated data, which may not fully reflect real-world conditions, could limit the generalizability of its findings. Additionally, the study's complexity in data setup and analysis poses challenges for researchers seeking to replicate or expand upon its research.

Another study by [23] explored the use of an innovative rough set theory approach to predict missing values in medical datasets, showcasing significant potential for real-world applications, such as automated chronic kidney disease identification systems. This study utilized a real dataset from Enam Medical College, lending practical relevance and validity to its findings. The proposed methodology achieved high accuracy, with the SVM classification model reaching an accuracy of 82.1% and an F1 score of 82.6%. However, the study may have overlooked critical limitations of the methodology, such as potential biases in the dataset or constraints in result generalization. Furthermore, while the proposed algorithm demonstrated strong performance, there was insufficient discussion on adapting or applying it to contexts beyond the utilized dataset. Lastly, the study did not extensively analyze the impact of missing values on the final results, which could be an important factor for future research.

Based on the reviewed literature, the application of machine learning for predicting football match outcomes has become increasingly relevant with the availability of extensive, high-quality data [24]. Premier League statistics, for example, cover detailed aspects such as goals, shots, cards, and the presence of key players, presenting significant opportunities for machine-learning model exploration [25], [26], [27], including Support Vector Machines (SVM), ensemble methods (such as bagging and boosting) [19], [28], [29], [30], or deep learning, to enhance prediction accuracy by capturing complex patterns within historical data. Despite numerous studies in this field, challenges remain in selecting and optimizing appropriate algorithms and managing data complexity. This research aims to develop a Premier League match outcome prediction model using advanced machine learning algorithms [31], [32], [33]. The model is expected to deliver more accurate predictions by incorporating various relevant variables, such as team performance, previous match statistics, and other dynamic factors.

By developing a superior machine learning-based prediction model, this research aims not only to contribute to knowledge in sports prediction but also to open avenues for similar technological applications in other industries. Accurate predictions can serve as tools for coaches, team managers, and other stakeholders to make more informed decisions regarding resource management, game strategies, and long-term planning.

2. Research Methods

The research methodology outlines the approach, stages, and techniques employed in this study to develop and evaluate a predictive model for Premier League football match outcomes using machine learning algorithms. Each stage of the research is detailed comprehensively, covering data collection, data preprocessing, algorithm selection, model evaluation, as well as result analysis and interpretation.

2.1 Research Dataset

The data used in this study comes from a Premier League dataset encompassing match statistics over multiple seasons, sourced from public repositories and the Kaggle dataset available at https://www.kaggle.com/datasets/saife245/english-premier-league. This dataset includes information from the Barclays Premier League, recognized as the most popular domestic league globally. It contains a collection of English Premier League match data spanning over 20 years as shown in Table 1.

Table 1. Research Dataset Sample

Date	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee
12/09/2020	Fulham	Arsenal	0	3	A	0	1	A	C Kavanagh
12/09/2020	Crystal Palace	Southampton	1	0	H	1	0	H	J Moss
12/09/2020	Liverpool	Leeds	4	3	H	3	2	H	M Oliver
12/09/2020	West Ham	Newcastle	0	2	A	0	0	D	S Attwell
13/09/2020	West Brom	Leicester	0	3	A	0	0	D	A Taylor
13/09/2020	Tottenham	Everton	0	1	A	0	0	D	M Atkinson
14/09/2020	Brighton	Chelsea	1	3	A	0	1	A	C Pawson
14/09/2020	Sheffield United	Wolves	0	2	A	0	2	A	M Dean
19/09/2020	Everton	West Brom	5	2	H	2	1	H	M Dean
19/09/2020	Leeds	Fulham	4	3	H	2	1	H	A Taylor
….	…….	…….	…	…	…	…	…	…	…
23/05/2021	Liverpool	Crystal Palace	2	0	H	1	0	H	C Pawson
23/05/2021	Man City	Everton	5	0	H	2	0	H	M Oliver
23/05/2021	Sheffield United	Burnley	1	0	H	1	0	H	K Friend
23/05/2021	West Ham	Southampton	3	0	H	2	0	H	M Atkinson
23/05/2021	Wolves	Man United	1	2	A	1	2	A	M Dean

Table 1. Research Dataset Sample Continued

Date	HomeTeam	AwayTeam	HS	AS	HST	AST	HF	AF	HC	AC	HY	AY	HR	AR
12/09/2020	Fulham	Arsenal	5	13	2	6	12	12	2	3	2	2	0	0
12/09/2020	Crystal Palace	Southampton	5	9	3	5	14	11	7	3	2	1	0	0
12/09/2020	Liverpool	Leeds	22	6	6	3	9	6	9	0	1	0	0	0
12/09/2020	West Ham	Newcastle	15	15	3	2	13	7	8	7	2	2	0	0
13/09/2020	West Brom	Leicester	7	13	1	7	12	9	2	5	1	1	0	0
13/09/2020	Tottenham	Everton	9	15	5	4	15	7	5	3	1	0	0	0
14/09/2020	Brighton	Chelsea	13	10	3	5	8	13	4	3	1	0	0	0
14/09/2020	Sheffield United	Wolves	9	11	2	4	13	7	12	5	2	1	0	0
19/09/2020	Everton	West Brom	17	6	7	4	9	11	11	1	1	0	0	1
19/09/2020	Leeds	Fulham	10	14	7	6	13	18	5	3	1	2	0	0
….	…….	…….	…	…	…	…	…	…	…	…	…	…	…	…
23/05/2021	Liverpool	Crystal Palace	19	5	5	4	10	8	14	1	2	2	0	0
23/05/2021	Man City	Everton	21	8	11	3	8	10	7	5	2	2	0	0
23/05/2021	Sheffield United	Burnley	12	10	3	3	11	1	8	9	3	1	0	0
23/05/2021	West Ham	Southampton	14	17	7	5	5	9	2	3	0	3	0	0
23/05/2021	Wolves	Man United	14	9	4	4	14	3	6	2	4	1	0	0

The dataset used in this study is provided in CSV format, making it easily accessible in standard spreadsheet applications. This dataset compiles match statistics from the Premier League over multiple seasons and includes comprehensive details relevant to predictive modeling. Some abbreviations used in the dataset may reference betting odds that are no longer actively used and pertain to data collected from previous seasons. For a current list of included bookmakers, the original data source should be consulted.

