Predicting the Risk of Post-Chikungunya Arthritis using Hybrid Ensemble Model

Md Wakil Ahmed
Department of CSE
Daffodil International University
Dhaka, Bangladesh
ahmedmdwakil512@gmail.com

Khushnor Rahman Meem
Department of CSE
Daffodil International University
Dhaka, Bangladesh
khushnorrahmanmeem@gmail.com

Md Shahnewaj Limon
Department of CSE
Daffodil International University
Dhaka, Bangladesh
mdshahnawajlimon@gmail.com

Shahrin Khan
Lecturer, Department of CSE
Daffodil International University
Dhaka, Bangladesh
shahrinkhan.cse@diu.edu.bd

***Abstract*—A mosquito-borne arboviral disease, chikungunya virus infection can cause long-term inflammatory rheumatism, including post-Chikungunya arthritis that develops in a large portion of patients and may lead to chronic disability. The gap is filled in this study as it constructs a hybrid ensemble model of XGBoost and Logistic Regression, which are trained on a combined dataset of 1001 real samples on Kaggle and 1129 synthetically generated samples on the basis of clinical studies. Using data preprocessing, feature engineering based on polynomial transformations to exploit non-linear relationships, class balancing through SMOTE, and hyperparameter optimization, the model obtained a 0.73 macro F1-score and 0.76 AUC on the combined dataset, or a 43% accuracy improvement over the baseline accuracy of 0.51 on the original dataset. Sensitivity analysis helped to confirm that the model is resistant to noise and fluctuations. The findings validate the importance of bigger and real-world data to enhance the model's predicting capacity in such diseases. Lack of biased synthetic data can be a limitation, and in the future, deep learning and interpretability tools such as SHAP can be applied for clinical validation.**

***Index Terms*—feature engineering, hybrid ensemble model, chikungunya, arthritis**

I. Introduction

Chikungunya is an arbovirus disease caused by the Chikungunya virus (CHIKV), with both the virus and the disease sharing the name "Chikungunya." The disease was first identified during an outbreak in Tanzania in 1952-1953 as a public health threat [1]. The term "Chikungunya" originates from the Kimakonde language spoken in southern Tanzania which translates to "that which bends up" or "to become contorted," reflecting the bended up posture of affected individuals due to severe joint pain [2].

Historically it was a tropical disease primarily affecting South Asia and Afric. Chikungunya spread globally mainly through travellers and military personnel returning from those endemic areas [3]. This spread has transformed it from a tropical illness into a worldwide concern reporting outbreaks reported in Europe, the Americas, and beyond. The virus is transmitted primarily by the Aedes aegypti and Aedes albopictus mosquitoes, which also serve as vectors for dengue virus. These two diseases often raise confusions due to overlapping clinical presentations [4].

During epidemics, human beings serve as the key reservoir of CHIKV, whereas in inter-epidemic time periods, the virus is retained in sylvatic cycles of non-human primates, rodents, birds, and other vertebrates in Africa [5]. Despite the similarities in the signs and symptoms of Chikungunya and dengue (high fever, severe headache, myalgia, and arthralgia), there are differences in the severity of the symptoms. As an example, a study conducted in Thailand studied the difference that the duration of fever might be reduced in Chikungunya but symptoms such as rash, conjunctival injection, and arthralgia are more severe than those observed in dengue [6]. Clinically, chikungunya is known as Chikungunya fever (CHIKF), an acute febrile disease that in most cases advances to chronic stages [7].

Of particular concern is the emergence of chronic inflammatory rheumatism (CIR) or polyarthritis that may cause joint destruction and permanent disability. Certain patients develop symptoms of rheumatoid arthritis (RA) or spondyloarthritis (SA), with a chronic synovitis and the presence of joint effusion [8] . Chronic symptoms occur in about 25-40% of infected people mainly of myalgia and severe arthralgia that can develop into arthritis [9]. In some areas, including the Americas, it has been documented that a high percentage(25-60%) of those affected experience post-Chikungunya arthritis of more than 12 weeks duration [10]. Older adults are particularly at risk of developing these chronic symptoms, and age above 40 is often characterized as a predictive factor because of other factors as well. [11].

