BAB II

TINJAUAN PUSTAKA

Berbagai macam permasalahan tentang *fraud detection* telah menarik banyak penelitian untuk memecahkan permasalahan tersebut. Masalah *fraud* pada perusahaan layanan keuangan terjadi karena dua faktor. Faktor pertama dari sisi pelanggan yang memberikan data dirinya secara asal-asalan atau tidak real (*fake document*) (Subudhi and Panigrahi, 2020). Faktor kedua dari sisi transaksi pelanggan ketika diketahui terjadi transaksi yang tidak wajar (Bagga et al., 2020;Sabau, 2012). Terdapat beberapa metode yang digunakan untuk menyelesaikan permasalahan tersebut. Di penelitian sebelumnya, kecerdasan buatan telah terbukti berhasil digunakan untuk menyelesaikan permasalahan *fraud detection*. Terdapat beberapa pendekatan, diantaranya menggunakan teknik *machine learning, deep learning, ensemble learning, transfer learning, Fuzzy*, dan lainnya. Pada penelitian sebelumnya, para peneliti membuat sebuah model atau mengkombinasikan model yang sudah ada untuk menyelesaikan permasalahan *fraud detection* tersebut. Pada bagian ini, peneliti akan membandingkan dan membahas hasil penelitian dari penelitian-penelitian yang telah dilakukan sebelumnya.

Pendekatan yang paling sering digunakan untuk menyelesaikan permasalahan *fraud detection* adalah menggunakan *machine learning*. Pada penelitian (Alwadain, Ali and Muneer, 2023), masalah yang terjadi adalah *transaction fraud* pada institusi finansial. Jumlah total dataset yang digunakan pada penelitian ini ada 6.362.620 transaksi. Dataset diambil dari *online source*, yaitu

Kaggle (*Synthetic Financial Datasets for Fraud Detection*). Dari dataset tersebut, terjadi *imbalanced dataset* dimana jumlah transaksi yang *non-fraud* jauh lebih banyak daripada transaksi yang *fraud*. Oleh karena itu, peneliti melakukan *generate* 5000 *sample* data lagi untuk menambah dataset pada transaksi *fraud* menggunakan model *Conditional Generative Adversarial Network for Tabular Data* (CTGAN). Model yang dibuat memiliki akurasi yang lebih tinggi dari model pada penelitian-penelitian sebelumnya. Hasil akhir diujikan dengan menggunakan 27 algoritma *machine learning* (klasifikasi). Algoritma *XGBoost* memiliki performa paling baik dibanding dengan algoritma yang lainnya dengan skor akurasi 0,999. Pengujian juga dilakukan menggunakan *10-fold cross-validation*, dan algoritma *XGBoost* tetap memiliki performa terbaik yaitu dengan akurasi sejumlah 0,998.

Di penelitian lain (Dornadula and Geetha, 2019), *fraud detection* digunakan untuk mendeteksi *fraud* pada kartu kredit di sebuah *e-commerce*. Peneliti menggunakan pendekatan *machine learning* untuk memprediksi manakah transaksi kartu kredit yang masuk dalam kategori *fraud*. Peneliti menggunakan data historis untuk pembuatan model prediksi, untuk kemudian menghasilkan sebuah pola dari data yang telah dibuat model tersebut. Data diambil dari *Streaming Data Transaction*, yang meliputi *fraud* pada kartu kredit di Eropa. Peneliti membuat sebuah klaster untuk mengelompokkan data transaksi kedalam beberapa kelompok berdasarkan jumlah transaksi. Fokus pada penelitian ini yang pertama adalah bagaimana mengatasi dataset yang *imbalance*. Hal tersebut dapat terjadi karena dataset yang digunakan terkadang tidak seimbang antara data transaksi yang *fraud* dengan yang tidak *fraud*. Yang kedua adalah dari dataset yang berbeda model, yaitu

