

Prediksi risiko cyber melalui analisis big data pada media sosial menggunakan statistical machine learning = Cyber risk prediction through social media big data analytics using statistical machine learning

Andri Apriyana SA, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20455552&lokasi=lokal>

Abstrak

ABSTRAK

Sebagai proses alamiah dalam mencapai titik ekuilibrium, perkembangan ekonomi digital akan selalu diikuti oleh peningkatan risiko keamanan cyber. Penelitian ini membahas analisis big data percakapan media sosial Twitter dengan tipe data yang tidak terstruktur untuk memprediksi risiko cyber berupa keberhasilan serangan exploit terhadap kerentanan sistem informasi yang dipublikasikan pada situs kerentanan global cvedetails.com common vulnerabilities and exposures CVE . Penelitian ini mengeksplorasi aspek kualitatif dan kuantitatif atas data yang bersumber dari twitter dan CVE menggunakan metode pemodelan algoritmik statistical machine learning. Prediksi dilakukan dengan membandingkan beberapa model klasifikasi yang dipilih dari sepuluh algoritma yang paling banyak digunakan dalam data mining berdasarkan survey yang dilakukan oleh IEEE pada International Conference on Data Mining tahun 2006. Hasil prediksi terbaik dihasilkan melalui model algoritma Artificial Neural Networks dengan tingkat akurasi 96,73 . Model prediksi dapat dimanfaatkan oleh perusahaan asuransi dengan produk perlindungan risiko cyber untuk mengurangi potensi penyebaran klaim terjadinya risiko. Model juga dapat dimanfaatkan oleh perusahaan umum untuk melakukan mitigasi risiko cyber secara efektif dan efisien dengan menghindari situasi false-negatives error dalam pengelolaan risiko.

ABSTRACT

As a natural process in achieving equilibrium state, digital economic progress will always be followed by the increase of cyber security risk exposure. This research is focusing on unstructured Twitter social media big data analytics to predict cyber risks event in terms of successful attack on exploit based software vulnerability published in global vulnerability information websites cvedetails.com common vulnerabilities and exposures CVE . This research explores qualitative and quantitative aspect of data extracted from Twitter and CVE using statistical machine learning algorithmic modeling method. Prediction result obtained by comparing and selecting classification model from several statistical machine learning algorithm based on top ten algorithms in data mining survey produced by IEEE in 2006 International Conference on Data Mining. The best prediction results provided through Artificial Neural Networks algorithm with 96,73 accuracy rate. This prediction model offers advantages for insurance company with cyber liability product by reducing claim spread probability over cyber risk loss event. Prediction model can also be beneficial for company in general by providing an effective and efficient way to mitigate cyber risks through false negatives error avoidance in risk management.