



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ ΕΜΠ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

Διπλωματική εργασία

**«Εκτίμηση της αποδοτικότητας ευρωπαϊκών σιδηροδρομικών
εμπορευματικών μεταφορών με εφαρμογή μοντέλου δύο σταδίων DEA και
τεχνητών νευρωνικών δικτύων»**

Πέτρου Δανιήλ



Επιβλέπουσα: Τατιάνα Π. Μοσχόβου, Ε.Δι.Π. ΕΜΠ

Αθήνα, Ιούλιος 2025

Ευχαριστίες

Η ολοκλήρωση της παρούσας διπλωματικής εργασίας σηματοδοτεί το τέλος ενός σημαντικού κεφαλαίου της ακαδημαϊκής μου πορείας και την αρχή μιας νέας περιόδου με στόχους, όνειρα και ευθύνες. Στο σημείο αυτό επιθυμώ να εκφράσω τις θερμές μου ευχαριστίες σε όλους εκείνους που συνέβαλαν με τον δικό τους τρόπο στην πραγματοποίηση αυτής της προσπάθειας.

Θα ήθελα να ευχαριστήσω ιδιαίτερα την επιβλέπουσά μου, κυρία Τατιάνα Μοσχόβου, Ε.ΔΙ.Π. ΕΜΠ για την πολύτιμη καθοδήγηση, τη διαρκή υποστήριξη, τις ουσιαστικές παρατηρήσεις, καθώς και τον χρόνο που αφιέρωσε σε όλα τα στάδια εκπόνησης της διπλωματικής εργασίας. Η συμβολή της υπήρξε καταλυτική για την επιστημονική τεκμηρίωση, τη μεθοδολογική ακρίβεια και την πληρότητα της μελέτης.

Επίσης θα ήθελα να ευχαριστήσω τα μέλη της επιτροπής κα Χ. Πλατή, Καθηγήτρια ΕΜΠ και κο. Κ. Γκιотσαλίτη, Επίκουρο Καθηγητή ΕΜΠ για την πρόθυμη συμμετοχή τους, και την υποστήριξη κατά τη διαδικασία αξιολόγησης.

Τέλος, ευχαριστώ από καρδιάς την οικογένειά μου, τους φίλους και τους ανθρώπους που στάθηκαν δίπλα μου με υπομονή, κατανόηση και εμπύχωση καθ' όλη τη διάρκεια των σπουδών, αλλά και της ζωής μου γενικότερα.

Τίτλος: Εκτίμηση της αποδοτικότητας ευρωπαϊκών σιδηροδρομικών εμπορευματικών μεταφορών με εφαρμογή μοντέλου δύο σταδίων DEA και τεχνητών νευρωνικών δικτύων