berlabel (*supervised*) dan tidak memiliki label (*unsupervised*). Yang ketiga adalah bagaimana meningkatkan kecepatan proses algoritma untuk mengolah data yang jumlahnya besar. Metode yang digunakan untuk mengatasi *imbalanced dataset* pada penelitian ini adalah SMOTE (*Synthetic Minority Over-sampling Technique*). Namun metode tersebut tidak memberikan hasil yang bagus. Lalu peneliti menggunakan metode *Matthew Coefficient Correlation* (MCC) dan satu metode lagi yaitu *one-class classifiers (one-class SVM)*. Pengujian dilakukan menggunakan beberapa algoritma *machine learning*. Algoritma *logistic regression*, *decision tree* dan *random forest* memiliki tingkat akurasi yang baik.

Pendekatan *machine learning* juga digunakan untuk menyelesaikan permasalahan *fraud detection* pada kartu kredit yang sering terjadi di sebuah perusahaan layanan keuangan (Manlangit et al., 2019). Peneliti menggunakan algoritma klasifikasi yaitu k-NN dan mengkombinasikan dengan metode PCA (*Principal Component Analysis*). Dataset pada penelitian ini diambil dari transaksi layanan keuangan selama dua hari. Jumlah transaksi ada 284.807 transaksi dan hanya 492 transaksi saja yang masuk dalam kategori *fraud* (0,172%). Pada penelitian ini juga terjadi masalah ketidakseimbangan data (*imbalanced dataset*) seperti pada penelitian sebelumnya. Proses deteksi anomali data dilakukan dengan menggunakan metode SMOTE (*Synthetic Minority Over-sampling Technique*). Pengujian dilakukan dengan mengukur presisi dan *recall* akurasi menggunakan AUPRC (*Area Under the Precision-Recall Curve*). Hasil akhir pada penelitian ini menunjukkan bahwa algoritma k-NN memiliki presisi dan *f1-score* masing-masing 98,32% dan 97,44%.

Pada penelitian lainnya, *deep learning* berhasil digunakan untuk menyelesaikan permasalahan pada *fraud detection*. Pada penelitian (Jan, 2021), peneliti menitikberatkan pada informasi yang terjadi pada layanan keuangan, misalnya pada status keuangan, informasi keuangan, dan laporan keuangan di sebuah perusahaan. Informasi yang ada kadang tidak simetris (tidak relevan). Untuk itu penelitian ini dilakukan untuk mendeteksi *financial statement fraud*. Dataset pada penelitian ini diambil dari 153 perusahaan finansial dari tahun 2001 hingga 2019 (*Taiwan Economic Journal*). Dari dataset tersebut, 51 perusahaan melaporkan *financial statement fraud*, sedangkan 102 perusahaan tidak melaporkan adanya *financial statement fraud*. Algoritma *deep learning* yang digunakan adalah RNN dan LSTM. Hasil akhir menunjukkan bahwa algoritma LSTM memberikan performa yang lebih baik daripada algoritma RNN. Hal tersebut terjadi karena memang algoritma LSTM merupakan pengembangan dari algoritma RNN. Algoritma LSTM memiliki akurasi sebesar 94,88%. Kedua algoritma ini memiliki performa pemrosesan data yang sangat baik ketika datasetnya berjumlah banyak. Pengujian pada penelitian ini menggunakan *confusion matrix* (akurasi, presisi, *recall*, *specificity*, dan *F1 score*), Type I dan Type II *error rate*, dan *ROC curve/AUC*.

