

About the Dataset / Σχετικά με το Σύνολο Δεδομένων

EN — About the dataset

The Bank Account Fraud (BAF) dataset suite (NeurIPS 2022) contains six synthetic tabular datasets designed to benchmark fraud detection and (fair) machine learning methods under realistic constraints: class imbalance (very low fraud prevalence), temporal dynamics (distribution shifts over time), controlled bias types, and privacy-preserving generation (noise, feature encoding, CTGAN-based synthesis).

Source (Kaggle): <https://www.kaggle.com/datasets/sgpjesus/bank-account-fraud-dataset-neurips-2022>

EL — Σχετικά με το σύνολο δεδομένων

Η σουίτα δεδομένων για την απάτη τραπεζικών λογαριασμών (BAF) δημοσιεύτηκε στο NeurIPS 2022 και περιλαμβάνει 6 συνθετικά σύνολα δεδομένων σε μορφή πίνακα. Είναι σχεδιασμένη ως ρεαλιστικό πεδίο δοκιμών για μεθόδους ML και δίκαιης ML: έντονη ανισορροπία κλάσεων (πολύ χαμηλή επικράτηση απάτης), χρονική δυναμική (κατανομές που μετατοπίζονται), ελεγχόμενοι τύποι προκατάληψης και τεχνικές προστασίας ιδιωτικότητας (θόρυβος/κωδικοποίηση/CTGAN).

Πηγή (Kaggle): <https://www.kaggle.com/datasets/sgpjesus/bank-account-fraud-dataset-neurips-2022>

Data Analysis / Ανάλυση Δεδομένων

Reading guide (EN):

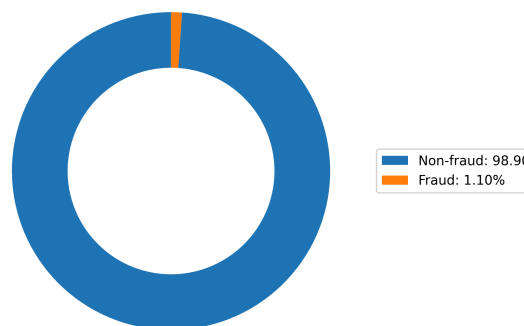
- *Fraud rate* is the fraction of fraud cases in a group.
- *Lift* is the fraud rate of a group divided by the overall fraud rate.
- lift = 1 → same as average risk
- lift > 1 → higher-than-average risk
- lift < 1 → lower-than-average risk

Οδηγός ανάγνωσης (EL):

- *Fraud rate* = ποσοστό απάτης μέσα σε μια ομάδα.
- *Lift* = (fraud rate ομάδας) / (συνολικό fraud rate).
- lift = 1 → ίδιο ρίσκο με τον μέσο όρο
- lift > 1 → υψηλότερο ρίσκο από τον μέσο όρο
- lift < 1 → χαμηλότερο ρίσκο από τον μέσο όρο

Section 1 — Prevalence / Ενότητα 1: Επικράτηση απάτης

Fraud vs Non-fraud (share)



Donut chart showing the share of fraud vs non-fraud in the dataset. / Διάγραμμα "donut" που δείχνει το ποσοστό απάτης έναντι μη-απάτης στο σύνολο δεδομένων.

Caption (EN): Overall class imbalance: fraud is a small minority of events. This immediately implies that accuracy alone is misleading; evaluation should focus on precision/recall-style metrics and top-k review performance.

Λεζάντα (EL): Η απάτη αποτελεί μικρό ποσοστό του συνόλου (έντονη ανισορροπία κλάσεων). Αυτό σημαίνει ότι η "ακρίβεια" (accuracy) από μόνη της είναι παραπλανητική: προτιμώνται μετρικές precision/recall και έλεγχος top-k (manual review των πιο ύποπτων).

