

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΠΜΣ «**ΔΙΟΙΚΗΣΗ, ΑΝΑΛΥΤΙΚΗ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΕΠΙΧΕΙΡΗΣΕΩΝ**»

**Δήλωση Διπλωματικής Εργασίας ή Project – ΕΝΤΥΠΟ Β**

(Συμπληρώστε στο σκιασμένο μέρος. Η περιγραφή οφείλει να είναι λιτή και σύντομη)

**Διπλωματική Εργασία ή Project**

(Υπογραμμίστε την επιλογή σας)

**Ακαδ. Έτος**: 2025-2026

**Επώνυμο και Όνομα**: Κουτσομπίνας Κωνσταντίνος

**Εποπτεύων / Εποπτεύουσα:** Αργυρίου Αθανάσιος

(Είναι απαραίτητο να δηλωθεί)

Ημερομηνία

**Α. Προτεινόμενος Τίτλος**

(Σύντομος και εστιασμένος στο αντικείμενο)

Fraud Detection on transactional data with XGBoost and LGBM

**Β. Σύντομη Περιγραφή**

(Έως 250 λέξεις. Ρητή περιγραφή του προβλήματος και των αποτελεσμάτων ως προς περιεχόμενο, μέθοδο και σημαντικότητα. Πχ μπορείτε να σκεφθείτε σε όρους: αριθμός κεφαλαίων, περιεχόμενο, δεδομένα, αναμενόμενα αποτελέσματα, μέθοδος ανάλυσης)

Η εργασία αυτή επικεντρώνεται στην ανίχνευση ηλεκτρονικής απάτης σε συναλλαγές, αξιοποιώντας το σύνολο δεδομένων **IEEE-CIS Fraud Detection**. Το ζήτημα της απάτης στις online συναλλαγές αποτελεί σοβαρή πρόκληση, λόγω του μεγάλου όγκου δεδομένων και της χαμηλής συχνότητας εμφάνισης δόλιων περιπτώσεων. Στόχος είναι η ανάπτυξη και αξιολόγηση μοντέλων μηχανικής μάθησης με έμφαση στους αλγορίθμους **XGBoost** και **LightGBM**, οι οποίοι είναι ιδιαίτερα κατάλληλοι για δεδομένα μεγάλης κλίμακας και με έντονη ανισορροπία κλάσεων. Μέσα από διεξοδική ανάλυση δεδομένων, πειράματα και αξιολόγηση με κατάλληλες μετρικές (ROC-AUC, Precision-Recall AUC), αναμένονται αποτελέσματα που θα συμβάλλουν τόσο στη θεωρητική κατανόηση όσο και στην πρακτική εφαρμογή τεχνικών ανίχνευσης απάτης.

**Γ. 1 Σχετικότητα προς ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ**

3 – Ισχυρή, 2- Μέτρια, 1-Ασθενής. Εξηγείστε πολύ σύντομα.

1. Ασθενής συσχέτιση, ενδεχομένως να αποτελεί βοήθεια σε επιχειρηματικές αποφάσεις.

**Γ. 2 Σχετικότητα προς ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΕΠΙΧΕΙΡΗΣΕΩΝ**

3 – Ισχυρή, 2- Μέτρια, 1-Ασθενής. Εξηγείστε πολύ σύντομα.

3. Απαιτείται γνώση python και ανάπτυξης αλγορίθμων μηχανικής μάθησης.

**Γ. 3 Σχετικότητα προς ΑΝΑΛΥΤΙΚΗ ΕΠΙΧΕΙΡΗΣΕΩΝ, ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ**

3 – Ισχυρή, 2- Μέτρια, 1-Ασθενής Εξηγείστε πολύ σύντομα.

3. Η μηχανική μάθηση έχει άμεση συσχέτιση με την ανάλυση δεδομένων.

**Γ.4 Συνδυασμός των ανωτέρω**

(Έως 80 λέξεις)

**Γ.5 Τίποτε εκ των ανωτέρω**

(Έως 80 λέξεις)

**Δ. ΤΑΞΙΝΟΜΗΣΗ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ**

JEL

ή

ACM Computing Classification System (CCS

ή

Mathematics Subject Classification – MSC2020

ή συνδυασμός

**Δ1. ΤΑΞΙΝΟΜΗΣΗ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ: Περιοχή Έρευνας** (θεωρία αποφάσεων, οικονομικά, διοίκηση, στρατηγική, πληροφοριακά συστήματα, αναλυτική) **.**  (Έως 80 λέξεις)

**Δ2. ΤΑΞΙΝΟΜΗΣΗ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ: Μεθοδολογία** (επισκόπηση, πρωτογενής έρευνα, δευτερογενής έρευνα) **και εργαλεία ανάπτυξης** (επιχειρησιακή έρευνα, παλινδρόμηση, στατιστική, μαθηματικά, κλπ). (Έως 80 λέξεις)

**Δ3. ΤΑΞΙΝΟΜΗΣΗ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ: Θεματική περιοχή** (διοίκηση, πληροφοριακά συστήματα, αναλυτική) (Έως 80 λέξεις)

**Ε. ΕΡΓΑΛΕΙΑ**

**Ποια θα είναι η μεθοδολογική προσέγγιση;**

Ανάλυση, παρουσίαση και επεξεργασία των δεδομένων με χρήση python, ανάπτυξη και εφαρμογή αλγορίθμων μηχανικής μάθησης και σύγκριση αποτελεσμάτων.