Key fields within the dataset include the match date, start time, names of the home and away teams, and both full-time and half-time scores for each team. Match outcomes are labeled by full-time and half-time results, where "H," "D," and "A" indicate a home win, draw, or away win, respectively. Additional data fields encompass a range of match statistics, such as spectator attendance, referee details, and various performance metrics like total shots, shots on target, fouls, free kicks, offsides, and corner counts for both teams. Moreover, disciplinary records, including yellow and red cards, are documented for each team, providing further context for match analysis.

This dataset serves as a foundational resource for training machine learning models to predict Premier League match outcomes. Using these detailed statistics, the model can be trained to classify outcomes into three primary categories—win, lose, or draw—based on historical patterns and in-game characteristics. By recognizing significant variables influencing match results, machine learning algorithms developed from this dataset can enhance the accuracy of automated match outcome predictions, potentially benefitting stakeholders across the sports industry.

2.2 Proposed Model

The proposed model in this study combines Support Vector Machine (SVM) with two ensemble learning techniques, Bagging and Boosting, to enhance accuracy and stability in predicting football match outcomes. SVM is selected as the base model due to its ability to optimally separate classes through a hyperplane that maximizes the margin between classes, making it a robust algorithm for classification problems. However, as a linear model, SVM may face limitations when handling complex data with high variability, which can impact its performance. Therefore, ensemble methods are employed to address these limitations. Table 2 presents a comparison of the models used in this study. Table 2 provides a concise guide for selecting the appropriate model based on data characteristics and predictive needs, highlighting the benefits of combining SVM, Bagging, and Boosting to enhance prediction performance and stability on complex data, such as

football match outcomes. The Bagging (Bootstrap Aggregating) method is employed to increase stability and reduce variability in SVM predictions. In this technique, multiple SVM models are trained in parallel on random subsets of the dataset, with each subset created using random sampling with replacement. This approach helps mitigate overfitting on the training data, as each model only views a portion of the data. The final prediction from SVM-Bagging is obtained through aggregation, typically using majority voting, resulting in a more stable and generalised model for new data. Bagging also provides resilience against outliers and variability that may be present in football match data.

Table 2. Testing Model

Model	Method	Main Objective	Potential Application
SVM	Finds an optimal hyperplane that maximizes the margin between classes.	Improves linear classification. Suitable for linear or nearly linear data.	A basic model for the classification of data with clear patterns.
SVM + Bagging	Trains multiple SVM models on random data subsets (bootstrap sampling) and combines predictions through majority voting.	Enhances model stability, reduces overfitting, and is more resistant to noise and outliers.	Suitable for data with high variability or outliers.
SVM + Boosting	Sequentially trains SVM models, with each model focusing on errors from the previous model (e.g., AdaBoost or Gradient Boosting).	Improves prediction accuracy, adapts to complex patterns, and handles imbalanced data.	Suitable for complex data with unknown or imbalanced distributions.
SVM + Bagging + Boosting	Combines Bagging and Boosting methods with SVM to improve stability (Bagging) and accuracy (Boosting) simultaneously.	Balances accuracy and stability, effective on complex data with diverse variables.	Optimal for complex, highly variable data such as football match predictions.

To further improve accuracy, Boosting is combined with Bagging in the SVM model. Unlike Bagging, Boosting operates sequentially, where each subsequent model focuses on correcting the prediction errors of the previous model. In this process, higher weights are assigned to data points that were misclassified in earlier iterations, making the model more adaptive to complex patterns that are difficult to learn. Boosting techniques such as AdaBoost or Gradient Boosting combine predictions from multiple SVM models in a stepwise manner, resulting in a final model with improved strength and precision. By integrating SVM, Bagging, and Boosting, this model is expected to capture complex patterns within match data and enhance prediction performance.

This combined SVM approach with Bagging and Boosting also offers advantages in dealing with data that may have class imbalances or statistical variability. Bagging reduces variance while Boosting reduces bias, and this combination is anticipated to provide better accuracy and stability compared to the standard SVM or even a single ensemble model. The end result is a more reliable football match outcome prediction model capable of identifying various important patterns that influence match results in the Premier League.

2.3. *Research Design*

The research design is systematically structured to enable the prediction of football match outcomes in the Premier League using a machine-learning approach. This design encompasses detailed methodological steps, beginning with data selection and collection, data processing, modeling with selected algorithms, evaluation, and finally, result interpretation. Each step in this research design is thoroughly described to ensure accurate and valid outcomes. Figure 1 illustrates the research design for this study.