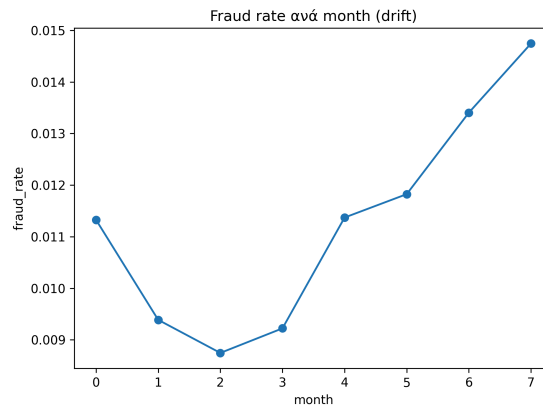
Section 2 — Drift over time / Ενότητα 2: Μετατόπιση με τον χρόνο (Drift)



Line chart showing fraud_rate per month (temporal drift). / Γραμμικό διάγραμμα που δείχνει fraud_rate ανά μήνα (χρονική μετατόπιση).

Caption (EN): Monthly fraud prevalence changes over time, indicating distribution drift. In a production setting, this is a monitoring signal: models and thresholds may need periodic recalibration.

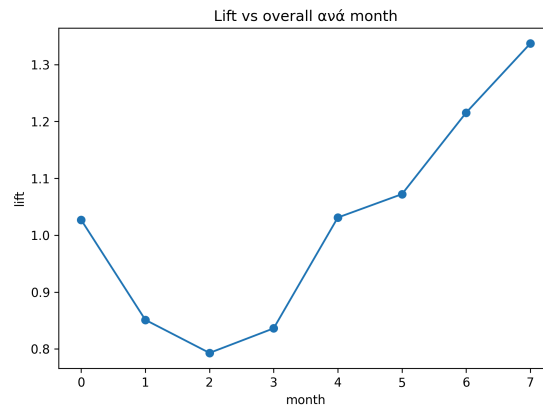
Λεξιάντα (EL): Το fraud rate μεταβάλλεται ανά μήνα, δείχνοντας drift. Σε πραγματικό σύστημα αυτό λειτουργεί ως “σήμα παρακολούθησης”: μπορεί να χρειάζονται αναπροσαρμογές σε thresholds ή/και επαναβαθμονόμηση μοντέλων.



Alternative view of monthly fraud_rate (secondary export). / Εναλλακτική απεικόνιση του fraud_rate ανά μήνα (δεύτερη εξαγωγή).

Caption (EN): Same idea as above, produced from a different export path; keep only one in the final report to avoid redundancy.

Λεξιάντα (EL): Ίδια πληροφορία με διαφορετική εξαγωγή· στο τελικό report κράτησε μόνο ένα για να αποφύγεις επανάληψη.

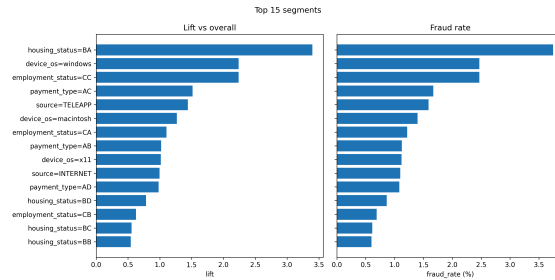


Line chart showing lift_vs_overall per month. / Γραμμικό διάγραμμα lift_vs_overall ανά μήνα.

Caption (EN): Lift over time highlights relative risk fluctuations compared to the global baseline. Useful for alerting: “risk level is drifting upward”.

Λεξιάντα (EL): Το lift ανά μήνα δείχνει πόσο “πάνω/κάτω” από τη συνολική βάση κινείται ο κίνδυνος. Είναι χρήσιμο για alerts τύπου “το ρίσκο ανεβαίνει”.

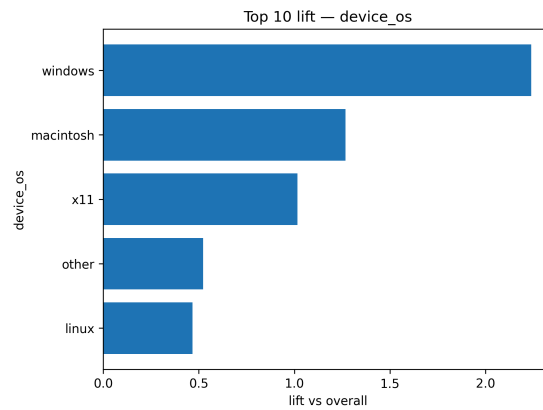
Section 3 — Risk segmentation / Ενότητα 3: Τμηματοποίηση ρίσκου (Segments)



Two-panel chart: left shows lift ranking; right shows fraud_rate (%) for top segments. / Διάγραμμα δύο πάνελ: αριστερά lift κατάταξη· δεξιά fraud_rate (%) για τα top segments.