**Από που θα αντλήσετε δεδομένα;** (Να αναφερθεί το είδος των δεδομένων, η πηγή τους, η διαθεσιμότητα)

Τα δεδομένα θα αντληθούν από τον Kaggle διαγωνισμό: [IEEE-CIS Fraud Detection | Kaggle](https://www.kaggle.com/competitions/ieee-fraud-detection)

**Θα χρησιμοποιήσετε λογισμικό;** (Αναφορά στο λογισμικό που θα χρησιμοποιηθεί για την ανάλυση των δεδομένων)

Χρήση python και βιβλιοθηκών της όπως pandas, numpy, sklearn κτλ.

**Προκαταρκτική Βιβλιογραφία (μερικές αναφορές)**

* Ali, A., Razak, S.A., Othman, S.H., Eisa, T.A., Al-Dhaqm, A., Nasser, M., Elhassan, T. and Elshafie, H., 2022. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied Sciences*, 12(19), p.9637. Available at: https://doi.org/10.3390/app12199637
* Ashfaq, T., Khalid, R., Yahaya, A.S., Aslam, S., Azar, A.T., Alsafari, S. and Hameed, I.A., 2022. A Machine Learning and Blockchain Based Efficient Fraud Detection Mechanism. *Sensors*, 22(19), p.7162. Available at: [https://www.mdpi.com/1424-8220/22/19/7162](https://www.mdpi.com/1424-8220/22/19/7162?utm_source=chatgpt.com)
* Höppner, S., Baesens, B., Verbeke, W. and Verdonck, T., 2020. Instance-Dependent Cost-Sensitive Learning for Detecting Transfer Fraud. *arXiv preprint arXiv:2005.02488*. Available at: [https://arxiv.org/abs/2005.02488](https://arxiv.org/abs/2005.02488?utm_source=chatgpt.com)
* Almalki, F. and Masud, M., 2025. Financial Fraud Detection Using Explainable AI and Stacking Ensemble Methods. *arXiv preprint arXiv:2505.10050*. Available at: https://arxiv.org/abs/2505.10050
* Huang, K., 2020. An Optimized LightGBM Model for Fraud Detection. *Journal of Physics: Conference Series*, 1651(1), p.012111. Available at: https://doi.org/10.1088/1742-6596/1651/1/012111
* Ben Boubker, M., Elmettat, T., Eddaoui, A. and Ouahabi, S., 2024. Fraud Detection in Financial Transactions Using Machine Learning with Oversampling Techniques: A Case Study of a Moroccan Bank. *Journal of Electrical Systems*, 20(10s). Available at: [https://journal.esrgroups.org/jes/article/view/6664](https://journal.esrgroups.org/jes/article/view/6664?utm_source=chatgpt.com)
* Tian, Y. et al., 2025. Financial Fraud Detection Based on an XGBoost and LSTM Fusion Model: A Comparative Study on the Enhancement of Time-Series Features. *Advances in Economics, Management and Political Sciences*, 170. Available at: [https://www.ewadirect.com/proceedings/aemps/article/view/23993](https://www.ewadirect.com/proceedings/aemps/article/view/23993?utm_source=chatgpt.com)
* Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.A., Caelen, O., Mazzer, Y. and Bontempi, G., 2017. SCARFF: a Scalable Framework for Streaming Credit Card Fraud Detection with Spark. *arXiv preprint arXiv:1709.08920*. Available at: [https://arxiv.org/abs/1709.08920](https://arxiv.org/abs/1709.08920?utm_source=chatgpt.com)
* Ramesh, A.C. and Uma, K., 2025. A Stacked LightGBM-XGBoost Model with SHAP-Based Fraud Detection for Financial Transactions. *International Scientific Journal of Engineering and Management*, 4(3), pp.1-7. Available at: <https://www.researchgate.net/publication/390260965>
* Vuppala, S.K. et al., 2025. Modeling Fraud Detection in Community Development Banking Through Machine Learning. *International Journal of Intelligent Systems and Applications in Engineering*. Available at: [https://ijisae.org/index.php/IJISAE/article/view/7547](https://ijisae.org/index.php/IJISAE/article/view/7547?utm_source=chatgpt.com)

**Στ. ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΟΛΟΚΛΗΡΩΣΗΣ**

(**Συνοπτική παρουσίαση με διάγραμμα GANTT**)

**Ζ. Σχόλια από τον ακαδημαϊκό υπεύθυνο**

**ΠΡΟΣΟΧΗ**

* Το Έντυπο Β οφείλει 1) να έχει όνομα επιβλέποντος και 2) την έγκριση του επιβλέποντος για το θέμα.
* Το Έντυπο Β συνοδεύεται από το Χρονοδιάγραμμα εκπόνησης – είναι στα Έγγραφα του eclass. Η ημερομηνία υποβολής της διπλωματικής / project είναι η 28/2/2026.
* Το Έντυπο Β αναρτάται στο «eclass\Διπλωματική Εργασία\Εργασίες – Έντυπο Β» και υποβάλλεται στη Γραμματεία ([bis-analytics@econ.uoa.gr](mailto:bis-analytics@econ.uoa.gr)) σε μορφή doc ή Latex ή text k και pdf με ταυτόχρονη κοινοποίηση στον επιβλέποντα.