Irrespective of these insights, the connection between the severity of acute symptoms and the development of post-Chikungunya arthritis has not been studied comprehensively. This gap prevents early diagnosis and treatment, especially in

resource-constrainted fields where long term follow ups are difficult. This research is expected to fill in these gaps by introducing a hybrid ensemble model of post-Chikungunya arthritis risk prediction on the basis of symptomatic and demographic characteristics. This work aims to improve the predictive accuracy and facilitate clinical decision-making to reduce the adverse effects of chronic conditions by applying advanced feature engineering techniques.

II. Related Works

Lazari et al. [9] in Brazil tracked 90 days of symptomatic CHIKV-confirmed patients and established that 45.3% of them fulfilled the criteria of chronic disease, and women were more likely to progress to the chronic disease. The article identifies the prognostic role of subacute symptoms and justifies the use of objective clinical and imaging outcomes to facilitate early detection and personalized treatment to limit the joint disability in the long term following CHIKV infection.

The study by Hossain et al. [12]is a longitudinal study carried out in a tertiary care hospital in Bangladesh after 143 Chikungunya virus-infected patients had been treated in the facility in 2017. They discovered that 41.9 percent of patients showed chronic inflammatory rheumatism in the course of a year, with the onset mainly as undifferentiated arthritis that exhibited polyarthralgia (55 percent) and polyarthritis (33.3 percent). This paper identifies the importance of CHIKV infection as a key long-term inflammatory factor in Bangladesh, which requires early clinical surveillance and risk evaluation to inform long-term management and rehabilitation of affected individuals.

Lozano-Parra et al. [13] conducted a nested case-control study in Colombia to determine acute immunological biomarkers to predict chronic rheumatologic disease following Chikungunya virus infection. They documented serum cytokine and chemokine concentrations, such as IL-4, IL-6, IL-8, IL-27, CCL-2, CXCL-9 and CXCL-10, at the acute and subacute phases in 46 confirmed CHIKV patients. They found that increased initial IL-8, CXCL-9 and CXCL-10 were strongly linked with a reduced likelihood of chronic post-Chikungunya rheumatism, indicating that an intense initial immune reaction could reduce chronic joint symptoms. This research contributes to the body of knowledge of immunological processes of chronic CHIKV arthritis and contributes to the possibility of using these biomarkers as an early warning of risk and further intervention to mitigate the long-term morbidity.

Chang et al. [14] carried out a prospective cohort study in order to establish a Chikungunya Arthritis Disease Activity Score (CHIK-DAS). Their work shows that the shift between acute infection and chronic arthritis correlates with immune deregulation, namely, the decrease of protective Th2 cytokines including IL-4 and IL-13, and regulatory cytokines including IL-2. Conversely. The article introduces the importance of CD4+ T cells in the inflammation of joints and indicates that early immunomodulation with cytokines may be a preventative intervention in chronic joint pain.

In a study done by Huits et al. [15] in Aruba, they sampled 498 symptomatic adults, of whom 269 (54%) had been confirmed to have CHIKV infection with a combination of IgM serology and RT-PCR. 44 percent of patients had developed chronic polyarthralgia, and 26 percent had one year or longer symptoms. There were acute presentations of arthralgia, fever, and rash, and inflammatory manifestations of edema and morning stiffness were found in 71% of arthralgia. Independent predictors of chronicity were found to be female gender, obesity, increased scores of joint involvement, and long-term viral RNA detection. The paper has underlined the necessity of simultaneous serologic and molecular testing to make correct diagnoses and early detection of patients who are at risk to proceed with specific intervention.

Pollett et al. [16] used a large cohort study based on the 195 confirmed CHIKV cases that occurred in the U.S. Military Health System in the period of 2005-2018. Their result revealed that 32.3% of infected individuals developed post-infection rheumatologic disorders, as compared to 20 in matched controls. Osteoarthritis, rheumatoid arthritis, and inflammatory polyarthropathy were the most common, and risk was dependent on demographic factors and comorbidities. Even though the cases were scarce, the study underlines the necessity to monitor them in the long-term and noted that vaccines may be used by high-risk populations like military staff and frequent travelers.