Caption (EN): The chart ranks the highest-risk segments (lift) and shows their absolute fraud rate. Segments with high lift and meaningful sample size (n) are candidates for tighter controls or manual review.

Λεξάντα (EL): Κατάταξη των πιο “επικίνδυνων” segments (με lift) και παράλληλα το απόλυτο fraud rate. Segments με υψηλό lift και σημαντικό n είναι υποψήφια για αυστηρότερους ελέγχους ή δειγματοληπτικό έλεγχο.



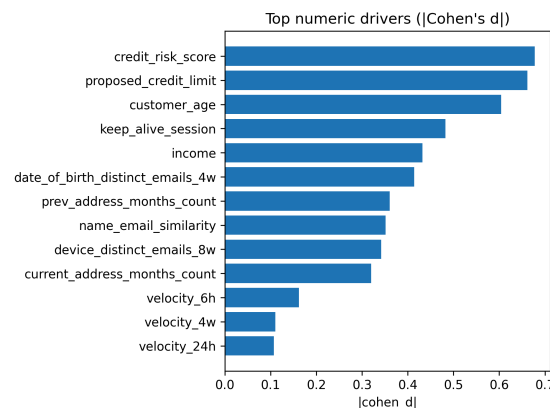
Horizontal bar chart of top 10 lift values for device_os categories. / Οριζόντιο bar chart για τα top 10 lift ανά κατηγορία device_os.

Caption (EN): Operating system categories show different relative risk levels. Interpret as “segment risk”, not as a causal claim.

Λεξάντα (EL): Οι κατηγορίες device_os εμφανίζουν διαφορετικά επίπεδα σχετικού ρίσκου. Αυτό διαβάζεται ως “segment risk” (στατιστικό μοτίβο), όχι ως αιτιότητα.

Note / Σημείωση: The figure 03_top10_device_os_lift.png appears twice in the source template. Keep it once to avoid duplication.
To 03_top10_device_os_lift.png εμφανίζεται διπλό· κράτα το μία φορά.

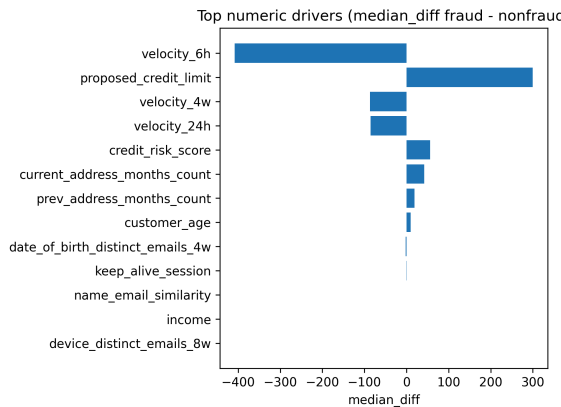
Section 4 — Numeric drivers (effect size) / Ενότητα 4: Αριθμητικοί “οδηγοί” (effect size)



Bar chart of top 15 numeric features by absolute Cohen's d. / Bar chart των top 15 αριθμητικών χαρακτηριστικών με βάση |Cohen's d|.

Caption (EN): |Cohen's d| provides a standardized difference between fraud vs non-fraud distributions. Larger values indicate stronger separation potential (as a univariate signal).

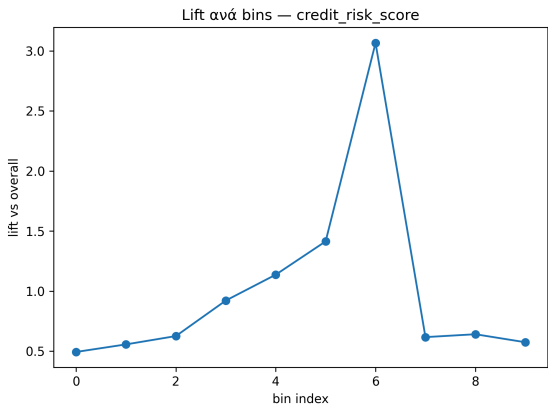
Λεξιάντα (EL): Το |Cohen’s d| δείχνει τυποποιημένη διαφορά κατανομών μεταξύ fraud και non-fraud. Όσο μεγαλύτερο, τόσο ισχυρότερο μονοδιάστατο σήμα διαχωρισμού.



