This observation aligns with the findings in earlier research, such as the study by [4], where KNN demonstrated good performance but was better at identifying stunted cases due to its higher recall compared to other algorithms.

On the other hand, the RF model achieved a higher accuracy of 99.22% with an F1-score of 96.94%. This highlights RF's advantage in classifying data with high accuracy and a better balance between precision and recall. This superior performance can be attributed to RF's ability to capture complex patterns and interactions within the data through its ensemble-based decision tree approach. Random Forest models are particularly well-suited for tasks involving multiple features and complex relationships, which is evident in this study where the data involves a mix of continuous and categorical attributes. This result is consistent with findings in previous studies, such as those by [4] and [7], which reported RF's robust performance in similar stunting prediction tasks. These studies demonstrated that RF outperforms KNN, particularly in datasets that require effective handling of feature interactions, even after addressing class imbalances.

IV. Conclusion

A comparative study of the K-Nearest Neighbors (KNN) and Random Forest (RF) algorithms for predicting stunting in Bekasi Regency demonstrates that machine learning approaches can effectively support early detection of stunting, particularly in cases of Low Birth Weight (LBW) and Low Birth Length (LBL). The KNN model demonstrated stable results across 20 runs, with an accuracy of 96.19% and an F1-score of 87.16%, highlighting consistent performance but limitations in capturing the dataset's complexity. In contrast, the RF model achieved a higher range of accuracy, with values between 98.99% and 99.22%, alongside stronger F1-scores ranging from 96.13% to 96.99%. This superior performance, attributed to RF's ensemble-based decision tree approach, demonstrates its capability to capture intricate patterns and maintain a balanced classification performance, particularly after addressing class imbalance using SMOTE and implementing comprehensive data preprocessing. These findings contribute to the advancement of machine learning applications in healthcare and provide practical tools for practitioners and policymakers to achieve stunting reduction targets. However, future research should consider incorporating additional features, exploring other advanced algorithms, and developing more comprehensive real-time analysis systems to enhance stunting prevention across various demographic groups and geographic regions.

References

[1] S. B. Gaffar, N. N. M. B, and M. Asri, "PKM Pencegahan Stunting melalui Pendidikan Keluarga," in *Seminar National Results of Community Service*, 2021, pp. 22–25.

[2] I. M. Apriliani, N. P. Purba, L. P. Dewanti, H. Herawati, and I. Faizal, "Stunting Risk Factors in Children Under Five in Indonesia: A Scoping Review," *Indonesian Journal of Health Promotion*, vol. 5, no. 6, pp. 654–661, 2022.

[3] M. D. Onis *et al.*, "The World Health Organization's global target for reducing childhood stunting by 2025: Rationale and proposed actions," *Maternal & Child Nutrition*, vol. 9, no. S2, pp. 6–26, September 2013.

[4] I. P. Putri, T. Terttiaavini, and N. Arminarahmah, "Analisis Perbandingan Machine Learning untuk Prediksi Stunting pada Anak," *MALCOM Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 257–265, 2024.

[5] H. H. Sutarno, R. Latuconsina, and A. Dinimaharawati, "Prediksi Stunting Pada Balita Dengan Menggunakan Algoritma Klasifikasi K-Nearest Neighbors," *e-Proceeding Engineering*, vol. 8, no. 5, p. 6657, 2021.

[6] F. Azzahra, N. Suarna, and Y. A. Wijaya, "Penerapan algoritma random forest dan cross validation untuk prediksi data stunting," *KOPERTIP: Scientific Journal of Informatics Management and Computer*, vol. 8, no. 1, pp. 1–6, 2024.

[7] M. G. Daffa and P. H. Gunawan, "Stunting classification analysis for toddlers in Bojongsoang: A data-driven approach," in *2024 2nd International Conference on Software Engineering and Information Technology (ICoSEIT)*, Feb 2024, pp. 42–46.

[8] R. Devika, S. V. Avilala, and V. Subramaniyaswamy, "Comparative study of classifier for chronic kidney disease prediction using naive bayes, KNN and random forest," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, March 2019, pp. 679–684.

[9] A. T. A. Sibuea and P. H. Gunawan, "Classifying stunting status in toddlers using k-nearest neighbor and logistic regression analysis," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, July 2024, pp. 6–11.

[10] J. B. Chandra and D. Nasien, "Application of Machine Learning K-Nearest Neighbour Algorithm to Predict Diabetes," *International Journal of Electrical, Energy and Power System Engineering*, vol. 6, no. 2, pp. 134–139, 2023.

[11] A. Almomany, W. R. Ayyad, and A. Jarrah, "Optimized implementation of an improved KNN classification algorithm using Intel FPGA platform: Covid-19 case study," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3815–3827, 2022.

[12] A. B. Adetunji, O. N. Akande, F. A. Ajala, O. Oyewo, Y. F. Akande, and G. Oluwadara, "House price prediction using random forest machine learning technique," *Procedia Computer Science*, vol. 199, pp. 806–813, 2022.

[13] C. Fannany, P. H. Gunawan, and N. Aquarini, "Machine Learning Classification Analysis for Proactive Prevention of Child Stunting in Bojongsoang: A Comparative Study," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, July 2024, pp. 1–5.

[14] M. Pal and S. Parija, "Prediction of heart diseases using random forest," *Journal of Physics: Conference Series*, vol. 1817, no. 1, p. 012009, March 2021.

[15] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," *The Journal of Supercomputing*, vol. 77, no. 5, pp. 5198–5219, 2021.

[16] H. Janawisuta and P. H. Gunawan, "Early detection of stunting in indonesian toddlers: A machine learning approach," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, July 2024, pp. 12–16.

[17] H. Guo, H. Nguyen, D.-A. Vu, and X.-N. Bui, "Forecasting mining capital cost for open-pit mining projects based on artificial neural network approach," *Resources Policy*, vol. 74, p. 101474, 2021.

[18] W. I. Rahayu, C. Prianto, and E. A. Novia, "Perbandingan Algoritma K-Means dan Naïve Bayes untuk Memprediksi Prioritas Pembayaran Tagihan Rumah Sakit Berdasarkan Tingkat Kepentingan Pada PT. Pertamina (Persero)," *Jurnal Teknik Informatika*, vol. 13, no. 2, pp. 1–8, 2021.

[19] G. A. F. Khansa and P. H. Gunawan, "Predicting Stunting in Toddlers Using KNN and Naïve Bayes Methods," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, July 2024, pp. 17–21.