Blettery et al. [17] published on chronic management of post-chikungunya rheumatic disease in Martinique, which involved 147 cases of the outbreak in 2013-2015. They identified a median delay to rheumatology consultation of 8 months following infection with four most common outcomes: mechanical pain reactivation (32%), fibromyalgia (6.1%), spondyloarthritis flare (30.6%), and new-onset bilateral symmetric chronic inflammatory arthritis (18.4%). Methotrexate has become the most important therapeutic agent with positive outcomes, and in some severe cases, anti-TNF agents were needed. The research also emphasized the diversity of so-called chronic chikungunya syndrome and emphasized the significance of early specialist attention and universal approaches to treatment with the aim of minimizing long-term loss.

III. Proposed Methodology

The approach integrates data from a Kaggle dataset [18] (1001 samples) with synthetically generated data (1129 samples) to create a merged dataset (2130 samples), enabling robust model training. Each subsection discusses techniques and improvements observed when using the merged dataset compared to the original.

The whole process of predicting the risk of post-Chikungunya Arthritis is divided into a few steps (Fig 1) while exploring different approaches and techniques. The following diagram shows the basic steps [3]

A. Data Acquisition

The primary dataset was collected from Kaggle [18], consisting of 1001 samples and 18 features. The features are

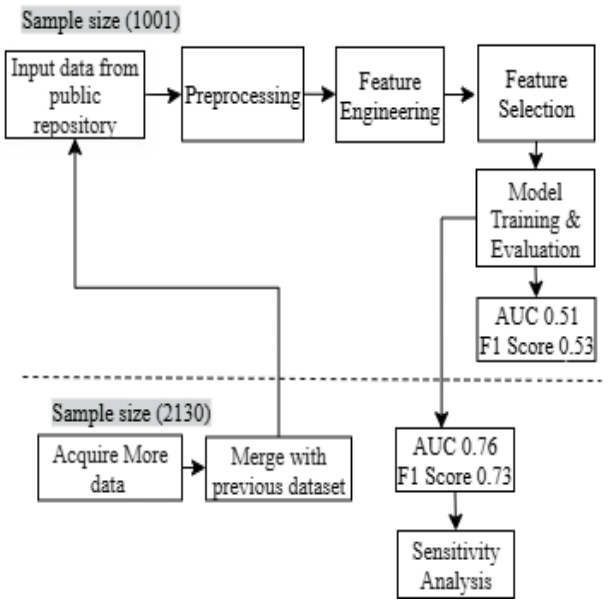

Fig. 1. The Workflow Diagram of Methodology

primarily binary symptoms and demographic variables. Another dataset was retrieved from a public repository of a summarized clinical study [19]. From this, a synthetic dataset of 1129 samples was generated using probability distribution [20]. Merging these two publicly available datasets (Fig 2) mitigates the issues of class imbalance and data scarcity. The combined total of 2130 samples significantly improved model performance. The macro F1 score increased to 0.73 from 0.51 (a 40% improvement) which supports the suitability of the chosen model.

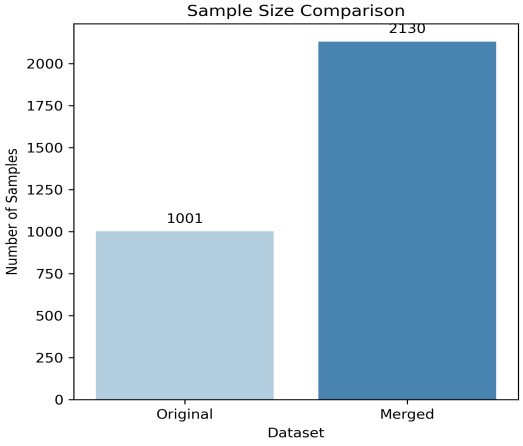

Fig. 2. Sample size of Two Datasets

B. Basic Preprocessing

The redundant columns ('Severe Chikungunya', 'Unnamed') were dropped initially to keep the dataset clean. The binary variables were encoded as yes = 1, no = 0, male = 1, female = 0, and the categorical variables were encoded using LabelEncoder [21]. Some tests like Feature Means, Feature Distribution, Target Distribution were done on the merged dataset to validate the newly merged dataset. The required preprocessing (encoding, filling missing values with appropriate values, dropping irrelevant columns) were done once again. Basic preprocessing couldn't improve the imbalanced classes of original dataset (59% arthritis) which caused the biases in the model training. Fig 3 shows class distribution for both dataset.

