BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Transaksi kartu kredit terus bertambah seiring meluasnya pembayaran digital, dan pertambahan ini diiringi naiknya risiko *fraud* yang menimbulkan kerugian finansial hingga miliaran dolar setiap tahun [1]. Kemunculan metode pembayaran baru, seperti *contactless payment*, memperluas celah yang dapat dimanfaatkan pelaku *fraud*, sehingga ancaman ini terus berkembang dari waktu ke waktu [2]. Membangun sistem deteksi *fraud* yang efektif karena itu menuntut model yang mampu bekerja pada karakteristik data yang sangat timpang, sekaligus dievaluasi dengan ukuran yang tidak menyesatkan pada kondisi tersebut.

Karakteristik utama data deteksi *fraud* kartu kredit adalah *class imbalance* yang ekstrem, yaitu transaksi *fraud* hanya menempati bagian yang sangat kecil dari keseluruhan transaksi, sehingga model kesulitan mempelajari pola kelas minoritas [3]. Pada *dataset* yang digunakan penelitian ini, proporsi *fraud* hanya 0,521% berbanding 99,479% transaksi normal. Ketimpangan sebesar ini menimbulkan dua konsekuensi. Pertama, pelatihan model perlu menyertakan mekanisme penyeimbangan kelas secara eksplisit, sebab tanpa itu model cenderung bias terhadap kelas mayoritas dan menghasilkan *recall* yang sangat rendah pada kelas minoritas [3, 4]. Kedua, metrik konvensional seperti *accuracy* menjadi kurang representatif, karena model yang selalu menebak kelas normal tetap memperoleh *accuracy* di atas 99% walaupun tidak pernah menangkap satu pun *fraud* [3]. Evaluasi karena itu dilakukan menggunakan sekumpulan metrik yang peka terhadap kelas minoritas, yaitu *precision*, *recall*, F1-*score*, *Precision-Recall AUC* (PR-AUC), dan *Matthews Correlation Coefficient* (MCC), dengan *accuracy* dan ROC-AUC tetap dilaporkan sebagai pelengkap agar hasil penelitian ini dapat dibandingkan langsung dengan literatur terkait pada *dataset* yang sama (Subbab 2.1).

Untuk menjawab tantangan tersebut, pendekatan *supervised learning* seperti *Random Forest* dan *XGBoost*, termasuk berbagai gabungan model (*ensemble*), telah banyak digunakan dan mencapai performa kompetitif pada metrik klasifikasi konvensional [4, 5, 6]. Kekuatan pendekatan ini terletak pada kemampuannya mempelajari pola statistik dari fitur transaksi berlabel melalui pemetaan langsung dari fitur ke label. Akan tetapi, model-model ini menilai tiap transaksi secara

independen dan seluruhnya bergantung pada kriteria pemisahan lokal seperti *impurity*, sehingga keberagaman antar model *supervised* yang digabungkan cenderung terbatas: model-model tersebut umumnya salah pada kasus yang sama karena mempelajari sinyal yang sama dengan cara yang serupa.

Sebagai alternatif, sejumlah penelitian memformulasikan deteksi *fraud* sebagai proses pengambilan keputusan sekuensial di bawah kerangka *reinforcement learning*, dengan alasan pendekatan ini belajar dari sinyal *reward* atas rangkaian keputusan, bukan dari pemetaan langsung fitur ke label [7, 8]. *Deep Q-Network* (*DQN*) khususnya memetakan deskripsi transaksi menjadi estimasi nilai tiap keputusan melalui *neural network*. Meski formulasi ini secara konseptual berbeda, keunggulannya sebagai model mandiri belum terbukti: performa *DQN* yang dilaporkan pada deteksi *fraud* kartu kredit masih terbatas dan lebih sering dikaitkan dengan potensi adaptasi jangka panjang daripada keunggulan metrik langsung [7]. Karena pendekatan *supervised* seperti *Random Forest* sudah mapan dan kompetitif sebagaimana dijelaskan sebelumnya, *DQN* lebih menjanjikan bila diposisikan untuk melengkapinya, bukan menggantikannya. Pemosisian ini didukung temuan empiris pada domain yang serupa: Papanastassiou et al. [9] mendapati bahwa model *tree* dan *Deep Q-Network* yang mengandalkan mekanisme belajar berbeda secara konsisten menempati titik operasi yang berbeda pula, yaitu model *tree* unggul pada *precision* sedangkan *Deep Q-Network* unggul pada *recall*, dan agen *reinforcement learning* mereka menangkap sekitar 30% *fraud* unik yang tidak tertangkap model *tree*. Pola serupa berpeluang muncul pada penelitian ini, sehingga penggabungan keduanya berpotensi memperbaiki deteksi melampaui masing-masing model secara terpisah.

