

Abstrak

Varikokel merupakan salah satu penyebab infertilitas yang terbanyak pada pria. Metode deteksi dini varikokel menggunakan studi metabolomik cairan semen dan *electronic nose* memiliki potensi yang cukup besar. Metode penelitian ini menggunakan sampel cairan semen dalam gelas beker yang dihubungkan dengan alat GeNose. Penelitian ini menggunakan model ekstraksi ciri berupa nilai *mean*, *max*, *grad*, *std*, *trapz* dan *var* dari dataset pada fase *sampling* dengan pembagian waktu tertentu. Dataset kemudian dilakukan normalisasi dengan *minmax scaler*, dibagi menjadi 80% data train dan 20% data test menggunakan *stratified k-fold cross validation*. Klasifikasi data menggunakan beberapa algoritma *machine learning* seperti *Random Forest*, *Gradient Boosting*, *k-Nearest Neighbour*, *Support Vector Classifier* dan *Multiple Layer Perceptron*. Hasil studi ini menunjukkan bahwa nilai *Total Motile Sperm Count* (TMSC) memiliki potensi untuk diklasifikasikan menggunakan teknologi *electronik nose* metode ekstraksi ciri nilai *trapz* pada detik ke 10 s.d 40 menggunakan algoritma *Gradient Boosting* dengan nilai *train accuracy* 1, *test accuracy* 0.78, *preciission* 0.76 dan *recall* 0.82. Hasil tersebut berdasarkan kategori nilai TMSC kurang dari 30, kategori nilai TMSC antara 30-60 dan kategori nilai TMSC lebih dari 60. Adapun varikokel memiliki potensi untuk diklasifikasikan menggunakan teknologi *electronik nose* metode ekstraksi ciri nilai *grad* detik ke 10 s.d 40 dengan algoritma *Gradient Boosting* dengan *train accuracy* 1, *test accuracy* 0.83, *preciission* 0.88 dan *recall* 0.88. Hasil tersebut berdasarkan kategori klinis non varikokel, varikokel unilateral dan varikokel bilateral.

Kata kunci : TMSC, varikokel, *electronic nose*, *machine learning*, dan metabolomik.

Abstract

Varicocele is one of the most common causes of infertility in men. The method of early detection of varicocele using semen metabolomic studies and electronic nose has considerable potential. This research method uses a sample of semen in a beaker connected to a GeNose device. This study uses a feature extraction model in the form of mean, max, grad, std, trapz and var values from the dataset in the sampling phase with a certain time division. The dataset was then normalized using a minmax scaler, divided into 80% train data and 20% test data using stratified k-fold cross validation. The data classification uses several machine learning algorithms such as Random Forest, Gradient Boosting, k-Nearest Neighbor, Support Vector Classifier and Multiple Layer Perceptron. The results of this study indicate that the Total Motile Sperm Count (TMSC) value has the potential to be classified using electronic nose technology, the trapz value feature extraction method in seconds 10 to 40 using the Gradient Boosting algorithm with a train accuracy value of 1, test accuracy 0.78, precision 0.76 and recall. 0.82. The results are based on the TMSC value category less than 30, the TMSC value category between 30-60 and the TMSC value category more than 60. The varicocele has the potential to be classified using electronic technology nose feature extraction method grad seconds to 10 to 40 seconds with Gradient Boosting algorithm with train accuracy 1, test accuracy 0.83, precision 0.88 and recall 0.88. The results were based on the clinical categories of non varicocele, unilateral varicocele and bilateral varicocele.

Keywords: TMSC, varicocele, electronic nose, machine learning, and metabolomics.