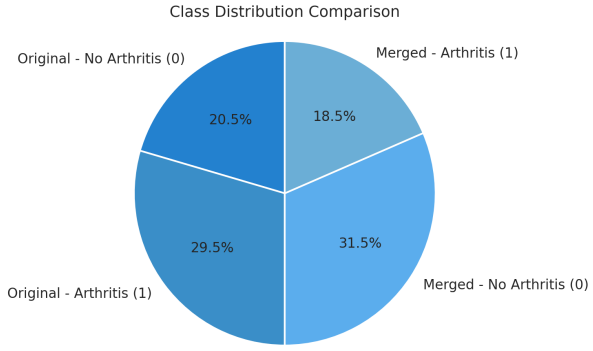

Fig. 3. Class Distribution of Target Variables

C. Feature Engineering

The significance level of variables was very weak against the target variable (arthritis). To unfold the hidden relationships between specific symptoms, new features were created. 'key_symptom_count' represented the summed influence of swelling, vomiting, and joint pains [9]. Additional features ('Swelling_joint_pains', 'joint_stiffness_pain_score', 'pain_score_duration', and 'joint_stiffness_duration') were designed (Fig 4) to directly capture the interactions of certain symptoms, which can serve as strong indicators for clinical studies [15].

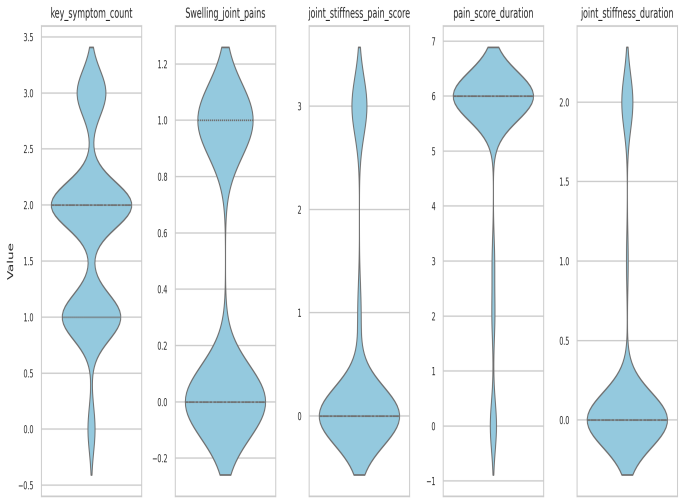

Fig. 4. Violin Plot for the Engineered Features

To capture the non-linear relationships among features new variables were generated using polynomial features of degree 2 [21]. Combining StandardScaler, this helped the linear model to learn more effectively from the training dataset. This technique contributed significantly to improve model performance and mitigate the overfitting problem caused by imbalanced data [20]. Original data engineering improved macro F1 from 0.46 (baseline) to 0.51. With merged data, polynomial features boosted this to 0.73 by capturing non-linear relationships (e.g., pain duration amplifying stiffness effects), demonstrating 43% improvement due to the dataset's scale allowing better interaction modeling [22].

D. Feature Selection

The dataset was divided into training and testing sets (80% and 20%) using the train_test_split function of scikit-learn with stratification to maintain the distribution of the arthritis classes and a fixed random seed for reproducibility (42) [21]. This gave a total of 804 training and 201 testing samples for the original dataset and 1704 training and 426 testing samples for the merged dataset. SMOTE analysis helped to reduce dataset bias by increasing the minority class of the target variable in the training dataset [20]. This improved both the sensitivity and specificity of the model. Significant features were also selected (Fig 5) using the K-Nearest Neighbor algorithm for training [21].

