

## ABSTRACT

Human Papilloma Virus can cause infections of the skin and mucous membranes. Infected skin can cause mild lumps with a rough shape called warts. Warts on the feet will provide discomfort to patients so that treatment is needed as soon as possible. At this time doctors have not identified which treatment is more appropriate. In addition, the treatment provided is still based on personal experience and medical knowledge possessed. Therefore, ML technology is applied to predict the wart treatment appropriately. ML technology that has been developed by researchers still requires better accuracy so that success in predicting treatment is more accurate. However, the limited number of samples and data that have imbalance class problems make it difficult to improve. This study aims to improve the accuracy of previous studies.

In an effort to realize the objectives of this study, several algorithms need to be proposed. First, feature selection is applied to choose which features can influence increased accuracy. The method applied for feature selection is called mutual information. This method can provide a value for each feature with a range of 0 to 1. It can provide justification for choosing which features are more relevant. Second, the algorithm applied to classify data is a combination of adaboost and random forest. Random forest algorithm can be used as a weak learner to classify cryotherapy and immunotherapy data. Then, adaboost is used to encourage the performance of weak learners to achieve even better accuracy. Both methods are suggested to have good abilities in imbalance class problems. Cryotherapy and immunotherapy data need to be randomized before that are induced into the algorithm. Cross-validation method is applied to shuffle data so that machine learning algorithms can learn various patterns.

The experiment results have shown the accuracy of cryotherapy and immunotherapy is 96.6 % and 91.1 %. This results can provide as a reference in further development and design of the system. Next, the system can help doctors to choose effective treatments and minimize the time and cost incurred by patients.

**Keywords :** Adaboost, Random Forest, Cryotherapy, Immunotherapy, Human Papilloma Virus

## INTISARI

*Human Papilloma Virus* dapat menyebabkan infeksi pada kulit dan selaput lendir. Kulit yang terinfeksi dapat menimbulkan benjolan ringan dengan bentuk kasar disebut kutil. Kutil yang berada di bagian kaki akan memberikan rasa tidak nyaman kepada pasien sehingga membutuhkan pengobatan segera mungkin. Saat ini dokter belum mengenali pengobatan mana yang lebih tepat. Selain itu, pengobatan yang diberikan masih berdasarkan pengalaman pribadi dan pengetahuan medis yang dimiliki. Oleh karena itu, teknologi ML diterapkan untuk memprediksi pengobatan kutil dengan tepat kepada pasien. Teknologi ML yang telah dikembangkan oleh para peneliti masih membutuhkan peningkatan akurasi lebih baik lagi agar keberhasilan dalam memprediksi pengobatan lebih akurat. Namun, jumlah sampel yang terbatas dan data yang memiliki masalah *imbalance class* mengakibatkan sulit untuk ditingkatkan. Penelitian ini bertujuan untuk meningkatkan akurasi dari penelitian sebelumnya.

Dalam upaya untuk mewujudkan tujuan dari penelitian ini maka terdapat algoritme yang diusulkan. Pertama, *mutual information* diterapkan untuk melakukan proses seleksi fitur. Metode tersebut hanya memberikan *non-negative value* atau memiliki kisaran 0 hingga 1 saja. Nilai itu dapat memberikan justifikasi dengan sangat jelas ketika memilih fitur mana yang lebih relevan. Kedua, algoritme klasifikasi yang digunakan yaitu kombinasi adaboost dan *random forest*. Algoritme *random forest* berperan sebagai *weak learner* untuk mengklasifikasi data *cryotherapy* dan *immunotherapy*. Selanjutnya, adaboost dapat digunakan untuk mendorong kinerja *weak learner* agar meraih akurasi lebih baik lagi. Kedua metode yang diusulkan memiliki kemampuan yang baik dalam masalah *imbalance class*. Data *cryotherapy* dan *immunotherapy* harus diacak terlebih dahulu sebelum diinduksikan ke dalam algoritme. Metode *10-fold cross-validation* diterapkan untuk mengacak data supaya algoritme *machine learning* dapat mempelajari pola yang bervariasi.

Hasil pengujian telah menunjukkan bahwa nilai akurasi sebesar 96,6% dan 91,1% pada *cryotherapy* dan *immunotherapy*. Hasil tersebut dapat dijadikan sebagai referensi dalam pengembangan selanjutnya dan merancang sistem yang akan datang. Nantinya, sistem tersebut dapat membantu para dokter untuk memilih pengobatan yang efektif dan meminimalisir waktu serta biaya yang dikeluarkan oleh pasien.

**Kata kunci** – Adaboost, *Random Forest*, *Cryotherapy*, *Immunotherapy*, *Human Papilloma Virus*.