Upaya menggabungkan *reinforcement learning* dengan *supervised classifier* sebenarnya sudah ada, tetapi peran *reinforcement learning* di dalamnya konsisten terbatas. Alkhozae et al. [10] menggabungkan *XGBoost*, *LightGBM*, *Random Forest*, *Gradient Boosting*, dan MLP, dengan *reinforcement learning* hanya menyetel ambang (*threshold*) keputusan secara adaptif, bukan menghasilkan skor deteksinya sendiri. Pola serupa tampak pada Cui et al. [11], yang memadukan *Graph Neural Network* dengan *reinforcement learning* untuk menyesuaikan ambang dan bobot fitur, serta pada Shen & Kurshan [12], yang memakai *DQN* untuk memilih ambang *alert* dari skor *classifier XGBoost* yang sudah ada. Pada ketiga penelitian ini, *reinforcement learning* selalu berdiri di belakang *classifier supervised* sebagai penyetel ambang, bukan sebagai sumber skor yang sejajar. Pada posisi lain, Papanastassiou et al. [9] menjalankan *Deep Q-Network* dan model *tree* sebagai dua

sistem paralel yang terpisah dan mendapati keduanya menangkap subset *fraud* yang sebagian berbeda, namun keduanya tetap dioperasikan sebagai dua skor terpisah tanpa pernah digabungkan menjadi satu keputusan. Baik sebagai penyetel ambang maupun sebagai sistem paralel, *reinforcement learning* pada penelitian-penelitian ini belum pernah diposisikan sebagai sumber skor yang sejajar dan disatukan dengan *classifier supervised*.

Menempatkan *reinforcement learning* sebagai sumber skor yang sejajar menuntut penyelesaian satu kendala teknis, yaitu skala keluaran kedua model yang berbeda. *Random Forest* menghasilkan probabilitas pada rentang $[0, 1]$, sedangkan *DQN* menghasilkan *Q-value* yang tidak terbatas, sehingga keduanya tidak dapat digabungkan langsung tanpa penyetaraan skala. Kedua alat untuk mengatasinya sudah tersedia dalam literatur, tetapi terpisah. Untuk penyetaraan skala, metode kalibrasi seperti *Platt scaling* dapat digunakan karena berlaku untuk skor skalar apa pun, tidak terbatas pada jenis model tertentu [13]. Untuk penggabungannya, Dal Pozzolo et al. [14] telah menggabungkan skor dua *classifier* melalui rata-rata probabilitas berbobot menjadi satu skor akhir, hanya saja kedua *classifier* di sana sama-sama *supervised* dan sudah berskala probabilitas sehingga tidak memerlukan kalibrasi. Dengan kata lain, kalibrasi skor dan rata-rata probabilitas berbobot sudah teruji sendiri-sendiri, tetapi belum pernah dirangkai untuk memasangkan *supervised classifier* dengan *reinforcement learning* sebagai dua sumber skor yang setara.

Dari sini muncul pertanyaan penelitian: sejauh mana *Q-value DQN*, setelah dikalibrasi menjadi probabilitas dengan cara yang sama seperti skor *classifier* lain, dapat digabungkan dengan probabilitas *Random Forest* melalui rata-rata berbobot untuk memperbaiki performa deteksi *fraud*, dan apakah perbaikan tersebut, jika ada, benar-benar berasal dari saling melengkapinya kedua model dan bukan sekadar variasi acak antar inisialisasi.

Penelitian ini memakai *dataset* sintetis Sparkov, yang validitasnya sebagai dasar pengujian metode deteksi *fraud* didukung oleh penelitian sebelumnya. Jemai et al. [15] mengujinya dengan *Random Forest*, *XGBoost*, dan *Naive Bayes* serta membandingkannya dengan data transaksi riil Eropa, sementara Al-Bulushi et al. [16] menguji rentang model yang lebih luas pada berkas Sparkov yang identik dengan penelitian ini. Kedua penelitian ini menegaskan Sparkov sebagai *dataset* yang representatif untuk menguji metode deteksi *fraud* kartu kredit, sekaligus menyediakan angka pembanding pada metrik yang sama. Keduanya, bagaimanapun, sama-sama tidak melibatkan *reinforcement learning* sebagai komponen *ensemble*, sehingga celah tersebut turut berlaku pada *dataset*

ini.

Berdasarkan pertimbangan tersebut, penelitian ini mengusulkan *pipeline hybrid ensemble* yang mengintegrasikan *Random Forest* sebagai *supervised classifier* dan *Deep Q-Network* berarsitektur *Dueling* sebagai model pelengkap berbasis *reinforcement learning*. *Q-value* dari *DQN* dikalibrasi dengan *Platt scaling* agar berada pada skala probabilitas yang sama dengan *Random Forest*, lalu kedua skor digabungkan melalui rata-rata probabilitas berbobot dengan bobot yang ditentukan pada data validasi. Pendekatan ini dirancang untuk menguji secara empiris, bukan mengasumsikan, apakah penggabungan terkalibrasi tersebut memperbaiki performa deteksi dibanding *Random Forest* sebagai *baseline*, tanpa mengganti *pipeline supervised* yang sudah berjalan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana *Q-value Deep Q-Network* yang telah dikalibrasi menjadi probabilitas melalui *Platt scaling* dapat digabungkan dengan probabilitas *Random Forest* melalui rata-rata probabilitas berbobot, dan sejauh mana penggabungan tersebut memperbaiki performa deteksi *fraud* kartu kredit dibandingkan *Random Forest* sebagai *baseline*?
2. Sejauh mana perbaikan tersebut, jika ada, berasal dari komplementaritas *Random Forest* dan *Deep Q-Network*, bukan dari variasi acak antar inisialisasi?