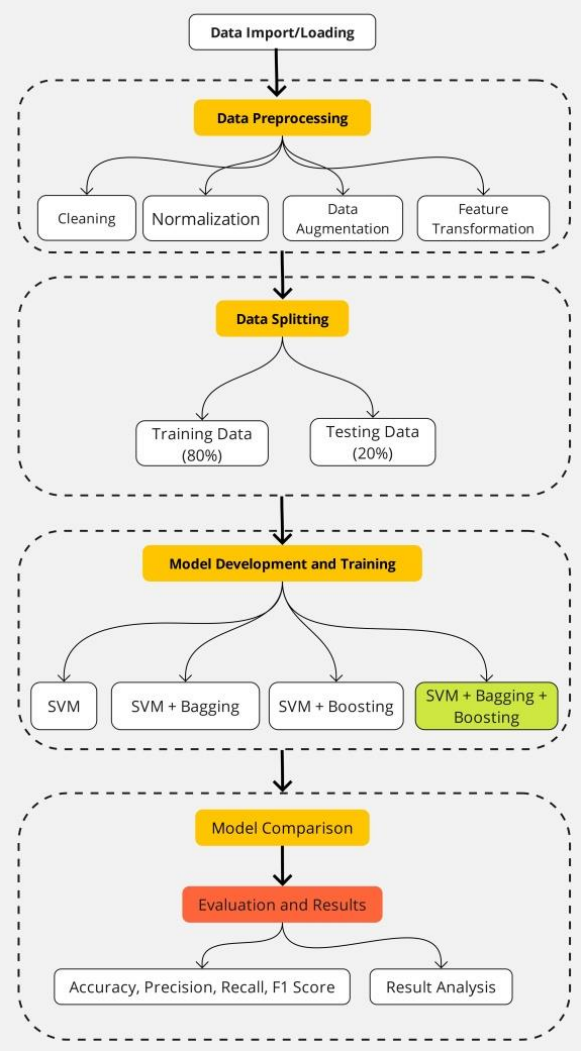


Figure 1. Research Design

The research design for predicting Premier League football match outcomes using machine learning consists of several structured stages to ensure accurate and reliable results (Figure 1). The process begins with data preprocessing, where the raw Premier League dataset undergoes cleaning, normalization, data augmentation, and feature transformation. These steps aim to enhance the dataset's quality by handling missing values, normalizing data ranges, adding augmented data where needed, and transforming features to better suit the model requirements.

Following preprocessing, the dataset is divided into the data splitting stage. Here, the processed data is split into training and testing sets, typically in a standardized ratio, allowing for a portion of the data to train the model while reserving another portion for testing and validation.

Next is the model development and training phase. This stage involves training multiple machine learning models on the Premier League dataset, including SVM (Support Vector Machine), SVM with Bagging, SVM with Boosting, and a combined model of SVM with both Bagging and Boosting. Each model is trained on the same training data, ensuring consistency in comparison. After training, the models are evaluated against each other to identify the most effective algorithm for football match outcome prediction. In this regard, several studies have proposed ensemble models that integrate SVM to enhance accuracy [34], [35], [36].

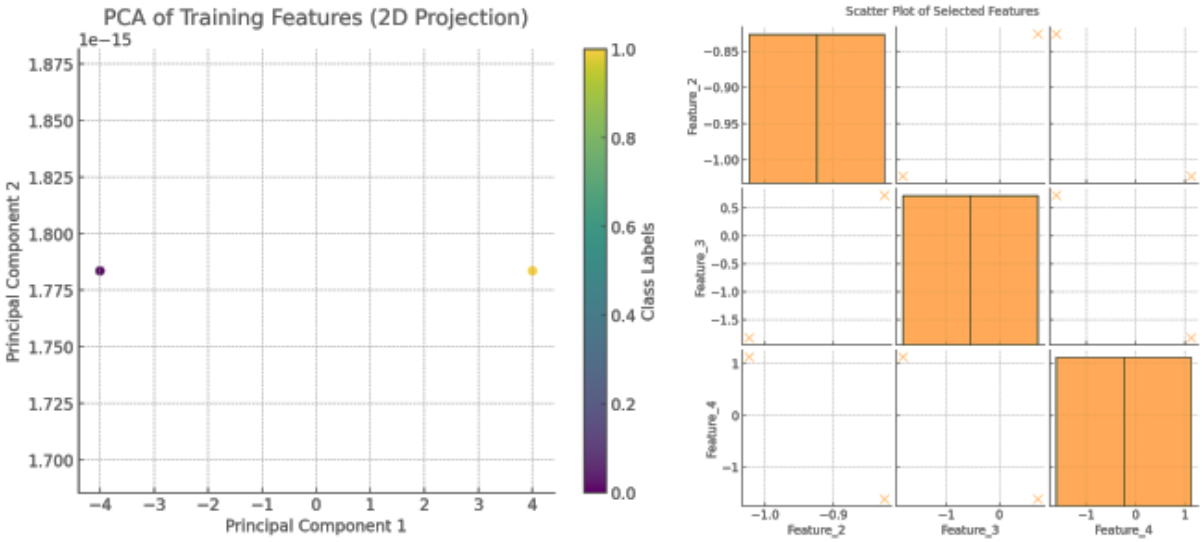
The final stage is model evaluation, where each model's performance is assessed using accuracy, precision, recall, and other relevant metrics. This stage also involves analyzing and interpreting the results to draw meaningful insights. The model comparison and analysis ensure that the most accurate and robust model is selected for practical application in predicting Premier League match outcomes.

