



ABSTRAK

SISTEM DETEKSI PENYALAHGUNAAN DANA BANK DUNIA MENGGUNAKAN METODE RANDOM FOREST

Oleh

Muhammad Rizki Hakim
15/383247/PA/16907

Praktek korupsi sudah menjadi masalah utama di Indonesia dan belum ada tindakan signifikan yang dilakukan untuk menghentikan korupsi di Indonesia. Menurut laporan Tahunan KPK pada tahun 2019, Indeks Persepsi Korupsi Indonesia naik dari tahun 2015 hingga tahun 2019 (KPK, 2019). Tak hanya di Indonesia, praktek korupsi juga terjadi di tingkat Internasional, salah satunya di Bank Dunia. Bank Dunia menyediakan pinjaman uang dengan bunga yang rendah untuk membangun infrastruktur negara-negara berkembang.

Proses deteksi korupsi yang dilakukan oleh Bank Dunia memakan waktu yang sangat lama karena sistemnya yang bergantung pada pengajuan komplain. Jika, tidak ada komplain yang diajukan, maka Bank Dunia tidak akan melakukan penyelidikan. Penelitian ini mengimplementasi metode klasifikasi *Random Forest* untuk penyelidikan lebih cepat dilakukan tanpa menunggu ada laporan pada pihak terkait dan mempercepat proses penyelidikan yang akan dilakukan Bank Dunia. Fitur yang digunakan ada tipe kategori dan *integer*. Untuk tipe kategori akan diubah menggunakan metode *dummy variable*.

Setelah model *random forest* berhasil dibangun, hasil *Mean Squared Error* terendah dicapai untuk jumlah dataset 50 baris kontrak. Metode *dummy variable* membuat jumlah fitur pada model sama dengan jumlah label pada fitur suatu dataset, sehingga label baru yang ada pada dataset baru yang akan diprediksi artinya tambahan fitur asing yang tidak dikenali oleh model yang telah dibangun. Artinya, model yang telah dibangun tidak bisa memprediksi label baru yang belum ada pada dataset pada saat proses fitting.

Kata kunci : Random Forest, Klasifikasi, Pohon Keputusan, Bank Dunia, Deteksi Penipuan



ABSTRACT

WORLD BANK FRAUD DETECTION SYSTEM USING RANDOM FOREST METHOD

by

Muhammad Rizki Hakim
15/383247/PA/16907

Corruption practices have become a major problem in Indonesia and no significant action has been taken to stop corruption in Indonesia. According to the KPK Annual report in 2019, Indonesia's Corruption Perception Index rose from 2015 to 2019 (KPK, 2019). Not only in Indonesia, corrupt practices also occur at the International level, one of them at the World Bank. The World Bank provides low interest money loans to develop the infrastructure of developing countries.

The fraud detection process carried out by the World Bank takes a very long time because the system is dependent on filing complaints. If no complaint is submitted, the World Bank will not conduct an investigation. This research implements the Random Forest classification method for faster investigation without waiting for reports from relevant parties and accelerates the investigation process to be carried out by the World Bank. The features used are of category type and integer type. For category type features will be changed by using the dummy variable method.

After the random forest model was successfully built, the lowest MSE results were achieved for the 50 contract rows dataset. The dummy variable method makes the number of features in the model the same as the number of labels in the features of a dataset, so that new labels that exist in the new dataset that will be predicted means the addition of foreign features that are not recognized by the model that has been built. This means that the model that has been built cannot predict new labels that do not yet exist in the dataset during the fitting process.

Keyword : Random Forest, Classification, Decision Trees, World Bank, Fraud Detection