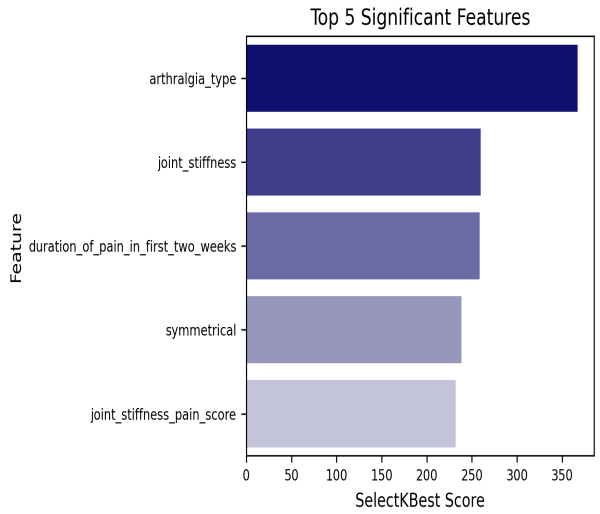

Fig. 5. Top 5 Significant Features for Arthritis=1

RandomUnderSampler was additionally experimented during training, though it showed poor results. A 5-fold Stratified Cross-Validation was also performed to validate the effectiveness of used class-balancing techniques [22].

E. Model Training and Evaluation

Both the original and merged datasets were used to train a hybrid ensemble model which was selected as the baseline model. To ensure robustness, the ensemble model comprises 70% XGBoost which captured non-linearity [23] and 30% Logistic Regression which provided interpretability [24]. This approach consistently gave the best performance with polynomial features. Other techniques such as undersampling, cost-sensitive learning (class weights 0:3.0, 1:1.0), and bagging with RandomForest and AdaBoost were also experimented with but all of them underperformed [25].

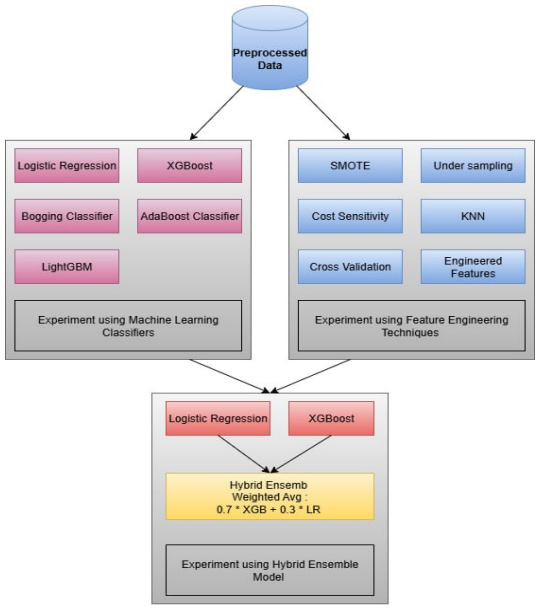

Fig. 6. Visualization of Model Training

Table I shows the value for different evaluation metrices in the mentioned methods for both dataset. Training on the original dataset achieved a maximum Macro F1 score of 0.56, whereas the merged dataset reached 0.78 both using the hybrid ensemble model. This clearly indicates the significance of larger dataset for the generalization of model [22].

TABLE I
MODEL EVALUATION SUMMARY

Model	Sample Size 1001				Sample Size 2130			
	Macro F1	AUC	CV Mean F1	CV Std F1	Macro F1	AUC	CV Mean F1	CV Std F1
Hybrid Ensemble	0.51	0.536	0.541	0.015	0.73	0.761	0.72	0.03
Hybrid Ensemble (Polynomial Features)	0.51	0.531	0.541	0.015	0.728	0.75	0.71	0.029
Hybrid Ensemble (Undersampling)	0.515	0.532	0.544	0.023	0.716	0.752	0.717	0.02
Hybrid Ensemble (Cost-Sensitive)	0.51	0.536	0.333	0.001	0.721	0.746	0.517	0.068
LightGBM	0.51	0.536	0.535	0.023	0.71	0.754	0.712	0.015
Bagging Classifier	0.457	0.536	0.535	0.023	0.721	0.752	0.711	0.031
AdaBoost Classifier	0.51	0.536	0.535	0.023	0.69	0.739	0.681	0.032