Bar chart of top 15 numeric features by median_diff (fraud - nonfraud). / Bar chart των top 15 αριθμητικών χαρακτηριστικών με βάση median_diff (fraud - nonfraud).

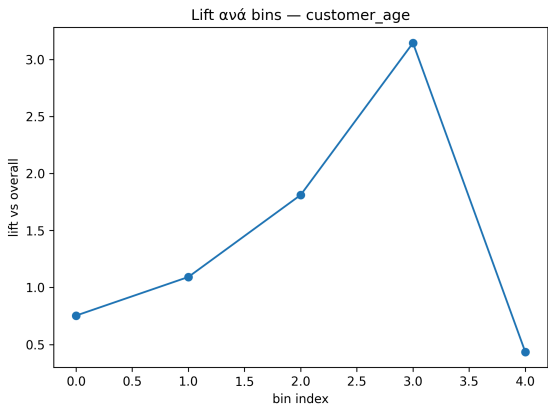
Caption (EN): Median difference is robust to outliers and complements Cohen’s d. It helps interpret direction: higher (or lower) values are associated with elevated risk.
Λεξιάντα (EL): Η διαφορά διαμέσου (median) είναι πιο ανθεκτική σε ακραίες τιμές και δείχνει κατεύθυνση: μεγαλύτερες (ή μικρότερες) τιμές σχετίζονται με αυξημένο ρίσκο.

Section 5 — Binned lift curves / Ενότητα 5: Lift ανά “κάδους” (bins)



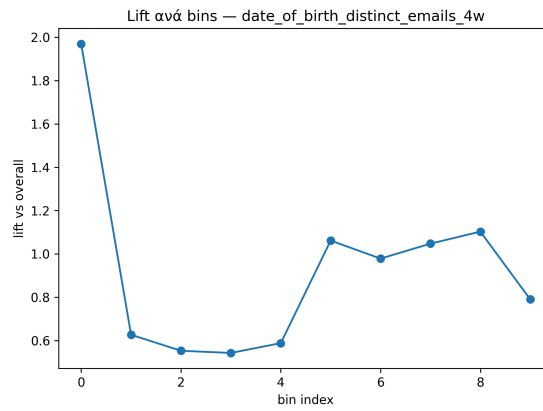
Lift by quantile bins for credit_risk_score. / Lift ανά quantile bins για credit_risk_score.

Caption (EN): As the score moves to higher bins, lift increases—evidence that the variable behaves like a risk score.
Λεξιάντα (EL): Όσο ανεβαίνουμε σε υψηλότερα bins, το lift αυξάνεται—ένδειξη ότι το πεδίο λειτουργεί σαν risk score.



Lift by quantile bins for customer_age. / Lift ανά quantile bins για customer_age.

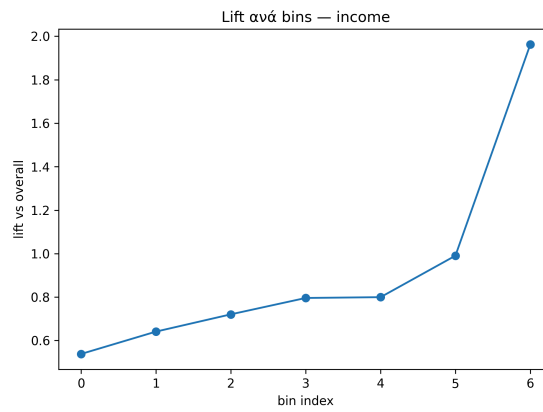
Caption (EN): Age bins show how risk changes across the distribution. Use as a descriptive pattern, not a normative interpretation.
Λεξιάντα (EL): Δείχνει πώς μεταβάλλεται το ρίσκο σε διαφορετικά επίπεδα ηλικίας. Διαβάζεται ως περιγραφικό μοτίβο.