Penelitian tentang *deep learning* untuk *fraud detection* juga digunakan untuk mengatasi *transaction fraud* pada kartu kredit (Alharbi et al., 2022). Peneliti mencermati di beberapa penelitian terdahulu yang menerapkan algoritma *machine learning* untuk kasus *fraud detection*, namun memiliki keterbatasan performa pada saat memproses data yang jumlahnya sangat besar. Selain itu, jumlah dataset yang

tidak seimbang (*imbalanced*) juga menjadi masalah utama. Peneliti mengusulkan metode baru, yaitu *deep learning*, dengan konsep *text to image* model. *Text* akan dikonversi kedalam sebuah *image*, dan kemudian akan diolah menggunakan algoritma *deep learning* yaitu CNN. Hasil akhir penelitian ini menghasilkan model akurasi sejumlah 99,87% untuk Coarse-KNN menggunakan *deep features* pada CNN. Pada penelitian ini, peneliti juga membandingkan beberapa penelitian terkait *computer vision* yang diterapkan untuk *fraud detection* pada transaksi kartu kredit. Dataset pada penelitian ini diambil dari situs *Kaggle*. Dataset terdiri dari 284.807 transaksi dengan 28 variabel dan 2 kelas. Dari dataset tersebut, terdapat 284.315 transaksi masuk dalam kategori non *fraud*, sedangkan 492 transaksi masuk dalam kategori *fraud*. Untuk mengatasi masalah *imbalanced dataset*, peneliti menggunakan pendekatan *class weight – inverse frequency method*.

Pendekatan lain yang digunakan untuk masalah *fraud detection* adalah *transfer learning*. Model *transfer learning* dibuat dengan mengadopsi model yang sudah ada sebelumnya, lalu menambahkan sedikit parameter untuk kasus yang baru (Lebichot et al., 2021). Penelitian ini melatarbelakangi sering terjadinya *transaction fraud* pada kartu kredit. Hal tersebut berpengaruh terhadap tingkat kepercayaan *customer* ketika akan melakukan transaksi di *e-commerce*. Fokus penelitian ini adalah bagaimana membangun sebuah sistem yang dapat mendeteksi *transaction fraud* dengan *short reaction time* dan *high precision*. Dataset diambil dari transaksi di beberapa *e-commerce* Eropa dalam rentang waktu enam bulan dengan jumlah dataset sebanyak lebih dari 200 juta transaksi. Peneliti menggunakan 15 model *transfer learning*, dari yang paling dasar adalah *Naïve Bayes* hingga yang paling

canggih. Peneliti juga membuat model baru pada penelitian ini dan membandingkannya dengan model yang sudah ada. Selain itu, peneliti mengusulkan sebuah pendekatan yang merupakan kombinasi *self-supervised* dan *semi-supervised classifiers*. Hasil menunjukkan bahwa model yang dibuat sangat akurat dan hampir tidak sensitif terhadap jumlah sampel berlabel.

Ensemble learning juga menjadi salah satu cara untuk kasus *fraud detection*. *Ensemble learning* dibuat dengan mengkombinasikan beberapa algoritma atau model yang sudah ada sehingga algoritma atau model tersebut menjadi lebih *optimal*. Penelitian (Bagga et al., 2020) menggunakan pendekatan *ensemble learning* untuk menyelesaikan kasus *fraud detection* pada perusahaan layanan keuangan. Peneliti melatarbelakangi banyaknya kasus *fraud* pada sebuah perusahaan layanan keuangan. Peneliti mencermati terdapat dua masalah pada *fraud detection*. Masalah pertama terdapat pada profil masing-masing penipu dan perubahan *behavioural transaction*. Masalah kedua adalah pada dataset yang digunakan seringkali tidak seimbang (*imbalanced dataset*). Peneliti mengusulkan sebuah model yaitu *ensemble learning*. Pada tahap pengujiannya, model yang dibuat oleh peneliti akan dibandingkan dengan beberapa model algoritma *machine learning*, yaitu *Ada Boost, k-nearest-neighbours, Logistic Regression, multilayer perceptron, Naïve Bayes, Quadrant Discriminative Analysis, Random Forest*, dan *Pipelining*. Peneliti menggunakan akurasi, presisi, *recall*, *F1-score* dan *confusion matrix* untuk membandingkan performa dari model yang dibuat. Model *ensemble learning* yang digunakan pada penelitian ini adalah *bagging classifier*. Peneliti juga menggunakan metode MCC (*Matthew Coefficient Correlation*), ADASYN, dan

BCR (*Balanced Classification Rate*) sebagai metode untuk mengatasi *imbalanced dataset*. Hasil akhir menunjukkan bahwa model *ensemble learning* yang dibuat peneliti memiliki akurasi yang sangat tinggi yaitu 99,99%.