F. Sensitivity Analysis

To ensure model robustness, an assessment was conducted by varying the hyperparameters of the hybrid ensemble model and introducing different noise levels (5–10%). The Polynomial Hybrid demonstrated promising performance with a stable Macro F1 score (0.71–0.73) and an AUC drop of less than

5% across 10 runs with different noise levels, outperforming the baseline Hybrid Ensemble Model. This robustness resulted from XGBoost's tolerance to outliers [23] as well as from the non-linearity captured by the polynomial features. The merged dataset improved the stability (std = 0.02 to 0.03) which validates the scalability of the models to larger datasets. Below Table II demonstrates the result of the assessment:

TABLE II
SENSITIVITY ANALYSIS OF HYBRID ENSEMBLE MODEL

Noise Level	Macro F1	Mean F1	Std F1
±5%	0.728633	0.727906	0.001258
±10%	0.726453	0.727906	0.001258
±15%	0.728633	0.727906	0.001258

IV. RESULT ANALSIS

Macro F1-score and AUC were chosen as primary metrics because they are robust to imbalanced medical datasets where a false negative (failure to diagnose arthritis) event is critical. The evaluation metrics are suitable for medical prediction since accuracy alone is misleading (e.g. majority class), F1 is balanced around minority recall (arthritis risk), and AUC measures overall ranking [22]. The confusion matrices (Fig 7) illustrate the predictive capacity of the chosen model. Due to data scarcity, the original training set (1001 samples)

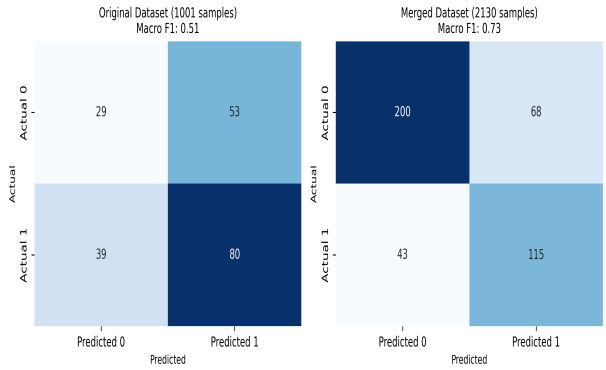

Fig. 7. Performance of model in Confusion Metrix

achieved a macro F1 score of 0.51 and 'Area Under the ROC Curve' (AUC) score of 0.537 (Fig 8). The hybrid ensemble performance improved to a macro F1-score of 0.73 and an AUC score of 0.750 when combined with synthetic data (2130 samples), showing that using larger datasets improves pattern capture and generalization by 43% F1 [20]. The F1 optimization threshold (0.3-0.7) improved the hybrid ensemble's performance, and Polynomial features further enhanced the hybrid ensemble by modelling non-linear symptom interactions [21]. In the below Table III shows the baseline model performance for both dataset: Model stability was evaluated by stratified 5-fold cross-validated sensitivity analysis (CV Std: 0.02-0.07) for medical decision support, which is essential in medical prediction . Threshold tuning (0.3-0.7) optimized F1 to some

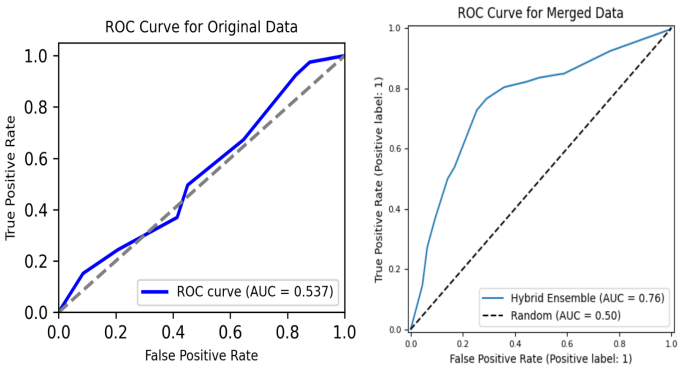

Fig. 8. ROC Curve for the chosen model

TABLE III
RESULTS FROM BASELINE MODEL

Sample	Hybrid Ensemble Model				
	Macro F1	AUC	CV Mean F1	CV Std F1	Accuracy
Original Dataset	0.51	0.536	0.541	0.015	0.54
Merged Dataset	0.73	0.761	0.72	0.03	0.74

extent. These results confirm the proposed method for arthritis prediction in Chikungunya patients.