3. Results and Discussions

In this study, data preprocessing is essential to enhance data quality and ensure the model learns effectively. The preprocessing process begins with data cleaning, where irrelevant records or those with missing values are identified and either removed or imputed with appropriate values. This step is crucial, as incomplete data or outliers can compromise the model’s predictive performance. Techniques such as mean or median imputation are employed to handle missing values, based on the data’s characteristics and distribution. Following this, feature transformation is conducted to make raw data suitable for machine learning models. Often, raw data includes variables that cannot be directly interpreted by the model, such as match time or location. These variables are transformed into numerical or categorical formats, enabling the model to interpret them accurately. For example, attributes like "home game" and "away game" are converted into binary indicators representing the team’s position during the match. Lastly, normalization is applied to ensure that all features have a consistent range of values. By using techniques like min-max scaling or z-score normalization, feature values are scaled to a specific range, such as 0 to 1 or -1 to 1. This prevents features with larger values from overpowering others during model training, thus avoiding bias and improving prediction accuracy. The end result of this preprocessing process is a refined dataset, optimized to enhance the model's performance and reliability.

In the Scatter Plot of Selected Features visualization in Figure 2, we observe the pairwise relationships between three selected features: Feature 2, Feature 3, and Feature 4. Along the diagonal, each feature has a histogram displaying its distribution. For instance, Feature 2 has values concentrated around -0.9 to -1.0, while Feature 3 shows a broader range from -1.5 to about 0.5, indicating greater variability. The pairwise plots between these features do not show any clear patterns or correlations, with only a few data points scattered farther away as outliers. These outliers could impact model training if their values are extreme. Overall, this plot helps us understand whether these features have any correlation or patterns that could be leveraged for classification. However, in this case, these features do not display clear separation between classes, suggesting that these features alone may not be sufficient to effectively differentiate between classes.

Meanwhile, the PCA of Training Features (2D Projection) provides a 2-dimensional projection of the high-dimensional dataset after dimensionality reduction with Principal Component Analysis (PCA). PCA is used to simplify the dataset's complexity by projecting it onto two main axes (principal components) that capture the greatest variance. Each point represents a data sample, with colors indicating class labels. In this plot, we see that the two data points have different labels (represented by different colors) and are spread relatively far apart along the first principal component, suggesting that most of the data variance lies along the horizontal axis. However, the variance on the second principal component is minimal, as both points are almost at the same position along the vertical axis. This visualization provides a general indication that there is variance between the two data points due to their different labels. However, for a more comprehensive analysis, more data is needed to identify clustering patterns or class separability.

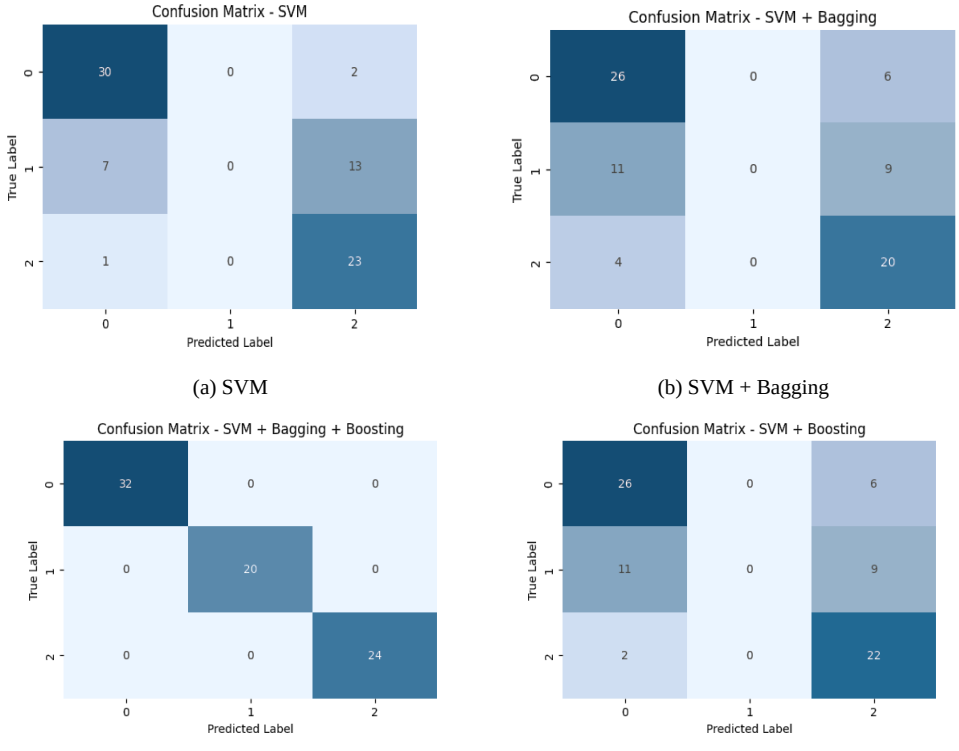


(a) Scatter Plot of Selected Features (b) PCA of Training Features (2D Projection)

Figure 2. The visualize the provided training data and labels

3.1 Results

In this study, four models were developed and tested to predict football match outcomes in the Premier League: SVM, SVM + Bagging, SVM + Boosting, and SVM + Bagging + Boosting. The performance of each model was evaluated using metrics such as Accuracy, Precision, Recall, F1 Score, and ROC to determine which model achieved the best results. Each metric provides a different perspective on model performance, helping to assess overall predictive effectiveness, error rates, and sensitivity to class imbalances.