Pendekatan *ensemble learning* juga digunakan pada penelitian (Xie et al., 2021) untuk kasus *Credit Card Fraud Detection* (CCFD) di sebuah perusahaan layanan keuangan. Studi kasus penelitian ini adalah sebuah bank. Masalah yang sering muncul pada CCFD yaitu tidak seimbangnya dataset (*imbalanced dataset*). Hal tersebut tentu sangat berpengaruh terhadap performa algoritma. Untuk itu, penelitian ini mengusulkan sebuah model *ensemble learning* yaitu *Heterogeneous Ensemble Learning Model Based on Data Distribution* (HELMDD) untuk mengatasi masalah *imbalanced dataset*. Dataset pada penelitian ini diambil dari dua sumber, yang pertama dari *Kaggle* (*public dataset*) dan yang kedua dari bank di China (*private dataset*). Dari sumber yang kedua (bank di China), jumlah dataset ada 24.024 transaksi, dengan 660 transaksi diantaranya (2,747%) masuk dalam kategori *fraud*. Pengujian model pada penelitian ini menggunakan *confusion matrix*. Hasil akhir menunjukkan bahwa model yang dibuat peneliti memiliki performa paling baik dibanding dengan beberapa model *machine learning* yang sudah ada. Model yang dibuat tidak hanya menghasilkan nilai *recall* yang baik, tetapi juga dapat meningkatkan *saving rates* bank masing-masing menjadi 0,8623 dan 0,6696.

Pada penelitian (Amini and Rabiei, 2022), peneliti menyoroti potensi metode *ensemble learning* untuk menyelesaikan permasalahan terkait klasifikasi, deteksi, prediksi pada penelitian-penelitian sebelumnya. Metode ini dapat memiliki

akurasi yang sangat tinggi dibanding dengan performa algoritma *machine learning* yang biasa digunakan untuk prediksi. Peneliti mengimplementasikan metode *ensemble learning* pada sebuah transaksi *e-commerce*. Terdapat enam metode *ensemble* yang digunakan pada penelitian ini, yaitu *Bagging, AdaBoost, Random Forest, XGBoost, Random SubSpace Neural Network* (NN) dan *Random SubSpace Support Vector Machine* (SVM). Dataset pada penelitian ini diambil dari *Kaggle*, dengan jumlah 6.362.620 baris data transaksi, dan 10 kolom. Peneliti membagi dataset menjadi 70% data *training* dan 30% data *testing*. Peneliti menguji performa model menggunakan metode *5-fold-cross-validation*, lalu menghitung nilai akurasi, presisi, *recall* dan *f1-score*. Hasil akhir penelitian menunjukkan bahwa algoritma *XGBoost* dan *Random Forest* memiliki performa paling baik dibanding dengan algoritma *ensemble learning* lainnya dengan akurasi masing-masing 99,87% dan 98,9%.

Penelitian (Baker, Mahmood and Shaker, 2022) meneliti tentang penggunaan *ensemble learning* untuk mendeteksi kasus *fraud* pada transaksi kartu kredit. Peneliti menggunakan dataset dari Kaggle. Data yang digunakan sejumlah 284.807 *rows* dengan 31 kolom. Dari dataset tersebut terdapat 492 data yang *fraud* dan 284.315 data *non-fraud*. Kemudian peneliti membagi dataset menjadi 80% data untuk *training* dan 20% untuk *testing*. Model yang digunakan pada penelitian ini yaitu *ensemble learning* dengan algoritma *supervised learning* pada *machine learning*. Untuk menguji performa model, peneliti menggunakan akurasi, *confusion matrix*, dan *classification report*. Hasil menunjukkan bahwa hasil terbaik menggunakan PCA (*Principal Component Analysis*) dengan skor akurasi mencapai

100%, skor presisi 97,3%, skor recall 73,5%, dan f1-score 83,7%. Peneliti menggunakan beberapa algoritma klasifikasi seperti *Random Forest* (RF), *Logistic Regression* (LR), *Bagging*, *Decision Tree* (DT), *Support Vector Machine* (SVM), *Naïve Bayes* (NB), dan *AdaBoost*.

