

This paper makes numerous major advances to the field of medical artificial intelligence diagnostics by presenting a complete multi-model ensemble system for early cancer diagnosis utilizing cytological data analysis: Technical creativity: To examine 10 different cytological features, the system effectively combines three complementary machine learning methods: logistic regression, decision trees, and XGBoost. Long-term clinical usefulness and regulatory compliance depend on building systems with natural flexibility and validation capacity. The preprocessing system addresses real-world clinical data challenges by sophisticated handling of missing values using median imputation for the BareNuc feature, outlier detection especially tailored for medical data where extreme values may represent genuine pathogenic findings, and comprehensive audit trails for all data manipulations to ensure clinical quality assurance. This multi-model architecture makes use of the interpretability of logistic regression coefficients reflecting pathologists' diagnostic reasoning, the hierarchical decision-making structure of decision trees translating directly into clinical protocols, and the advanced feature interaction capabilities of XGBoost for complex pattern recognition. Medical data compliance standards, including HIPAA and GDPR requirements, conform with FDA guidelines for AI/ML-based Software as Medical Device (SaMD) and follow accepted healthcare data exchange standards, including HL7 FHIR protocols are included in the system framework. The system runs effectively on normal clinical workstation hardware without requiring specialist infrastructure; the user interface mirrors genuine pathology laboratory procedures, input forms mirror common medical reports, and the system operates as such. Multi-Model Approach Validation: The ensemble technique demonstrated better than single-model approaches not only in diagnostic accuracy but also in offering educational value for comprehending many elements of cellular morphology analysis. Performance Achievement in Clinical Settings: With a goal true positive rate of at least 95% for early cancer detection, the system delivers extraordinary diagnostic accuracy, hence greatly lowering the danger of missed diagnosis. Considerations of Resources: Constraints Designing for real-world clinical situations calls for great attention to computing resource restrictions, current infrastructural capabilities, and budgetary restraints.