1.3 Batasan Permasalahan

Penelitian ini memiliki batasan-batasan sebagai berikut:

1. *Dataset* yang digunakan adalah *dataset* sintetis *Sparkov* ($\approx$ 1,8 juta transaksi dengan *fraud rate* $\approx$ 0,5%). Hasil belum divalidasi pada *dataset* transaksi riil.
2. Penanganan *class imbalance* dibatasi pada mekanisme internal algoritma, yaitu pembobotan kelas *cost-sensitive* pada *Random Forest* dan *stratified experience replay* pada *Deep Q-Network*. Teknik *resampling* eksternal seperti SMOTE atau *undersampling* tidak dieksplorasi sebagai variabel perbandingan.

3. Varian *Deep Q-Network* yang digunakan adalah arsitektur *Dueling* dengan konfigurasi *environment* dan *reward* yang tetap. Eksplorasi terhadap varian arsitektur *neural network* lain maupun sensitivitas terhadap parameter *environment* berada di luar cakupan penelitian ini.

4. Kombinasi *ensemble* difokuskan pada pasangan *Random Forest* dan *Deep Q-Network* dengan strategi penggabungan berupa rata-rata probabilitas berbobot. Kombinasi dengan lebih dari dua model dan strategi penggabungan yang lebih kompleks seperti *stacking* tidak dieksplorasi.

5. Evaluasi dilakukan secara *offline* pada *test set* kronologis dengan *policy* yang tidak di-*update* selama pengujian (*fixed policy evaluation*). Skenario *online learning*, integrasi ke sistem produksi, latensi *end-to-end*, dan kebijakan operasional tidak dibahas.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, penelitian ini memiliki tujuan sebagai berikut:

1. Mengimplementasikan *hybrid ensemble Random Forest* dan *Deep Q-Network* yang dikalibrasi dengan *Platt scaling* lalu digabungkan melalui rata-rata probabilitas berbobot, serta mengevaluasi performanya terhadap *Random Forest* sebagai *baseline* menggunakan *accuracy*, *precision*, *recall*, F1-*score*, F2-*score*, MCC, ROC-AUC, dan PR-AUC, dengan signifikansi perbedaan diuji menggunakan uji *Wilcoxon signed-rank* pada beberapa *seed*.

2. Menganalisis komplementaritas *Random Forest* dan *Deep Q-Network* melalui perbandingan himpunan transaksi *fraud* yang terdeteksi oleh masing-masing konfigurasi (*orthogonality analysis*), serta menguji konsistensi perbaikan antar inisialisasi, untuk memastikan perbaikan performa, jika ada, berasal dari komplementaritas tersebut dan bukan dari variasi acak.

1.5 Manfaat Penelitian

1. **Manfaat akademik**: penelitian ini memberikan bukti empiris bahwa *Deep Q-Network* memiliki nilai sebagai model pelengkap bagi *supervised classifier* pada deteksi *fraud* kartu kredit, bahkan ketika performa *DQN* sebagai model

tunggal masih di bawah *supervised baseline*. Selain itu, penelitian ini mengisi kekosongan dalam literatur deteksi *fraud* kartu kredit yang belum membahas integrasi *supervised classifier* dengan model berbasis pembelajaran interaksi melalui kalibrasi probabilitas eksplisit.

2. **Manfaat praktis**: pendekatan yang diusulkan memungkinkan institusi keuangan yang sudah memiliki *Random Forest* sebagai *baseline* untuk menambahkan *Deep Q-Network* sebagai model pelengkap tanpa mengganti arsitektur *supervised* yang sudah ada.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

1. **Bab 1 Pendahuluan** membahas latar belakang, rumusan masalah, tujuan, batasan, manfaat, dan sistematika penulisan.
2. **Bab 2 Landasan Teori** memuat tinjauan pustaka terkait deteksi *fraud*, *class imbalance*, *Random Forest*, *Deep Q-Network* dengan arsitektur *Dueling*, *Platt scaling*, strategi penggabungan skor (*ensemble fusion*), serta metrik evaluasi.
3. **Bab 3 Metodologi Penelitian** menjelaskan *dataset* dan pra-pemrosesan, seleksi fitur, penanganan *class imbalance*, formulasi proses pengambilan keputusan, arsitektur *Random Forest* dan *Deep Q-Network Dueling*, prosedur *Platt scaling*, strategi penggabungan skor, serta protokol evaluasi.
4. **Bab 4 Hasil dan Diskusi** menyajikan hasil eksperimen, perbandingan dengan *baseline* dan penelitian terkait, uji statistik, serta analisis komplementaritas antara *Random Forest* dan *Deep Q-Network*.
5. **Bab 5 Simpulan dan Saran** menyimpulkan temuan penelitian, kontribusi, keterbatasan, dan memberikan saran penelitian lanjutan.