Dari penelitian yang sudah ada, membuktikan bahwa *fraud* pada transaksi keuangan menjadi masalah yang sangat krusial bagi industri atau perusahaan yang bergerak di sektor layanan keuangan. Secara umum, masalah *transaction fraud* dapat diselesaikan menggunakan teknik klasifikasi maupun *clustering*, tergantung pada kasus dan dataset yang ada. Terdapat beberapa metode yang bisa digunakan untuk memecahkan masalah tersebut, beberapa diantaranya adalah menggunakan *machine learning* (Alwadain, Ali and Muneer, 2023; Dornadula and Geetha, 2019;Manlangit et al., 2019), *deep learning* (Jan, 2021;Alharbi et al., 2022), *transfer learning* (Lebichot et al., 2021), *ensemble learning* (Bagga et al., 2020;Xie et al., 2021;Amini and Rabiei, 2022;Baker, Mahmood and Shaker, 2022), serta beberapa teknik yang mengkombinasikan algoritma. Masalah yang sering muncul pada kasus *transaction fraud* yaitu masalah *imbalanced dataset*. Data transaksi yang *non-fraud* jauh lebih banyak daripada transaksi yang *fraud*. *Imbalanced dataset* dapat mengakibatkan model menjadi bias, lalu model lebih mengarah pada *majority class* sehingga dapat berpengaruh pada akurasi performa model. Kasus ini hampir muncul di setiap penelitian tentang *transaction fraud*. Masalah *imbalanced dataset* dapat diselesaikan menggunakan beberapa cara seperti CTGAN (*Conditional Generative Adversarial Network for Tabular Data*) (Alwadain, Ali and Muneer, 2023), SMOTE (*Synthetic Minority Over-sampling*

Technique) (Dornadula and Geetha, 2019;Manlangit et al., 2019;Baker, Mahmood and Shaker, 2022), MCC (*Matthew Coefficient Correlation*) (Dornadula and Geetha, 2019;Bagga et al., 2020), RMDD (*Resampling method based on the distribution of data*) (Xie et al., 2021), ADASYN (Bagga et al., 2020).

Pada penelitian ini, peneliti akan menggunakan metode *ensemble learning* untuk mengatasi masalah *fraud* pada transaksi keuangan. Teknik yang akan digunakan pada *ensemble learning* yaitu *single stacking*. Metode ini dibuat dengan cara melatih sejumlah model dasar, lalu menggabungkannya. Penggabungan dilakukan dengan melatih suatu *meta-model* untuk menghasilkan prediksi akhir berdasarkan prediksi-prediksi dari sejumlah model dasar tersebut. Terdapat tiga model dasar yang digunakan pada penelitian ini, yaitu *Support Vector Machine* (SVM), *K-Nearest Neighbour* (KNN), dan *Logistic Regression* (LR). Dan untuk meta model, peneliti menggunakan algoritma *Random Forest* (RF). Pada penelitian ini, peneliti menggunakan metode SMOTE (*Synthetic Minority Oversampling Technique*) untuk mengatasi masalah *imbalanced dataset*. Untuk menguji performa model, peneliti menggunakan *confusion matrix* dengan menghitung nilai akurasi, presisi, *recall*, dan *F1-score*.

