

## ABSTRAK

Yogyakarta adalah sebuah wilayah dengan aktivitas tektonik yang tinggi sebagai akibat dari keberadaan zona subduksi di bagian selatan dan Zona Sesar Opak yang melintang di wilayah ini. Pergerakan zona sesar tersebut pada tanggal 27 Mei 2006 menghasilkan guncangan berkekuatan 6,3Mw. Disamping korban jiwa dan kerugian ekonomi, sebanyak lebih dari 180 titik mengalami gerakan massa di wilayah Yogyakarta dan sekitarnya. Hal tersebut kemudian melatarbelakangi dilakukannya penelitian ini, yang bertujuan untuk melakukan zonasi bahaya gerakan massa akibat gempa bumi. Penelitian dilakukan dengan menggunakan lima metode berbeda yang termasuk kedalam metode pengolahan pembelajaran mesin, yakni *nearest neighbour*, regresi logistik, *support vector machine*, *decision tree*, dan *random forest*, dimana akan dibandingkan performa dari kelima modul tersebut. Variabel independen yang digunakan meliputi nilai *peak ground acceleration*, *Vs30*, *slope*, *aspect*, *roughness*, *wetness index*, *stream power index*, NDVI, tataguna lahan, jarak sungai, jarak sesar, formasi batuan, dan kelengkungan lereng. Berdasarkan nilai AUC, metode *random forest* menghasilkan performa terbaik dengan nilai sebesar 0,96. Adapun nilai *recall* dan akurasi modul *random forest* sebesar 0,83 dan 0,87. Variabel yang berkorelasi signifikan terhadap kejadian gerakan massa akibat gempa bumi adalah *stream power index*, *slope*, dan jarak sesar. Mayoritas wilayah kerentanan terhadap gerakan massa tinggi berada di Kabupaten Gunung Kidul dan Kabupaten Bantul dengan proporsi 85,58% dari seluruh wilayah kerentanan tinggi pada wilayah penelitian.

Kata kunci : Zona Sesar Opak, gerakan massa akibat gempa bumi, pembelajaran mesin tersupervisi

## ABSTRACT

*Yogyakarta is known as seismically active area due to presence of subduction and Opak Fault Zone around the area. On May 27<sup>th</sup>, 2006, it experienced a 6,3Mw earthquake resulting from Opak Fault activity. Besides of economic loss and casualties, more than 180 landslides occurred throughout the region. This research aims to generate a landslide susceptibility zoning using supervised machine learning method. Five modules are applied, they are nearest neighbours, logistic regression, support vector machine, decision tree, and random forest. Their performances will be compared using recall, accuracy, and AUC score. Thirteen independent variables used are peak ground acceleration, Vs30, slope, aspect, roughness, wetness index, stream power index, NDVI, landuse, distance-to-river, distance-to-fault, stratigraphy unit, and curvature. Random forest module yields best performance with AUC score of 0.96, with recall and accuracy scores are 0.83 and 0.87 respectively. Stream power index, distance-to-fault, and slope are three variables significantly correlated with landslide events. Gunungkidul and Bantul Regency owns most areas with high landslide vulnerability, with the proportion of 85,58% of high susceptibility areas throughout research area.*

*Keywords : Opak Fault Zone, earthquake-induced landslide, machine learning*