ASEC 2025 Conference
The 6th International Electronic Conference on Applied Sciences
09–11 December 2025 | Online
MDPI
sciforum-141468

Machine Learning-Based Hybrid Model for Improved Crop Yield Correlation Analysis: A Data-Driven Assessment

Nasina Harshavardhan, Yamini Kodali, Yellapragada Venkata Pavan Kumar
Corresponding Author: Yellapragada Venkata Pavan Kumar (Email ID:pavankumar.yv@vitap.ac.in)

INTRODUCTION & AIM

- Climate change greatly affects farming by reducing crop yields, thereby challenging sustainable farming.
- Real-time data and analytics help identify key factors to boost crop yield and support smart farming through better decision making.
- Since climate factors often change in complex, nonlinear ways, this paper suggests a path analysis approach that combines the direct and indirect effects for better analysis.
- This study aims to build correlation inbetween the determinants through path analysis.

METHOD

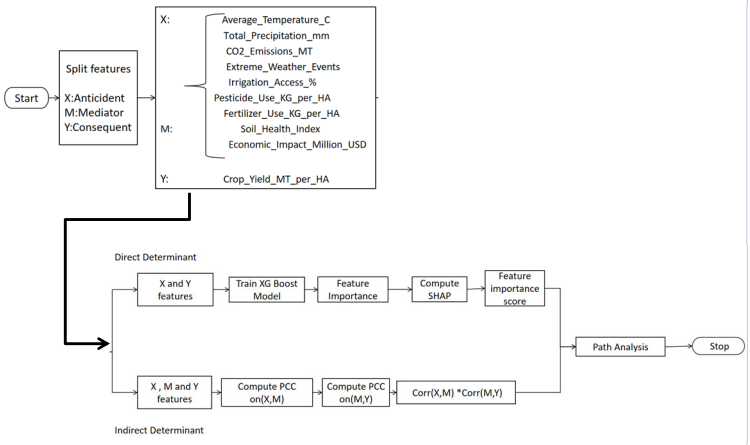

Fig. 1. Conceptual workflow for Crop Yield Correlation Analysis

RESULTS & DISCUSSION

- The target determinant Crop_Yield_MT_per_HA is considered as consequent (Y).
- The XGBoost feature importance was performed and found as Economic_Impact_Million_USD has the highest impact of 50% on the target determinant Crop_Yield_MT_per_HA which reduces the loss. The feature importance score was computed using SHAP as direct effect.
- For Eg, Pearson Correlation Coefficient Value of X and M CO2_Emissions_MT and Economic_Impact_Million_USD(0.061) and the Pearson Correlation Coefficient Value of M and Y Economic_Impact_Million_USD and Crop_Yield_MT_per_HA (-0.041)was calculated as Indirect Effect.

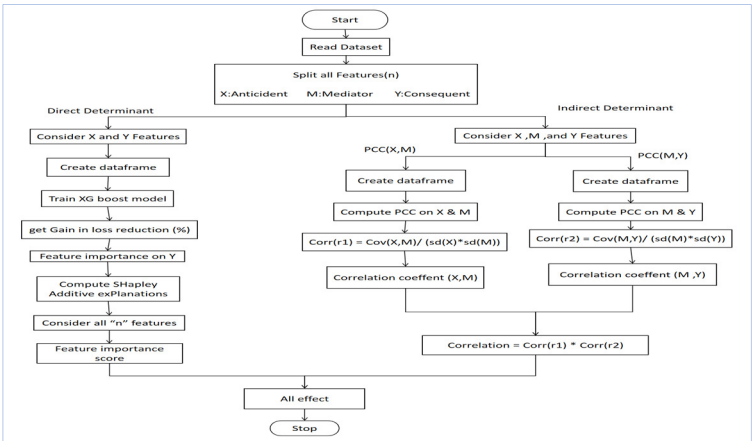

Fig. 2. Implementation workflow for Crop Yield Correlation Analysis

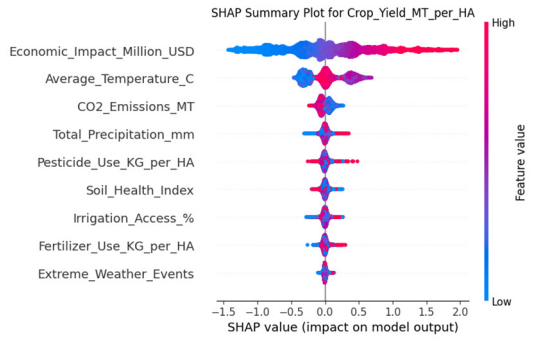

Fig. 3. Feature Importance Values using SHAP

CONCLUSION AND FUTURE SCOPE

- ✓ The correlation of crop yield was largely explained by its high direct and indirect effect on yield, as revealed by path analysis, indicating that, the trait Crop_Yield_MT_per_HA is a key determinant for improving yield.
- ✓ Future work will focus further on regression analysis and partial correlation which enhance more insights on crop yields.

KEY REFERENCES

[1] Y. Gao et al., "A Data-driven Correlation Analysis Method for Cities Carbon Emissions with Power System," in 2021 Smart City Challenges & Outcomes for Urban Transformation (SCOUT), Bhubaneswar, India: IEEE, Dec. 2021, pp. 209–213. doi: https://doi.org/10.1109/SCOUT54618.2021.00051.

[2] M. Zhou, H. Liu, and Z. Wang, "Can Smart City Construction Promote the Level of Public Services? Quantitative Evidence From China," IEEE Access, vol. 10, pp. 120923–120935, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3221759.

[3]A. Youssef and P. Hajek, " in 2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC), Rome, Italy: IEEE, Nov. 2021, pp. 276–279. doi: https://doi.org/10.1109/ISCSIC54682.2021.00057.