Kebaharuan dari penelitian ini dibanding dengan penelitian sebelumnya adalah pada penggunaan metode *ensemble learning*. Pada penelitian sebelumnya, *ensemble learning* dibangun menggunakan metode *boosting* dan *bagging*, sedangkan pada penelitian ini menggunakan metode *stacking*. Perbedaan dari kedua metode ini terletak pada penggunaan model dasar. Pada *boosting* dan *bagging*, model dasar yang digunakan bersifat homogen, sedangkan pada *stacking*, model

dasar yang digunakan bersifat heterogen (model dasar yang dilatih menggunakan metode pembelajaran berbeda-beda). Selain itu, *stacking* belajar untuk menggabungkan sejumlah model dasar menggunakan satu model dasar lain (yang biasanya disebut *meta-learner*) sedangkan *bagging* dan *boosting* menggabungkan sejumlah model dasar berbasis algoritma deterministik (Suyanto et al., 2020).

Tabel 1. Daftar Penelitian Terdahulu

No	Fokus	Metode	Hasil	Referensi
1.	Membangun sebuah sistem yang dapat mendeteksi *transaction fraud* dengan *short reaction time* dan *high precision.*	- *Transfer learning* untuk mendeteksi *transaction fraud* kartu kredit pada sebuah *e-commerce*. - *Ensemble method* berdasarkan *self-supervised* dan *semi-supervised domain.*	- Akurasi dari beberapa model *transfer learning* sangat bergantung terhadap jumlah sampel yang diberi label pada *target domain*.	(Lebichot et al., 2021)
2.	Membangun sebuah pendekatan baru untuk memprediksi adanya *transaction fraud* menggunakan teknik *machine learning.*	- CTGAN untuk mengatasi *imbalanced dataset*. - 27 algoritma *machine learning* (klasifikasi).	Algoritma *XGBoost* memiliki performa paling baik dibanding dengan algoritma yang lainnya dengan skor akurasi 0,999.	(Alwadain, Ali and Muneer, 2023)
3.	- Mengatasi *imbalanced dataset*. - Dataset yang berbeda model, yaitu berlabel (*supervised*) dan tidak memiliki label (*unsupervised*). - Meningkatkan kecepatan proses algoritma untuk mengolah data yang jumlahnya besar.	- SMOTE. - *Matthew Coefficient Correlation* (MCC). - *One-class classifiers.* - Algoritma *Local Outlier Factor*, *Isolation Forest*, *Logistic Regression*, *Decision Tree*, *Random Forest.*	*Logistic regression*, *decision tree* dan *random forest* memiliki tingkat akurasi yang baik. *Logistic regression* memiliki akurasi sebesar 97,18%, presisi sebesar 98,31% dan MCC sebesar 94,38%. *Decision Tree* memiliki akurasi sebesar 97,08%, presisi sebesar 98,14% dan MCC sebesar 94,20%. *Random Forest* memiliki akurasi sebesar 99,98%, presisi sebesar 99,96% dan MCC sebesar 99,96%.	(Dornadula and Geetha, 2019)
4.	Proses deteksi anomali data pada kartu kredit yang sering terjadi di sebuah perusahaan layanan keuangan.	- Algoritma klasifikasi yaitu k-NN dan mengkombinasikan dengan metode PCA (*Principal Component Analysis*). - SMOTE. - AUPRC.	Hasil akhir pada penelitian ini menunjukkan bahwa algoritma k-NN memiliki presisi dan *f1-score* masing-masing 98,32% dan 97,44%.	(Manlangit et al., 2019)
5.	- Menitikberatkan pada informasi yang terjadi pada layanan keuangan,	- Algoritma *deep learning* yaitu RNN dan LSTM.	Algoritma LSTM memiliki akurasi sejumlah 94,88%. Kedua algoritma memiliki performa yang sangat baik	(Jan, 2021)