Συγγραφέας Διπλωματικής Εργασίας: Δανιήλ Πέτρου

Επιβλέπουσα: Τατιάνα Π. Μοσχόβου

Σύνοψη

Η παρούσα διπλωματική εργασία διερευνά την αποδοτικότητα των ευρωπαϊκών οργανισμών εμπορευματικών σιδηροδρόμων χρησιμοποιώντας μια υβριδική μεθοδολογική προσέγγιση που ενσωματώνει την Περιβάλλουσα Ανάλυση Δεδομένων (DEA) με τα Τεχνητά Νευρωνικά Δίκτυα (ANN). Η DEA χρησιμοποιείται ως μη παραμετρική μέθοδος για την αξιολόγηση της σχετικής αποτελεσματικότητας των σιδηροδρομικών εταιρειών, αντιμετωπίζοντάς τις ως μονάδες λήψης αποφάσεων (DMUs), χρησιμοποιώντας εισροές όπως το μήκος του δικτύου, το τροχαίο υλικό, το ανθρώπινο δυναμικό, τις κεφαλαιουχικές δαπάνες, το λειτουργικό κόστος και την κατανάλωση ενέργειας, και εκροές όπως τα τονοχιλιόμετρα και τα έσοδα. Παρόλο που η DEA παρέχει πολύτιμες γνώσεις συγκριτικής αξιολόγησης, έχει επίσης εγγενείς περιορισμούς, όπως η μη προβλεπτική της φύση, η ευαισθησία στη σύνθεση των δεδομένων και η τάση να προτείνει στόχους βελτίωσης που μπορεί να είναι μη εφαρμόσιμοι στην πράξη. Για να ξεπεραστούν αυτοί οι περιορισμοί, η παρούσα διπλωματική εμβαθύνει περισσότερο με την ενσωμάτωση Τεχνητών Νευρωνικών Δικτύων, τα οποία εκπαιδεύονται στα αποτελέσματα της DEA, επιτρέποντας στο μοντέλο ANN να μάθει τις πολύπλοκες μη γραμμικές σχέσεις μεταξύ των λειτουργικών μεταβλητών και των αποτελεσμάτων αποδοτικότητας. Τα εκπαιδευμένα μοντέλα ANN μπορούν να προβλέψουν βελτιστοποιημένες μεταβλητές εισόδου για κάθε οργανισμό, επιτρέποντας την ανάλυση σεναρίων και την προσαρμοστική πρόβλεψη αποτελεσματικότητας. Επιπλέον, τα ANN ενισχύουν την ερμηνευσιμότητα και τη δυνατότητα εφαρμογής των ευρημάτων της DEA, προτείνοντας πιο ρεαλιστικές και εφικτές βελτιώσεις αποτελεσματικότητας. Η υβριδική προσέγγιση DEA-ANN εφαρμόστηκε σε ένα σύνολο δεδομένων 13 ευρωπαϊκών εμπορευματικών σιδηροδρομικών οργανισμών. Τα αποτελέσματα καταδεικνύουν ότι το ενισχυμένο με ANN μοντέλο μπορεί όχι μόνο να αναπαράγει τη λογική της DEA, αλλά και να παρέχει εφαρμόσιμες οδηγίες για βελτιστοποίηση της λειτουργίας της. Τα αποτελέσματα υποδηλώνουν ότι τέτοιες ολοκληρωμένες μέθοδοι μπορούν να υποστηρίξουν τη λήψη αποφάσεων βάσει δεδομένων, τη συγκριτική αξιολόγηση βέλτιστων πρακτικών και τη συνεχή βελτίωση στον ευρωπαϊκό τομέα των σιδηροδρομικών εμπορευματικών μεταφορών.

Λέξεις-κλειδιά: Εμπορευματικές σιδηροδρομικές μεταφορές, Περιβάλλουσα Ανάλυση Δεδομένων (DEA), Τεχνητά Νευρωνικά Δίκτυα (ANN), πρόβλεψη αποδοτικότητας, Τεχνητή Νοημοσύνη (AI)

Title: Assessment of the efficiency of European rail freight transport using a two-stage DEA model and artificial neural networks

Thesis Author: Daniil Petrou

Supervisor: Tatiana P. Moschovou

Abstract

The current research investigates the efficiency of European freight railway organizations using a hybrid methodological approach that integrates Data Envelopment Analysis (DEA) with Artificial Neural Networks (ANN). DEA is employed as a non-parametric method to evaluate the relative efficiency of railway companies, treating them as decision-making units (DMUs), using operational inputs such as network length, rolling stock, staff, capital expenditures, operating costs, and energy consumption, and outputs such as ton-kilometers and revenue. While DEA provides valuable benchmarking insights, it also has some limitations, such as its non-predictive nature, sensitivity to data composition, and the tendency to propose improvement targets that may be impractical in real-world applications. To overcome these limitations this diploma thesis extends the framework by incorporating Artificial Neural Networks trained on DEA results, enabling the ANN to learn the complex non-linear relationships between operational variables and efficiency scores. The trained ANN models can forecast optimized input configurations for each company, allowing scenario analysis and adaptive efficiency forecasting. Additionally, the ANN enhances the interpretability and applicability of DEA findings by suggesting more realistic and achievable efficiency improvements. The hybrid DEA–ANN approach was applied to a dataset of 13 European freight railway organizations. Results demonstrate that the ANN-enhanced framework can not only reproduce DEA efficiency logic but also provide actionable guidance for operational optimization. The findings suggest that such integrated methods can support data-driven decision-making, best-practice benchmarking, and continuous improvement in the European rail freight sector.