Lift by quantile bins for date_of_birth_distinct_emails_4w. / Lift ανά quantile bins για date_of_birth_distinct_emails_4w.

Caption (EN): Behavioral signals (e.g., many distinct emails associated with the same DOB window) can indicate anomalous activity.

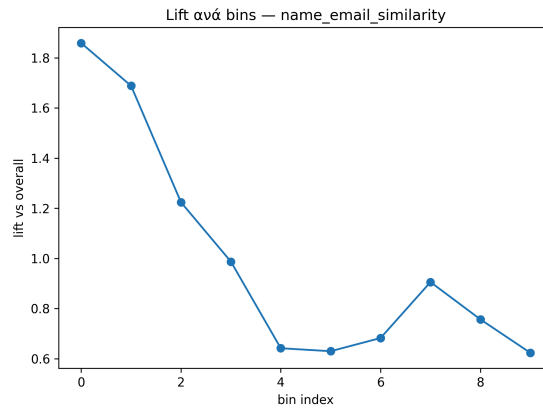
Λεξάντα (EL): Συμπεριφορικά σήματα (π.χ. πολλά διαφορετικά emails σε σύντομο παράθυρο) μπορεί να δείχνουν ασυνήθιστη δραστηριότητα.



Lift by quantile bins for income. / Lift ανά quantile bins για income.

Caption (EN): Income bins show relative risk gradients. In real deployments, such variables require careful fairness and governance checks.

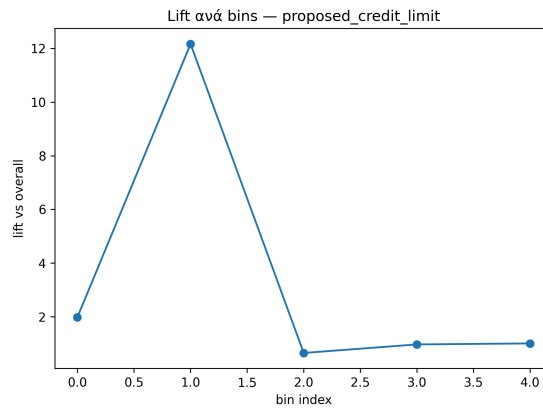
Λεξάντα (EL): Δείχνει κλίση ρίσκου ανά επίπεδα εισοδήματος. Σε πραγματική εφαρμογή απαιτείται προσοχή (fairness/governance).



Lift by quantile bins for name_email_similarity. / Lift ανά quantile bins για name_email_similarity.

Caption (EN): Lower similarity between name and email may correlate with higher risk. This is a heuristic behavioral pattern.

Λεξάντα (EL): Χαμηλότερη ομοιότητα ονόματος-email μπορεί να σχετίζεται με αυξημένο ρίσκο ως ευριστικό μοτίβο.

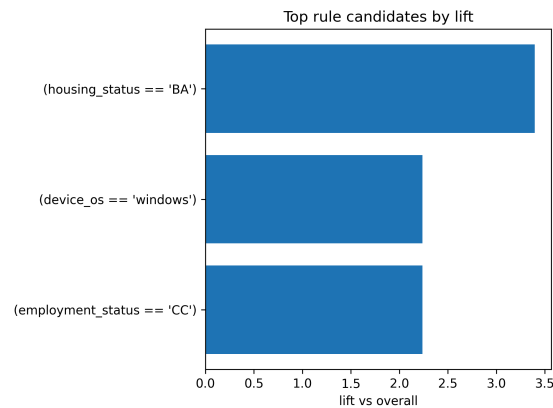


Lift by quantile bins for proposed_credit_limit. / Lift ανά quantile bins για proposed_credit_limit.

Caption (EN): Higher proposed credit limits often align with higher lift in this dataset, suggesting risk concentration at the upper tail.

Λεξιάντα (EL): Υψηλότερα προτεινόμενα όρια πίστωσης εμφανίζουν αυξημένο lift, δείχνοντας συγκέντρωση ρίσκου στο “πάνω άκρο”.