No	Fokus	Metode	Hasil	Referensi
	misalnya pada status keuangan, informasi keuangan, dan laporan keuangan di sebuah perusahaan. Informasi yang ada kadang tidak simetris (tidak relevan). - Mendeteksi *financial statement fraud*.	- Pengujian menggunakan *confusion matrix* (akurasi, presisi, *sensitivity*(recall), *specificity*, dan F1 score), *Type* I, *Type* II *error rate*, dan ROC curve/AUC.	ketika datasetnya berjumlah banyak, ditunjukkan dengan kecepatan pemrosesan data.	
6.	Metode baru, yaitu *deep learning*, dengan konsep *text to image model* untuk deteksi *transaction fraud*.	- Coarse-KNN. - *Class weight – inverse frequency method*.	Hasil akhir penelitian ini menghasilkan model akurasi sejumlah 99,87% untuk *Coarse-KNN* menggunakan *deep features* pada CNN.	(Alharbi et al., 2022)
7.	- Masalah pertama terdapat pada profil masing-masing penipu dan perubahan *behavioural transaction*. - Masalah kedua adalah pada dataset yang digunakan seringkali tidak seimbang.	- *Ensemble learning*. - LR, KNN, RF, NB, MLP, *AdaBoost, quadrant discriminative analysis, pipelining*. - akurasi, presisi, *recall*, F1 score dan *confusion matrix*. - MCC, ADASYN, dan BCR untuk *imbalanced dataset*.	Model yang dibuat peneliti, yaitu *ensemble learning*, mampu menghasilkan akurasi sebesar 99,99%.	(Bagga et al., 2020)
8.	- Studi kasus pada sebuah bank. - Masalah yang sering muncul pada CCFD (*Credit Card Fraud Detection)* yaitu tidak seimbangnya dataset (*imbalanced dataset*). - Pengujian pada dua dataset yang berbeda.	- *Ensemble learning*. - *Heterogeneous Ensemble Learning Model Based on Data Distribution* (HELMDD). - *Resampling method based on the distribution of data* (RMDD).	- Model yang dibuat memiliki performa paling baik dibanding dengan beberapa model *machine learning* yang sudah ada. - Model yang dibuat tidak hanya menghasilkan nilai *recall* yang baik, tetapi juga dapat meningkatkan *saving rates* bank masing-masing menjadi 0,8623 dan 0,6696.	(Xie et al., 2021)
9.	- Implementasi model *ensemble learning* untuk deteksi *fraud* pada transaksi *e-commerce*.	- *Ensemble learning*. - Algoritma *ensemble learning*: *Bagging, AdaBoost, Random Forest, XGBoost, Random SubSpace Neural Network* (NN) dan *Random SubSpace Support Vector Machine* (SVM). - 5-*fold-cross-validation*, lalu menghitung nilai akurasi, presisi, *recall* dan *f1-score*.	- Algoritma *XGBoost* dan *Random Forest* memiliki performa paling baik dibanding dengan algoritma *ensemble learning* lainnya dengan akurasi masing-masing 99,87% dan 98,9%.	(Amini and Rabiei, 2022)

No	Fokus	Metode	Hasil	Referensi
10.	- *Ensemble learning* untuk mendeteksi *fraud* pada transaksi kartu kredit.	- *Ensemble learning*. - PCA, SMOTE. - *Decision Tree* (DT), *Bagging*, *AdaBoost*, *Random Forest* (RF), *Naïve Bayes* (NB), *Logistic Regression* (LR), dan *Support Vector Machine* (SVM).	- Hasil terbaik menggunakan PCA (*Principal Component Analysis*) dengan skor akurasi mencapai 100%, skor presisi 97,3%, skor *recall* 73,5%, dan *f1-score* 83,7%.	(Baker, Mahmood and Shaker, 2022)
11.	- Deteksi fraud pada transaksi keuangan. - Cara mengatasi bias pada *imbalanced dataset*.	- *Ensemble learning* dengan model dasar: SVM, LR, kNN, dan meta model: RF. - Metode SMOTE untuk mengatasi masalah imbalanced dataset. - Pengujian model menggunakan *confusion matrix*: akurasi, presisi, recall, f1-score.		Penelitian penulis