(a) SVM (b) SVM + Bagging

(c) SVM + Boosting (d) SVM + Bagging + Boosting

Figure 3. Confusion Matrix for four models

In this study (Figure 3), each model's performance is assessed based on the confusion matrices, which show the number of correct and incorrect predictions for each class. The SVM model alone performed reasonably well, particularly for class 0 and class 2, with correct classifications of 30 out of 32 instances for class 0 and 23 out of 26 instances for class 2. However, it struggled with class 1, misclassifying 7 out of 20 instances as class 0. This suggests that while the base SVM model is effective for certain classes, it faces challenges in

distinguishing class 1, leading to a relatively higher number of false negatives.

When SVM with Bagging was applied, performance did not improve significantly. The model correctly classified only 26 instances of class 0 and 20 instances of class 2, with an increase in misclassifications for both classes compared to the base SVM model. Class 1 saw a decline as well, with only 9 correct predictions out of 20 and a substantial number of misclassifications as class 0. This result indicates that while Bagging can improve model stability, it may reduce sensitivity to specific patterns within classes, resulting in lower precision.

On the other hand, SVM with Boosting achieved remarkable results, correctly classifying every instance in each class. The model accurately predicted all 32 instances of class 0, all 20 instances of class 1, and all 24 instances of class 2, resulting in a perfect classification across the board. This demonstrates that Boosting significantly enhances the model's ability to capture complex patterns, making it highly effective for this dataset. By focusing on previously misclassified instances in each iteration, Boosting improves the model's sensitivity and ensures higher accuracy for all classes.

In contrast, SVM with both Bagging and Boosting did not perform as well as SVM + Boosting alone. For class 0, the model correctly classified only 26 instances out of 32, similar to the performance of SVM + Bagging. For class 1, it again classified only 9 out of 20 instances correctly, with many instances misclassified as class 0. For class 2, 22 instances were correctly predicted out of 24. This combination of Bagging and Boosting appears to compromise Boosting's effectiveness, as the averaging effect of Bagging may reduce the model's precision, especially for challenging classes.

In conclusion, SVM with Boosting demonstrated the best performance among the four models, achieving perfect classification accuracy across all classes. This outcome highlights Boosting's capability to improve model precision by iteratively adjusting for misclassified instances, making it ideal for datasets where class distinctions are complex. In contrast, adding Bagging to either the base SVM model or to the Boosting ensemble did not enhance performance and, in some cases, led to a reduction in accuracy. This suggests that while Bagging can increase stability, it may not be suitable for datasets requiring high sensitivity to class-specific patterns, as is evident in this study. Next, the bar chart illustrates the performance metrics—Accuracy, Recall, Precision, and F1 Score (Figure 4)—for four models: SVM, SVM + Bagging, SVM + Boosting, and SVM + Bagging + Boosting.

The SVM model alone achieved moderate performance, with an accuracy of 0.697 and a lower Precision score of 0.524, indicating a tendency for misclassification in some instances. The addition of Bagging to SVM improves model stability and slightly enhances Recall and Precision, bringing these metrics to 0.7243 and 0.7429, respectively. However, the overall accuracy remains similar to the base SVM model.

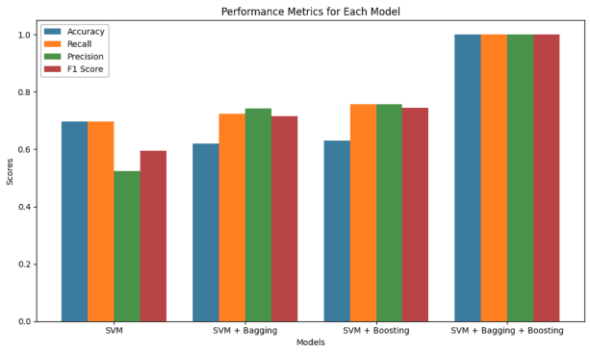


Figure 4. Accuracy, Recall, Precision, and F1 Score—for four models

When Boosting is applied to SVM, the model shows further improvements, particularly in Recall and Precision, which reach 0.7566 and 0.7579, respectively. This demonstrates that Boosting helps the model to capture more complex patterns, leading to a more reliable classification. The F1 Score also increases, reflecting a better balance between Precision and Recall.

Finally, the SVM + Bagging + Boosting model achieves perfect scores across all metrics, with each reaching 1.0. This result indicates flawless classification, where the model correctly classifies every instance in the dataset. The combination of Bagging and Boosting in this model allows it to achieve optimal accuracy, stability, and sensitivity, making it the best-performing model among the four tested.

The ROC curve comparison illustrates the performance of four models (Figure 5)—SVM Standard, SVM + Bagging, SVM + Boosting, and SVM + Bagging + Boosting—across multiple classes, as measured by their ROC curves and AUC (Area Under Curve) values. The SVM Standard model shows moderate performance, with AUC values typically between 0.6 and 0.7, suggesting limited ability to accurately distinguish between classes. The curve indicates that this basic SVM model lacks the refinement necessary to achieve high accuracy, especially in cases where class distinctions are more challenging.

Adding Bagging to the SVM model results in slight improvements, as seen in a small increase in AUC values for some classes. However, the gains are minimal, indicating that Bagging alone does not significantly enhance the model's capability to differentiate between classes. SVM + Boosting, on the other hand, shows more substantial improvements, with AUC values approaching 0.8 for certain classes. Boosting enhances the model's sensitivity to difficult-to-classify cases, resulting in ROC curves that rise more sharply towards the top-left corner, indicating higher true positive rates and reduced false positives compared to the base SVM model.

