IMPLEMENTATION OF CLASSIFICATION USING SUPPORT VECTOR MACHINE METHOD FOR DATA OF UNBALANCE BIOPSY DIAGNOSIS USING SMOTE

(Case Study: Patients Coming to Examine Cervical Cancer Diagnosis in the Hospital of Universitasrio de Caracas, Caracas Venezuela)

Elsa Murni Nasution

Statistics Study Program Faculty of Mathematics and Natural Sciences

Indonesian Islamic University

ABSTRACT

Cervical cancer is a disease that affects every part of the body that is characterized by abnormal cells or tissues that spread and damage normal cells around the patient's body quickly. One of the diagnoses of cervical cancer is by biopsy examination, performed if the results of a pelvic examination find growth or injury to the cervix, or when a Pap smear test shows something abnormal or cancerous. The risk factors for cervical cancer are alcohol consumption, smoking, age, having an uncircumcised male partner, a history of contraceptive use, and a lot of childbirth or high parity. Other factors of cervical cancer are sexually transmitted infections caused by the Hepatitis B virus, Genital Herpes, HPV, Syphilis, Gonorrhea, HIV and IDS. In this research the classification system used is the Vector Machine Support method with the Plynomial Kernel function and the Radial Basis Function (RBF) Kernel, due to unbalance in the Biopsy diagnosis data class for 679 patients who came to check the diagnosis of cervical cancer at the Universitario de Caracas Hospital in Caracas Venezuela it is necessary to do SMOTE. The results of the method in this study the best Vector Machine Support model is the Polynomial kernel function with an accuracy rate of 85.5172% with parameters C = 1 and d = 2 with a kappa value of 33.44%.

Keywords: *Cervical Cancer, Cervical Cancer Factors, Biopsy, Classification, Support Vector Machine (SVM), Polynomial Kernel, Radial Basis Function (RBF) Kernel, SMOTE, Classification Performance Measurement.*