Keywords: Freight rail transport, DEA, Artificial Neural Networks (ANNs), efficiency prediction, AI-driven benchmarking

Περίληψη

Η βελτίωση της αποδοτικότητας στις μεταφορές αποτελεί διαχρονικό στόχο της επιστημονικής έρευνας και της δημόσιας πολιτικής, καθώς σχετίζεται με την αποτελεσματική αξιοποίηση των πόρων, τη βιωσιμότητα και την οικονομική ανάπτυξη. Η αποδοτικότητα των σιδηροδρομικών οργανισμών εμπορευματικών μεταφορών αποτελεί ένα από τα πλέον καίρια ζητήματα που απασχολούν τις ευρωπαϊκές πολιτικές μεταφορών, καθώς σχετίζεται άμεσα τόσο με την περιβαλλοντική βιωσιμότητα όσο και με την επιχειρησιακή αποτελεσματικότητα των εφοδιαστικών αλυσίδων. Σε ένα πλαίσιο όπου η στρατηγική προτεραιότητα της Ευρωπαϊκής Ένωσης εστιάζει στη μείωση των εκπομπών διοξειδίου του άνθρακα και την ενίσχυση των σιδηροδρομικών μεταφορών έναντι των οδικών, η ανάγκη για αξιόπιστη μέτρηση και στοχευμένη βελτίωση της αποδοτικότητας των σιδηροδρομικών φορέων καθίσταται επιτακτική. Η παρούσα διπλωματική εργασία αποσκοπεί στην ικανοποίηση της ανάγκης αυτής μέσω της ανάπτυξης και εφαρμογής ενός μοντέλου δύο σταδίων, το οποίο συνδυάζει τη μέθοδο της Περιβάλλουσας Ανάλυσης Δεδομένων (Data Envelopment Analysis - DEA) με τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks - ANN). Αντικείμενο της μελέτης αποτελεί η συγκριτική αξιολόγηση δεκατριών ευρωπαϊκών σιδηροδρομικών οργανισμών που δραστηριοποιούνται στον τομέα των εμπορευματικών μεταφορών, με σκοπό όχι μόνο την αποτύπωση του επιπέδου αποδοτικότητας αλλά και την πρόβλεψη εναλλακτικών σεναρίων που οδηγούν σε ουσιαστική επιχειρησιακή βελτίωση.

Στο πλαίσιο της εν λόγω διπλωματικής εργασίας, πραγματοποιήθηκε μία ολοκληρωμένη μελέτη με στόχο την αξιολόγηση και βελτίωση της αποδοτικότητας των υπό μελέτη ευρωπαϊκών σιδηροδρομικών εμπορευματικών οργανισμών, αξιοποιώντας προηγμένες τεχνικές από τον χώρο της επιχειρησιακής έρευνας και της τεχνητής νοημοσύνης. Αρχικά, παρουσιάστηκε το θεωρητικό υπόβαθρο των μεθόδων που υιοθετήθηκαν, δίνοντας ιδιαίτερη έμφαση στη λογική, τα πλεονεκτήματα και τους περιορισμούς της Περιβάλλουσας Ανάλυσης Δεδομένων (DEA) και των Τεχνητών Νευρωνικών Δικτύων (ANN). Εν συνεχεία, αναλύθηκε διεξοδικά η μεθοδολογική προσέγγιση που ακολουθήθηκε, τόσο ως προς την κατασκευή του συνόλου δεδομένων, όσο και ως προς τα βήματα εφαρμογής των δύο μεθόδων. Η συλλογή των δεδομένων πραγματοποιήθηκε από επίσημες πηγές, ετήσιες εκθέσεις και οικονομικά δελτία των δεκατριών οργανισμών, ενώ σε περιπτώσεις ελλείπων στοιχείων εφαρμόστηκαν συγκεκριμένοι υπολογισμοί και τεκμηριωμένες παραδοχές για τη διασφάλιση της εγκυρότητας του συνόλου των μεταβλητών. Τα δεδομένα αυτά διαρθρώθηκαν κατάλληλα ώστε να τροφοδοτήσουν το υπόδειγμα DEA και έπειτα να αξιοποιηθούν για την εκπαίδευση των ANN.