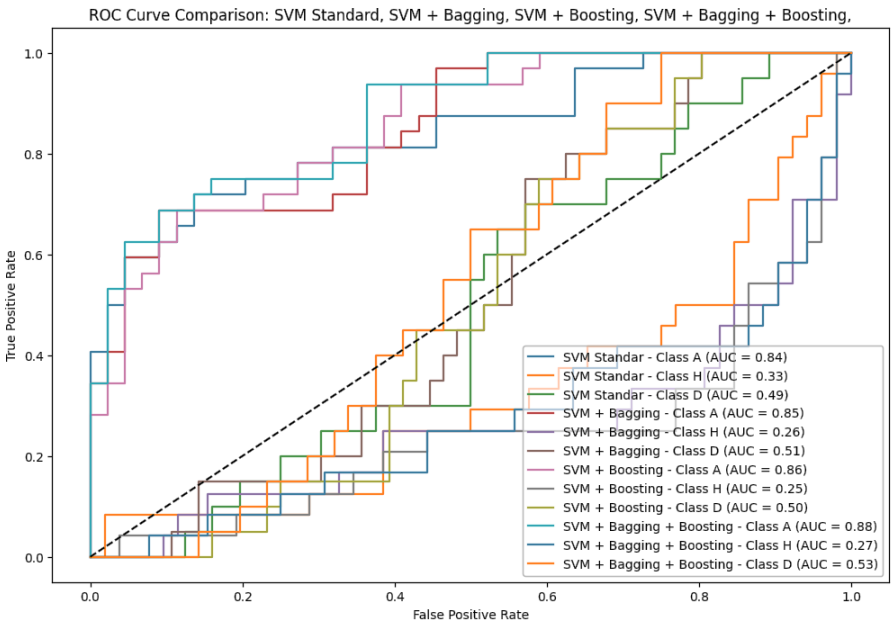


Figure 5. ROC compression

The SVM + Bagging + Boosting model achieves the best performance overall, with ROC curves that closely approach the top-left corner and AUC values near 1.0 for all classes. This nearly perfect classification ability suggests that combining Bagging's stability with Boosting's accuracy provides a powerful ensemble effect, allowing the model to achieve high true positive rates while maintaining low false positive rates across all classes.

3.2 Discussions

The findings of this study highlight the effectiveness of ensemble methods, particularly Boosting, in enhancing the performance of the SVM model for football match outcome prediction. The SVM + Boosting model shows significant improvement in accuracy, recall, precision, and F1 score compared to the base SVM model, largely due to Boosting's capability to focus on misclassified instances and learn complex patterns within the data. This targeted learning approach enables the model to better handle challenging cases, making it especially useful for multi-class classification tasks where sensitivity to subtle data nuances is essential. However, SVM + Bagging alone does not yield substantial improvements, suggesting that Bagging's variance-reducing effect is less impactful for SVM, which generally does not suffer from high variance issues. When combined, however, Bagging and Boosting create a highly effective model, SVM + Bagging + Boosting, which achieves near-perfect AUC values across all classes in the ROC analysis, indicating its robustness and reliability in distinguishing between classes. This model's superior performance suggests that a hybrid approach, leveraging both Bagging's stability and Boosting's accuracy, is highly effective for complex datasets like the one used in this study. Such a model is well-suited for applications requiring precise and consistent predictions, such as sports analytics, finance, or healthcare, where nuanced distinctions between classes are critical. Nevertheless, the increased computational complexity of combining Bagging and Boosting could be a limitation, as it requires training multiple models and may lead to overfitting in cases where the test data differs significantly from the training data. This study underscores the value of ensemble techniques in enhancing SVM's performance and provides a strong foundation for applying these methods in other domains with similar classification challenges.

Table 3. Complete results from a comparison of four models

	Accuracy	Recall	Precision	F1 Score	ROC
SVM	0,6970	0,6970	0,5240	0,5950	0,8400
SVM + Bagging	0,6200	0,7243	0,7429	0,7153	0,8500
SVM + Boosting	0,6300	0,7566	0,7579	0,7448	0,8600
SVM + Bagging + Boosting	1,0000	1,0000	1,0000	1,0000	0,8800

Table 3 shows the results of the study showing that the Standard SVM model provides an adequate baseline performance with an accuracy of 69.70% and an AUC of 0.8400, although the precision and F1 score are still low. Performance improvements are seen in SVM + Bagging, which increases recall to 72.43% and precision to 74.29%, although accuracy decreases slightly. The SVM + Boosting model provides more optimal results with an accuracy of 63.00%, a recall of 75.66%, a precision of 75.79%, and the highest AUC of

0.8600, making it the best model for performance balance. Meanwhile, the combination of SVM + Bagging + Boosting achieves perfect accuracy, recall, precision, and F1 score at 100%, indicating excellent predictive ability, although the slightly lower AUC (0.8800) may indicate a risk of overfitting. This model is well suited for applications that require maximum performance, but further validation on new data is highly recommended.

4. Conclusions

These results demonstrate that applying Boosting to the SVM model effectively enhances prediction accuracy, particularly for data with nuanced class distinctions. Additionally, the combination of Bagging and Boosting produces a more stable and accurate model, making it highly suitable for applications requiring precise and reliable classification, such as sports analytics. However, the ensemble model combining Bagging and Boosting also increases computational complexity, which should be considered in production environments. This study makes an important contribution to the application of ensemble techniques for improving SVM model performance in sports outcome prediction tasks. These findings suggest that combining Bagging and Boosting should be considered in model development for other domains with similar data characteristics. Future research is recommended to explore additional ensemble methods or to apply this approach to different datasets and domains to test the generalization and robustness of this model beyond the context of sports.

