

PEMETAAN BANJIR MENGGUNAKAN CROWDSOURCING DATA TWITTER BERDASARKAN DATA CURAH HUJAN OBSERVASI PERMUKAAN (KASUS: JAKARTA JANUARI 2020)

Shabira Putri Malahayati

18/423659/GE/08708

INTISARI

Setiap tahunnya, Jakarta yang merupakan ibukota Negara Indonesia selalu mengalami kejadian banjir. Saat terjadinya bencana alam seperti banjir, pemerintah wajib melakukan serangkaian tanggap darurat bencana, salah satunya suatu pemodelan bencana. Pemodelan bencana dapat dilakukan dengan berbagai sumber data, namun sumber data survei lapangan, penginderaan jauh, dan foto udara dianggap kurang efisien untuk digunakan. Survei lapangan memerlukan waktu yang lebih lama, sedangkan penginderaan jauh dan foto udara memiliki keterbatasan saat musim hujan yaitu terganggu akibat intensitas awan yang tinggi. Media sosial khususnya Twitter pada saat ini telah mendapat banyak perhatian dari berbagai kalangan sebagai sumber data untuk pemodelan banjir. Penelitian ini bertujuan untuk membangun basis data spasial, melakukan pemodelan banjir, menguji tingkat akurasi yang dihasilkan dan menguji data *tweet* dengan kejadian hujan di Jakarta.

Pemodelan banjir dilakukan dengan metode *Kernel Based Flood Mapping Model* yang menggunakan data DEMNAS, data persebaran sungai, batas administrasi, data kedalaman air di setiap pintu air sungai, data *tweet*, peta wilayah banjir BPBD, dan data curah hujan observasi permukaan yang dikeluarkan oleh BMKG. Hasil dari pemodelan banjir diuji akurasi dengan metode *overall accuracy* terhadap seluruh data *tweet* yang didapatkan dan peta wilayah banjir BPBD. Selain itu, dilakukan uji regresi untuk mengetahui hubungan antara curah hujan dengan data *tweet* terkait banjir di Jakarta. Hasil penelitian menunjukkan sebanyak 149 dari 12.345 *tweet* dapat disusun ke dalam basis data yang digunakan sebagai dasar pemodelan. Hasil pemodelan banjir menunjukkan nilai akurasi sebesar 70% berdasarkan perhitungan total titik banjir dan 57% berdasarkan zona banjir BPBD, nilai ini termasuk kategori rendah. Uji regresi antara titik banjir terhadap curah hujan bulan Januari menunjukkan hubungan yang tidak mempengaruhi. Nilai uji regresi antara curah hujan pada lokasi banjir dengan jumlah *tweet* sebesar 0,020822, sementara kedalaman banjir sebesar 0,049214 yang artinya variabel curah hujan hanya mempengaruhi variabel jumlah *tweet* sebesar 2% dan kedalaman banjir sebesar 4,9% serta sisanya dipengaruhi oleh faktor-faktor lain.

Kata Kunci: *data mining*, big data geospasial, banjir, Twitter, *kernel based flood mapping model*

***FLOOD MAPPING USING TWITTER CROWDSOURCING DATA
BASED ON RAINFALL DATA SURFACE OBSERVATIONS (CASE:
JAKARTA JANUARY 2020)***

Shabira Putri Malahayati

18/423659/GE/08708

ABSTRACT

Every year, Jakarta, which is the capital city of Indonesia, always experiences floods. When natural disasters such as floods occur, the government is obliged to carry out a series of disaster emergency responses, one of which is disaster modeling. Disaster modeling can be done using various data sources, but field survey data sources, remote sensing, and aerial photography are less efficient to use. Field surveys require more time, while remote sensing and aerial photography have limitations during the rainy season and are disrupted due to high cloud intensity. Social media, especially Twitter at this time, has received a lot of attention from various groups as a source of data for flood modeling. This study aims to build a spatial database, conduct flood modeling, test the resulting level of accuracy, and test tweet data with rain events in Jakarta.

Flood modeling is carried out using the Kernel-Based Flood Mapping Model method, which uses DEMNAS data, river distribution data, administrative boundaries, water depth data at each river sluice, tweet data, BPBD flood area maps, and surface observation rainfall data issued by BMKG. The flood modeling results were tested for accuracy with the overall accuracy method for all tweet data obtained and the BPBD flood area map. In addition, a regression test was carried out to determine the relationship between rainfall and flood-related tweet data in Jakarta. The results showed that 149 of the 12,345 tweets could be compiled into a database that was used as the basis for modeling. The flood modeling results show an accuracy value of 70% based on the calculation of the total flood points and 57% based on the BPBD flood zone; this value is in a low category. The regression test between rainfall and flood points shows no relationship. The regression test value between rainfall at the location of the flood and the number of tweets is 0,020822, while the depth of the flood is 0,049214, which means that the rainfall variable only affects the variable number of tweets by 2% and flood depth by 4.9%, and the rest is influenced by other factors.

Keywords: data mining, geospatial big data, floods, Twitter, kernel based flood mapping Model