PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRAK

Transaksi kartu kredit membentuk aliran data yang sangat tidak seimbang, dengan pola penipuan yang terus bergeser akibat *concept drift*. Model *supervised* yang dilatih sekali tidak dapat mengenali pola baru, sedangkan *contextual bandit* belajar secara *online* dari *feedback* atas keputusannya sendiri. Penelitian ini membandingkan tiga kebijakan *contextual bandit* (ε-greedy, LinUCB, dan Linear Thompson Sampling) dalam desain *original* dan *shared-parameter* dengan dua model *supervised frozen* (Random Forest dan XGBoost), dalam kerangka *cost-sensitive* dengan tiga aksi (Approve, MFA, dan Review). Simulasi sekuensial dijalankan pada dataset Sparkov dengan drift semi-sintetis *sudden, gradual, incremental, recurring,* dan *adversarial*, serta *delayed feedback*. Kinerja diukur pada 10 seed menggunakan *Total Financial Loss* (TFL), yaitu nominal penipuan yang lolos ditambah biaya MFA dan *review*. Dari tiga biaya yang diuji, pembahasan berfokus pada biaya dengan TFL terendah (c_mfa = $0,20 dan c_rev = $1). *Shared-parameter bandit* menghasilkan TFL terendah pada empat dari enam skenario ($71 ribu hingga $102 ribu), sedangkan model *supervised* hanya unggul ketika pola penipuan stabil dan pada *gradual drift*. Saat terjadi *drift*, model *supervised* meloloskan 522 hingga 3.597 transaksi penipuan, sedangkan *shared-parameter bandit* hanya 169 hingga 948 transaksi. *Original bandit* menghasilkan TFL tertinggi ($171 ribu hingga $211 ribu) akibat tingginya biaya *review*. *Recall shared-parameter bandit* tetap 79–96% di semua skenario, sedangkan *recall* model *supervised* turun dari 98–99% pada S0 menjadi 60–65% pada S1, S2, dan S4, serta 19–21% pada S5. Temuan ini menunjukkan bahwa *shared-parameter bandit* lebih mampu beradaptasi terhadap *concept drift*, tetapi keunggulannya mengecil pada biaya intervensi yang lebih tinggi.

Kata kunci: deteksi penipuan kartu kredit, *contextual bandit*, *concept drift*, data tidak seimbang, *cost-sensitive learning*.

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

ABSTRACT

Credit card transactions form highly imbalanced data streams in which fraud patterns shift through concept drift. Supervised models trained once cannot recognise new patterns, whereas contextual bandits learn online from feedback on their own decisions. This study compares three contextual bandit policies (ε-greedy, LinUCB and Linear Thompson Sampling) in original and shared-parameter designs with two frozen supervised models (Random Forest and XGBoost), within a cost-sensitive framework with three actions (Approve, MFA and Review). Sequential simulations on the Sparkov dataset included semi-synthetic sudden, gradual, incremental, recurring and adversarial drift, and delayed feedback. Performance was measured over 10 seeds by Total Financial Loss (TFL): fraud let through plus MFA and review costs. Of the three cost pairs tested, the discussion focuses on the one with the lowest TFL (c_mfa = $0.20, c_rev = $1). The shared-parameter bandits achieved the lowest TFL in four of six scenarios ($71 thousand to $102 thousand), while the supervised models won only under stable fraud patterns and gradual drift. Under drift, the supervised models let through 522 to 3,597 fraudulent transactions, compared with 169 to 948 for the shared-parameter bandits. The original bandits had the highest TFL ($171 thousand to $211 thousand) due to high review costs. The recall of the shared-parameter bandit remained at 79–96% across all scenarios, while the recall of the supervised model dropped from 98–99% in S0 to 60–65% in S1, S2, and S4, and to 19–21% in S5. These findings show that shared-parameter bandits adapt better to concept drift, but their advantage shrinks at higher intervention costs.

Keywords: *credit card fraud detection, contextual bandit, concept drift, imbalanced data, cost-sensitive learning.*