V. LIMITATIONS

Due to the weak correlations found in the original dataset, model generalizability was limited and may have been caused by incomplete symptom capture [9]. The synthetic data may introduce biases if real-world distributions vary even though it was probabilistically aligned [20]. Sensitivity analysis showed low variance, but effects of increasing noise could show more as assumptions on noise may not reflect real heterogeneity [22]. The studied binary or categorical features do not represent all types of possible biomarkers which may include continuous variables [13]. Small sample sizes (1001 original, 1129 synthetic) limited robustness and combining presumed feature correspondence without external validation raises ethical questions in clinical use [1]. For the larger dataset, computation of polynomial features and ensemble may require high-constrained settings [23].

VI. FUTURE WORK

These tests provide evidence of strong potential for such large-scale multi-centers to further develop and test the technique on larger datasets. Future directions include the use of deep learning methods for automatically extracted features and the conduct of prospective clinical trials to evaluate real-time prediction of arthritis risk. In addition, interpretability techniques (e.g. SHAP [26]) will be studied to identify feature importance and improve the reliability of medical decision-making.

VII. CONCLUSION

This study illustrates a successful approach to predicting arthritis risk with Chikungunya through a hybrid ensemble with polynomial features with macro F1=0.73 and AUC=0.750

on merged data, a significant improvement compared with the baselines on original data. This increased model performance resulted from improved class balancing, feature engineering and model generalization in the merged dataset (2130 samples). This strategy (70% XGBoost + 30% Logistic Regression) would potentially enable early intervention with reduced chronic complications with larger datasets. Real world validation is what should be aimed at in the future.

REFERENCES

[1] H. Zeller, W. Van Bortel, and B. Sudre, "Chikungunya: its history in africa and asia and its spread to new regions in 2013–2014," *The Journal of infectious diseases*, vol. 214, no. suppl_5, pp. S436–S440, 2016.

[2] M. C. Robinson, "An epidemic of virus disease in southern province, tanganyika territory, in 1952–1953," *Transactions of the royal society of tropical medicine and hygiene*, vol. 49, no. 1, pp. 28–32, 1955.

[3] S. C. Weaver and M. Lecuit, "Chikungunya virus and the global spread of a mosquito-borne disease," *New England Journal of Medicine*, vol. 372, no. 13, pp. 1231–1239, 2015.

[4] E. S. Paixao, L. C. Rodrigues, M. d. C. N. Costa, M. Itaparica, F. Barreto, P. Gerardin, and M. G. Teixeira, "Chikungunya chronic disease: a systematic review and meta-analysis," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 112, no. 7, pp. 301–316, 2018.

[5] M. Diallo, J. Thonnon, M. Traore-Lamizana, and D. Fontenille, "Vectors of chikungunya virus in senegal: current data and transmission cycles.," *The American journal of tropical medicine and hygiene*, vol. 60, no. 2, pp. 281–286, 1999.

[6] S. Nimmannitya, S. B. Halstead, S. N. Cohen, and M. R. Margiotta, "Dengue and chikungunya virus infection in man in thailand, 1962-64. i. observations on hospitalized patients with hemorrhagic fever.," 1969.

[7] G. Pialoux, B.-A. Gaüzère, S. Jauréguiberry, and M. Strobel, "Chikungunya, an epidemic arbovirosis," *The Lancet infectious diseases*, vol. 7, no. 5, pp. 319–327, 2007.

[8] A. Zaid, P. Gérardin, A. Taylor, H. Mostafavi, D. Malvy, and S. Mahalingam, "Chikungunya arthritis: Implications of acute and chronic inflammation mechanisms on disease management," *Arthritis Rheumatology*, vol. 70, no. 4, pp. 484–495, 2018.