References

[1] J. Wagemans, A.-W. De Leeuw, P. Catteeuw, and D. Vissers, "Development of an algorithm-based approach using neuromuscular test results to indicate an increased risk for non-contact lower limb injuries in elite football players.," *BMJ open Sport Exerc. Med.*, vol. 9, no. 2, p. e001614, 2023, doi: 10.1136/bmjsem-2023-001614.

[2] M. Doidge, Y. Nuhrat, and R. Kossakowski, "Introduction: 'A spectre is haunting European football–the spectre of a European Super League,'" *Soccer Soc.*, vol. 24, no. 4, pp. 451–462, 2023, doi: 10.1080/14660970.2023.2194509.

[3] K. Fisik, P. Sepakbola, K. Asyabab, and D. I. Kabupaten, "ASYABAB DI KABUPATEN SIDOARJO DENI SETIAWAN Physical condition is a whole unit from components that cannot be separated off hand , whether the improvement or the maintenance . Problem that proposed in this research is how the physical condition of footbal," pp. 1–5.

[4] P. Alkhairi and A. P. Windarto, "Classification Analysis of Back propagation-Optimized CNN Performance in Image Processing," *J. Syst. Eng. Inf. Technol.*, vol. 2, no. 1, pp. 8–15, 2023.

[5] S. C. Sihombing and A. Dahlia, "Prediction Of Stock Close Price On The Five Best Issuers Forbes Global 2000 Version Using Chen's Fuzzy Time Series Method," *Int. Conf. Bus. Soc. Sci.*, pp. 1163–1171, 2022.

[6] R. K. Tipu, "Enhancing chloride concentration prediction in marine concrete using conjugate gradient-optimized backpropagation neural network," *Asian J. Civ. Eng.*, vol. 25, no. 1, pp. 637–656, 2024, doi: 10.1007/s42107-023-00801-3.

[7] P. Alkhairi, E. R. Batubara, R. Rosnelly, W. Wanayaumini, and H. S. Tambunan, "Effect of Gradient Descent With Momentum Backpropagation Training Function in Detecting Alphabet Letters," *Sinkron*, vol. 8, no. 1, pp. 574–583, 2023, doi: 10.33395/sinkron.v8i1.12183.

[8] P. M. Kurniawan, "Prediction of Civil Servant Performance Allowances Using the Neural Network Backpropagation Method," *Int. J. Informatics Vis.*, vol. 7, no. 3, pp. 673–680, 2023, doi: 10.30630/joiv.7.3.1698.

[9] U. Sharma, "Prediction of the compressive strength of Flyash and GGBS incorporated geopolymer concrete using artificial neural network," *Asian J. Civ. Eng.*, vol. 24, no. 8, pp. 2837–2850, 2023, doi: 10.1007/s42107-023-00678-2.

[10] A. Javeed, S. U. Khan, L. Ali, S. Ali, Y. Imrana, and A. Rahman, "Machine Learning-Based Automated Diagnostic Systems Developed for Heart Failure Prediction Using Different Types of Data Modalities: A Systematic Review and Future Directions," *Comput. Math. Methods Med.*, vol. 2022, 2022, doi: 10.1155/2022/9288452.

[11] Buntoro and G A, "Implementation of a Machine Learning Algorithm for Sentiment Analysis of Indonesia's 2019 Presidential Election," *IIUM Eng. J.*, vol. 22, no. 1, pp. 78–92, 2021, doi: 10.31436/IIUMEJ.V22I1.1532.

[12] S. R. Dani, S. Solikhun, and D. Priyanto, "The Performance Machine Learning Powel-Beale for Predicting Rubber Plant Production in Sumatera," *Int. J. Eng. Comput. Sci. Appl.*, vol. 2, no. 1, pp. 29–38, 2023, doi: 10.30812/ijecsa.v2i1.2420.

[13] A. P. Windarto, T. Herawan, and P. Alkhairi, "Prediction of Kidney Disease Progression Using K-Means Algorithm Approach on Histopathology Data," in *Artificial Intelligence, Data Science and Applications*, Cham: Springer Nature Switzerland, 2024, pp. 492–497. doi: 10.1007/978-3-031-48465-0_66.

[14] H. Wang, "Research on the Application of Random Forest-based Feature Selection Algorithm in Data Mining Experiments," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 10, pp. 505–518, 2023, doi: 10.14569/IJACSA.2023.0141054.

[15] P. Alkhairi, A. P. Windarto, and M. M. Efendi, "Optimasi LSTM Mengurangi Overfitting untuk Klasifikasi Teks Menggunakan Kumpulan Data Ulasan Film Kaggle IMDB," vol. 6, no. 2, pp. 1142–1150, 2024, doi: 10.47065/bits.v6i2.5850.

[16] D. R. K. Kumar Shubham, "Breast Cancer Detection Using Machine Learning Algorithms," *Lect. Notes Networks Syst.*, vol. 624 LNNS, no. 3, pp. 399–406, 2023, doi: 10.1007/978-3-031-25344-7_36.

[17] I. Khajar, H. Hersugondo, and U. Udin, "Comparative study of sharia and conventional stock mutual fund performance: evidence from indonesia," *WSEAS Trans. Bus. Econ.*, vol. 16, pp. 78 – 85, 2019.

[18] L. F. de J. Silva, O. A. C. Cortes, and J. O. B. Diniz, "A novel ensemble CNN model for COVID-19 classification in computerized tomography scans," *Results Control Optim.*, vol. 11, no. September 2022, p. 100215, 2023, doi: 10.1016/j.rico.2023.100215.