Η μεθοδολογική προσέγγιση που υιοθετήθηκε βασίστηκε στη σταδιακή εφαρμογή ενός μοντέλου δύο σταδίων το οποίο επιδιώκει να εφαρμόσει τη DEA, να αναδείξει τις αδυναμίες τις και με τις δυνατότητες που προσφέρει το ANN να καταφέρει να τις υπερβεί. Σε πρώτο στάδιο, εφαρμόστηκε το μοντέλο DEA με προσανατολισμό στις εισροές (input-oriented) και υπό την παραδοχή σταθερών αποδόσεων κλίμακας (Constant Returns to Scale - CRS), όπως αυτό ορίζεται στο κλασικό μοντέλο των Charnes, Cooper και Rhodes (CCR). Η επιλογή αυτή ανταποκρίνεται στη φύση του υπό μελέτη προβλήματος, όπου στόχο αποτελεί η ελαχιστοποίηση της κατανάλωσης πόρων διατηρώντας αμετάβλητο το επίπεδο παρεχόμενης υπηρεσίας.

Η ανάλυση DEA πραγματοποιήθηκε μέσω του λογισμικού DEAP, χρησιμοποιώντας ένα σύνολο δεδομένων που αφορά δεκατρείς ευρωπαϊκούς σιδηροδρομικούς εμπορευματικούς οργανισμούς:

- DB Cargo - Γερμανία
- Hexafret (πρώην SNCF Fret) - Γαλλία
- Renfe Mercancías - Ισπανία
- CD Cargo - Τσεχία
- OBB Rail Cargo Group - Αυστρία
- ZSSK Cargo - Σλοβακία
- Hupac - Ελβετία
- Green Cargo - Σουηδία
- PKP Cargo - Πολωνία
- VR Transpoint - Φινλανδία
- Rail Cargo Hungaria Zrt - Ουγγαρία
- Mercitalia Rail - Ιταλία
- DB Cargo Netherlands - Ολλανδία

Οι μεταβλητές που χρησιμοποιήθηκαν ως εισροές περιλαμβάνουν κρίσιμους πόρους, όπως το μήκος του σιδηροδρομικού δικτύου, τον αριθμό του τροχαίου υλικού, τον αριθμό των εργαζομένων, τις ετήσιες κεφαλαιουχικές δαπάνες, τα λειτουργικά κόστη και την ενεργειακή κατανάλωση. Ενώ, ως εκροές επιλέχθηκαν ο όγκος μεταφοράς σε τονοχιλιόμετρα (billion ton-kilometres) και τα ετήσια έσοδα, με σκοπό την ταυτόχρονη αξιολόγηση τόσο της τεχνικής όσο και της οικονομικής αποδοτικότητας.

Στο δεύτερο στάδιο, αναπτύχθηκαν και εκπαιδεύτηκαν δύο νευρωνικά δίκτυα τεχνητής νοημοσύνης (το πρώτο αντιστοιχεί στην τεχνική και το δεύτερο στην οικονομική αποδοτικότητα) με στόχο την υπέρβαση ορισμένων εγγενών περιορισμών της DEA, όπως η στατικότητα και η αδυναμία (ρεαλιστικής) πρόβλεψης. Το ANN μοντέλο σχεδιάστηκε ώστε να λαμβάνει ως είσοδο (input) το μήκος δικτύου (X1) που σε κάθε περίπτωση παραμένει σταθερό, την επιλεγμένη εκροή (είτε τα τονοχιλιόμετρα είτε τα έσοδα) καθώς και τη βαθμολογία αποδοτικότητας (θ) που προκύπτει από τη DEA. Παράλληλα, προβλέπει ως έξοδο (output) ένα σύνολο πέντε βελτιστοποιημένων εισροών (X2–X6), προσαρμοσμένων στις συνθήκες λειτουργίας κάθε οργανισμού. Για την εκπαίδευση του μοντέλου δημιουργήθηκε ένα εκτενές σύνολο δεδομένων μέσω συστηματικής μεταβολής των αρχικών εισροών, με ελεγχόμενη κλιμάκωση και διατήρηση της εξόδου, το οποίο επανειλημμένα αξιολογήθηκε μέσω DEA ώστε να αποδοθεί σε κάθε εγγραφή η αντίστοιχη τιμή αποδοτικότητας.

