

Comparative Analysis of 1-Stage and 2-Stage Classification Models Using IndoBERT

Crysantha Monica Lim – 2602090076 – crysantha.lim@binus.ac.id
School of Computer Science, Bina Nusantara University

Abstract

This study explores the performance of **1-stage** and **2-stage classification models** using **IndoBERT** for textual classification in sports data. The models are evaluated on their ability to classify text into subcategories, such as **Liga Inggris, Liga Indonesia, Liga Spanyol, and Liga Italia**, while addressing challenges with minority class performance. Results demonstrate that the **2-stage models**, particularly the second iteration, outperform **1-stage models** in terms of precision, recall, and F1-score, especially for minority classes like **Liga Italia**. The findings provide insights into the benefits of hierarchical classification in imbalanced datasets.

Introduction

This study investigates the performance of two distinct classification approaches, 1-stage and 2-stage hierarchical models, in the context of sports text classification using IndoBERT. The 1-stage model directly classifies text into multiple subcategories, while the 2-stage model utilizes a hierarchical approach where the first stage determines broader categories (e.g., Sepak Bola vs Non-Sepak Bola) before refining the classification in the second stage. The evaluation metrics include accuracy, precision, recall, and F1-score across training, validation, and testing datasets. Results indicate that the 2-stage approach, particularly its second iteration, outperforms the 1-stage model in precision and F1-score, especially for subcategories with smaller representation. This research highlights the trade-offs between simplicity and accuracy, offering insights into model selection for real-world text classification tasks.

Materials

- Dataset:** Sports text data categorized into:
 - Broad categories: Sepak Bola, Non-Sepak Bola.
 - Subcategories: Liga Inggris, Liga Indonesia, Liga Spanyol, Liga Italia.
- Modeling Framework:**
 - IndoBERT pre-trained language model for text representation.
 - Tokenized inputs with max length of 512 tokens, padding, and truncation.
- Evaluation Metrics:** Accuracy, precision, recall, and F1-score across datasets (training, validation, testing).

Methodology

- Model Architecture:**
 - 1-Stage Model:
 - Single-step classification into 5 categories.
 - 2-Stage Model:
 - Stage 1: Classifies into Sepak Bola and Non-Sepak Bola.
 - Stage 2: For Sepak Bola, further classification into Liga subcategories.
- Training Details:**
 - Optimizer: AdamW.
 - Loss Function: CrossEntropyLoss.
 - Batch Size: 2 (Model 1) and 8 (Model 2).
 - Dropout: 0.5 (Model 1) and 0.7 (Model 2).
 - Early Stopping: Patience of 3 epochs for 1-Stage Model and 2-Stage Model 2.
- Evaluation:**
 - Metrics computed for training, validation, and testing datasets.
 - Classification performance analyzed for each category.

Results

The evaluation of the models highlights key differences between the 1-stage and 2-stage approaches. In the 1-stage model, both Model 1 and Model 2 achieve a test accuracy of 95%, indicating strong overall performance. However, for the minority class Liga Italia, Model 2 significantly outperforms Model 1 with an F1-Score of 1.00 compared to 0.67, demonstrating the effectiveness of added regularization through higher dropout. Additionally, Model 2 achieves higher Macro Avg F1-Score (0.93) and Weighted Avg F1-Score (0.97) compared to Model 1, showing improved consistency across all categories.

Metric	Model 1	Model 2
Test Accuracy	84.2%	100%
F1-Score (Liga Italia)	0.50	1.00
Macro Avg F1-Score	0.84	1.00
Weighted Avg F1-Score	0.85	1.00

In the 2-stage model, the first stage, which classifies data into Sepak Bola and Non-Sepak Bola, achieves 100% test accuracy and an F1-Score of 1.00 for both Model 1 and Model 2. This indicates that both models excel in broad classification. In the second stage, where the model further classifies Sepak Bola into subcategories, Model 2 demonstrates superior performance with a test accuracy of 100% compared to Model 1's 84.2%. Model 2 also achieves a perfect F1-Score (1.00) for all subcategories, including the minority class Liga Italia, compared to Model 1's F1-Score of 0.50 for the same class. The Macro Avg F1-Score and Weighted Avg F1-Score for Model 2 in Stage 2 are also perfect (1.00), showcasing its ability to generalize well across all classes.

Metric (Stage 1)	Model 1	Model 2
Test Accuracy	100%	100%
F1-Score	1.00	1.00

Metric (Stage 2)	Model 1	Model 2
Test Accuracy	95%	95%
F1-Score (Liga Italia)	0.67	1.00
Macro Avg F1-Score	0.90	0.93
Weighted Avg F1-Score	0.95	0.97

In summary, while the 1-stage model provides a simpler approach with good performance overall, it struggles with handling minority classes. On the other hand, the 2-stage hierarchical model, particularly Model 2, offers substantial improvements in both overall performance and the handling of underrepresented classes like Liga Italia, making it a superior approach for datasets with complex subcategories.

Conclusion

- The 1-stage model offers a simpler approach with good performance for balanced datasets but struggles with minority class performance, especially Liga Italia (F1-Score 0.67).
- The 2-stage model, particularly its second iteration, demonstrates superior performance, achieving 100% accuracy and F1-Score across all subcategories, including minority classes.
- Hierarchical classification in the 2-stage approach enables better handling of complex datasets with multiple subcategories.

Recommendations

- Use 2-Stage Approach:** The hierarchical structure is recommended for datasets with multiple levels of classification and subcategories.
- Optimize Minority Class Performance:** Further experiments with data augmentation or synthetic data generation for minority classes like Liga Italia may improve robustness.
- Trade-Off Consideration:** The 1-stage model remains a viable option for tasks where simplicity and speed are prioritized over precision in minority classes.

Acknowledgements

Crysantha Monica Lim – 2602090076 – crysantha.lim@binus.ac.id
School of Computer Science, Bina Nusantara University