Section 6 — Rule candidates / Ενότητα 6: Υποψήφιοι κανόνες (rules)



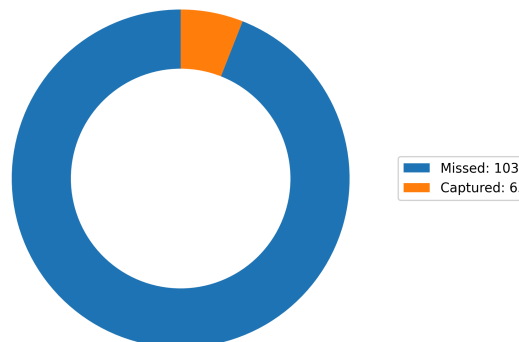
Bar chart ranking simple single-feature rule candidates by lift. / Bar chart κατάταξης απλών κανόνων (μονο-χαρακτηριστικό) με βάση lift.

Caption (EN): Simple rules can be used for baseline screening or manual review routing. They are interpretable but typically weaker than multi-feature models.

Λεξιάντα (EL): Απλοί κανόνες βοηθούν σε baseline screening ή δρομολόγηση για manual review. Είναι ερμηνεύσιμοι, αλλά συνήθως υστερούν από πολυπαραγοντικά μοντέλα.

Section 7 — Top-k review performance / Ενότητα 7: Απόδοση ελέγχου Top-k

captured vs missed (top-k) — proxy=credit_risk_score, k=0.01



Donut chart showing fraud captured vs missed when reviewing the top-k highest-risk cases. / Donut chart που δείχνει πόση απάτη “πιάνεται” και πόση χάνεται, όταν ελέγχουμε μόνο τα top-k πιο ύποπτα.

Caption (EN): This is an operational view: if we only review the top fraction of highest-risk applications, how much fraud do we capture vs miss?

Λεζάντα (EL): Επιχειρησιακή απεικόνιση: αν ελέγχουμε μόνο ένα μικρό top ποσοστό των πιο ύποπτων αιτήσεων, πόση απάτη “πιάνουμε” και πόση χάνουμε;

Conclusions / Συμπεράσματα

EN — What do these results imply?

1. **Strong class imbalance** means evaluation must emphasize *top-k capture*, precision/recall, and cost-sensitive decisions—not accuracy alone.
2. **Temporal drift exists**, so a real system needs monitoring (drift dashboards) and periodic recalibration.
3. **Certain segments show elevated lift** (relative risk). These are candidates for targeted controls—provided sample sizes are sufficient and governance constraints are respected.
4. **Numeric variables show clear risk gradients** (binned lift curves), supporting the use of scoring models and threshold policies.

How safe should we feel?

This analysis demonstrates how to detect risk patterns **in a synthetic benchmark dataset**. It can inform methodology and monitoring design, but **it does not measure the real-world safety of any bank or system**. In practice, “feeling safe” depends on: model quality, monitoring maturity, human review capacity, fraud response procedures, and fairness/governance safeguards.

EL — Τι σημαίνουν αυτά πρακτικά;

1. Η **έντονη ανισορροπία κλάσεων** απαιτεί αξιολόγηση τύπου *top-k capture*, precision/recall και αποφάσεις με βάση κόστος—όχι σκέτο accuracy.
2. Υπάρχει **χρονικό drift**, άρα ένα πραγματικό σύστημα χρειάζεται monitoring και περιοδικές αναπροσαρμογές.
3. Ορισμένα **segments έχουν αυξημένο lift** (σχετικό ρίσκο). Είναι υποψήφια για στοχευμένους ελέγχους, εφόσον το δείγμα (n) είναι επαρκές και τηρούνται περιορισμοί governance/fairness.
4. Οι **αριθμητικές μεταβλητές δείχνουν κλίσεις ρίσκου** (lift ανά bins), κάτι που υποστηρίζει πολιτικές scoring/threshold.

Πόσο ασφαλείς πρέπει να νιώθουμε;

Η ανάλυση δείχνει τη μεθοδολογία εντοπισμού μοτίβων ρίσκου **σε ένα συνθετικό benchmark dataset**. Δεν αποτελεί μέτρηση “ασφάλειας” πραγματικής τράπεζας/συστήματος. Στην πράξη, η αίσθηση ασφάλειας εξαρτάται από: ποιότητα μοντέλου, monitoring, δυνατότητα ανθρώπινου ελέγχου, διαδικασίες απόκρισης σε απάτη, και πλαίσιο governance/fairness.