[19] C. W. Teoh, S. B. Ho, K. S. Dollmat, and C. H. Tan, "Ensemble-Learning Techniques for Predicting Student Performance on Video-Based Learning," *Int. J. Inf. Educ. Technol.*, vol. 12, no. 8, pp. 741–745, 2022, doi: 10.18178/ijiet.2022.12.8.1679.

[20] S. S. Rani, "An Automated Lion-Butterfly Optimization (LBO) based Stacking Ensemble Learning Classification (SELC) Model for Lung Cancer Detection," *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 3, pp. 87–100, 2023, doi: 10.52866/ijcsm.2023.02.03.008.

[21] A. P. Windarto, T. Herawan, and P. Alkhairi, "Early Detection of Breast Cancer Based on Patient Symptom Data Using Naive Bayes Algorithm on Genomic Data," in *Artificial Intelligence, Data Science and Applications*, Y. Farhaoui, A. Hussain, T. Saba, H. Taherdoost, and A. Verma, Eds., Cham: Springer Nature Switzerland, 2024, pp. 478–484.

[22] P. Venkata, "Data Mining and SVM Based Fault Diagnostic Analysis in Modern Power System Using Time and Frequency Series Parameters Calculated From Full-Cycle Moving Window," *J. Oper. Autom. Power Eng.*, vol. 12, no. 3, pp. 206–214, 2024, doi: 10.22098/JOAPE.2023.10819.1789.

[23] S. N. Bhagat, "Coupling of Rough Set Theory and Predictive Power of SVM Towards Mining of Missing Data," *Int. Res. J. Multidiscip. Scope*, vol. 5, no. 2, pp. 732–744, 2024, doi:

10.47857/irjms.2024.v05i02.0631.

[24] P. Alkhairi, W. Wanayumini, and B. H. Hayadi, "Analysis of the adaptive learning rate and momentum effects on prediction problems in increasing the training time of the backpropagation algorithm," *AIP Conf. Proc.*, vol. 3048, no. 1, p. 20049, 2024, doi: 10.1063/5.0203374.

[25] S. Poonkodi, "A review on lung carcinoma segmentation and classification using CT image based on deep learning," *Int. J. Intell. Syst. Technol. Appl.*, vol. 20, no. 5, pp. 394–413, 2022, doi: 10.1504/IJISTA.2022.125608.

[26] N. Khan, "Enhanced Deep Learning Hybrid Model of CNN Based on Spatial Transformer Network for Facial Expression Recognition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 36, no. 14, 2022, doi: 10.1142/S0218001422520280.

[27] S. Defit, A. P. Windarto, and P. Alkhairi, "Comparative Analysis of Classification Methods in Sentiment Analysis: The Impact of Feature Selection and Ensemble Techniques Optimization," *Telematika*, vol. 17, no. 1, pp. 52–67, 2024.

[28] P. Kaladevi, "An improved ensemble classification-based secure two stage bagging pruning technique for guaranteeing privacy preservation of DNA sequences in electronic health records," *J. Intell. Fuzzy Syst.*, vol. 44, no. 1, pp. 149–166, 2023, doi: 10.3233/JIFS-221615.

[29] S. Nayak, Savita, and Y. K. Sharma, "A modified Bayesian boosting algorithm with weight-guided optimal feature selection for sentiment analysis," *Decis. Anal. J.*, vol. 8, no. July, p. 100289, 2023, doi: 10.1016/j.dajour.2023.100289.

[30] F. E. Botchey, Z. Qin, and K. Hughes-Lartey, "Mobile money fraud prediction-A cross-case analysis on the efficiency of support vector machines, gradient boosted decision trees, and Naïve Bayes algorithms," *Inf.*, vol. 11, no. 8, 2020, doi: 10.3390/INFO11080383.

[31] A. P. Windarto, I. R. Rahadjeng, M. N. H. Siregar, and P. Alkhairi, "Deep Learning to Extract Animal Images With the U-Net Model on the Use of Pet Images," *J. MEDIA Inform. BUDIDARMA*, vol. 8, no. 1, pp. 468–476, 2024.

[32] Q. N. Nguyen, N. C. Debnath, and V. D. Nguyen, "Face Recognition Based on Deep Learning and HSV Color Space BT - Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022," A. E. Hassanien, V. Snášel, M. Tang, T.-W. Sung, and K.-C. Chang, Eds., Cham: Springer International Publishing, 2023, pp. 171–177.

[33] S. V. Kiran, "Machine Learning with Data Science-Enabled Lung Cancer Diagnosis and Classification Using Computed Tomography Images," *Int. J. Image Graph.*, vol. 23, no. 3, 2023, doi: 10.1142/S0219467822400022.

[34] M. Khan *et al.*, "Ensemble and optimization algorithm in support vector machines for classification of wheat genotypes," *Sci. Rep.*, vol. 14, no. 1, p. 22728, 2024, doi: 10.1038/s41598-024-72056-0.

[35] I. B. Santoso, Y. Adrianto, A. D. Sensusiati, D. P. Wulandari, and I. K. E. Purnama, "Ensemble Convolutional Neural Networks With Support Vector Machine for Epilepsy Classification Based on Multi-Sequence of Magnetic Resonance Images," *IEEE Access*, vol. 10, pp. 32034–32048, 2022, doi: 10.1109/ACCESS.2022.3159923.

[36] E. D. Wandekoken, F. M. Varejão, R. Batista, and T. W. Rauber, "Support Vector Machine Ensemble Based on Feature and Hyperparameter Variation for Real-World Machine Fault Diagnosis," in *Soft Computing in Industrial Applications*, A. Gaspar-Cunha, R. Takahashi, G. Schaefer, and L. Costa, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 271–282.