[9] C. d. S. Lazari, M. S. Ramundo, F. Ten-Caten, C. S. Bressan, A. M. B. de Filippis, E. R. Manuli, I. de Moraes, G. M. Pereira, M. F. Côrtes, D. d. S. Candido, *et al.*, "Clinical markers of post-chikungunya chronic inflammatory joint disease: A brazilian cohort," *PLoS Neglected Tropical Diseases*, vol. 17, no. 1, p. e0011037, 2023.

[10] J. J. Miner, H. X. Aw Yeang, J. M. Fox, S. Taffner, O. N. Malkova, S. T. Oh, A. H. Kim, M. S. Diamond, D. J. Lenschow, and W. M. Yokoyama, "Brief report: chikungunya viral arthritis in the united states: a mimic of seronegative rheumatoid arthritis," *Arthritis & rheumatology*, vol. 67, no. 5, pp. 1214–1220, 2015.

[11] K. A. Sacco and R. M. Chirila, "Postchikungunya chronic inflammatory rheumatism," *Case Reports in Rheumatology*, vol. 2016, no. 1, p. 7068901, 2016.

[12] S. Hossain, M. R. Choudhury, M. A. Islam, M. M. Hassan, S. Yeasmin, F. Hossain, and M. M. Zaman, "Post-chikungunya arthritis: a longitudinal study in a tertiary care hospital in bangladesh," *Tropical Medicine and Health*, vol. 50, no. 1, p. 21, 2022.

[13] A. Lozano-Parra, V. Herrera, L. Á. Villar, S. Urcuqui-Inchima, J. F. Valdés-López, and E. M. R. Garrido, "Acute immunological biomarkers for predicting chronic rheumatologic disease after chikungunya virus infection," *Tropical Medicine and Infectious Disease*, vol. 10, no. 7, p. 195, 2025.

[14] A. Y. Chang, S. Simmens, H. Watson, R. L. Amdur, A. Siqueira, A. Proctor, S. Tritsch, C. A. H. Gomez, L. Encinales, A. S. Hernández, *et al.*, "Development of a chikungunya arthritis disease activity score (chik-das) based on a prospective cohort study," *Journal of cellular immunology*, vol. 6, no. 6, p. 236, 2024.

[15] R. Huits, J. De Kort, R. Van Den Berg, L. Chong, A. Tsoumanis, K. Eggermont, K. Bartholomeeusen, K. K. Ariën, J. Jacobs, M. Van Esbroeck, *et al.*, "Chikungunya virus infection in aruba: Diagnosis, clinical features and predictors of post-chikungunya chronic polyarthralgia," *PloS one*, vol. 13, no. 4, p. e0196630, 2018.

[16] S. Pollett, H.-C. Hsieh, D. Lu, M. Grance, S. Richard, G. Nowak, C. Lanteri, D. Tribble, and T. Burgess, "The risk and risk factors of chikungunya virus infection and rheumatological sequelae in a cohort of us military health system beneficiaries: implications for the vaccine era," *PLOS Neglected Tropical Diseases*, vol. 18, no. 8, p. e0011810, 2024.

[17] M. Blettery, L. Brunier, K. Polomat, F. Moinet, C. Deligny, S. Arfi, G. Jean-Baptiste, and M. De Bandt, "Brief report: management of chronic post-chikungunya rheumatic disease: the martinican experience," *Arthritis & Rheumatology*, vol. 68, no. 11, pp. 2817–2824, 2016.

[18] R. Bernat, "Chikungunya symptom data," 2017. Accessed: July. 24, 2025.

[19] M. S. Hossain, M. M. Hasan, M. S. Islam, S. Islam, M. Mozaffor, M. A. S. Khan, N. Ahmed, W. Akhtar, S. Chowdhury, S. M. Y. Arafat, M. A. Khaleque, Z. J. Khan, T. F. Dipta, S. M. Z. H. Asna, M. A. Hossain, K. S. Aziz, A. A. Mosabbir, and E. Raheem, "Arthralgia profile of chikungunya patients (n = 1129, cases with past history of arthritis were excluded) in bangladesh," 2018. Accessed: Aug. 29, 2025.

[20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[22] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.

[23] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.

[24] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013.

[25] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[26] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.