Η προσέγγιση αυτή κατέστησε δυνατή την κατασκευή ενός μοντέλου πρόβλεψης που δεν περιορίζεται στο αρχικό δείγμα, αλλά δύναται να προτείνει νέους, πιο ρεαλιστικούς και πρακτικά εφαρμόσιμους συνδυασμούς εισροών, ικανούς να μεγιστοποιήσουν την αποδοτικότητα σε μελλοντικά ή εναλλακτικά σενάρια λειτουργίας. Με τον τρόπο αυτόν, δημιουργήθηκε ένα εργαλείο στρατηγικού σχεδιασμού και υποστήριξης στη λήψη αποφάσεων για τους υπεύθυνους των σιδηροδρομικών οργανισμών, το οποίο συμπληρώνει και ενισχύει τη χρηστικότητα των παραδοσιακών μεθόδων ανάλυσης της αποδοτικότητας. Η εφαρμογή του μοντέλου δύο σταδίων DEA-ANN αποκάλυψε κρίσιμες διαφορές ως προς τα επίπεδα αποδοτικότητας μεταξύ των υπό μελέτη σιδηροδρομικών οργανισμών. Μέσω της DEA, έγινε εφικτός ο εντοπισμός τόσο πλήρως αποδοτικών μονάδων όσο και μονάδων με σημαντικά περιθώρια βελτίωσης, παρέχοντας σαφή εικόνα της επιχειρησιακής απόδοσης στον ευρωπαϊκό

σιδηροδρομικό χώρο. Τα αποτελέσματα κατέδειξαν ότι, σε αρκετές περιπτώσεις, οι μη αποδοτικοί οργανισμοί διέθεταν υπερβάλλουσες εισροές σε βασικούς επιχειρησιακούς πόρους, γεγονός που υποδηλώνει δυνατότητες εξορθολογισμού και ανακατανομής. Ωστόσο, οι προτάσεις της DEA για βελτίωση βασίζονται σε θεωρητικά βέλτιστα αλλά συχνά μη ρεαλιστικά σενάρια, τα οποία απαιτούν δυσανάλογες μειώσεις πόρων.

Σε πλαίσιο αυτό, η συνεισφορά των ANN αποδείχθηκε καθοριστική. Το νευρωνικό δίκτυο κατάφερε να ενσωματώσει την πολυπλοκότητα των δεδομένων και να προτείνει πιο ισορροπημένες και εφαρμόσιμες λύσεις, προβλέποντας συνδυασμούς εισροών που είναι προσαρμοσμένοι στις πραγματικές συνθήκες λειτουργίας κάθε οργανισμού. Η επανεκτίμηση αυτών των προβλέψεων μέσω DEA έδειξε σαφή βελτίωση της αποδοτικότητας, γεγονός που επιβεβαιώνει την πρακτική αξία του μοντέλου. Η ενσωμάτωση της προβλεπτικής ικανότητας των ANN στη διαδικασία ανάλυσης δημιούργησε ένα ευέλικτο και δυναμικό πλαίσιο, ικανό όχι μόνο να αποτιμήσει την υφιστάμενη κατάσταση αλλά και να υποδείξει εφαρμόσιμες πρακτικές βελτίωσης με προγνωστικό χαρακτήρα.

Η παρούσα εργασία συνεισφέρει πολλαπλώς στη μελέτη της αποδοτικότητας στον τομέα των σιδηροδρομικών εμπορευματικών μεταφορών, προσφέροντας ένα ολοκληρωμένο μεθοδολογικό πλαίσιο το οποίο συνδυάζει τη μεθοδολογική συνέπεια της DEA με τη δυναμική προσαρμοστικότητα των ANN. Η καινοτομία του προτεινόμενου υποδείγματος έγκειται στην ικανότητά του να μην περιορίζεται σε μία στατική αποτύπωση της αποδοτικότητας, αλλά να παρέχει ρεαλιστικές και εξατομικευμένες προτάσεις βελτιστοποίησης, βασισμένες σε εμπειρικά δεδομένα και προγνωστικά πρότυπα.

Παρά τα ενθαρρυντικά αποτελέσματα, η μεθοδολογία δεν στερείται περιορισμών. Ο αριθμός των εξεταζόμενων οργανισμών, αν και αντιπροσωπευτικός, παραμένει σχετικά μικρός, γεγονός που ενδέχεται να επηρεάζει τη γενικευσιμότητα των συμπερασμάτων. Επιπλέον, η προσέγγιση επικεντρώνεται αποκλειστικά σε θετικές εκροές (όπως όγκος μεταφοράς και έσοδα), χωρίς να ενσωματώνει αρνητικούς δείκτες απόδοσης, όπως περιβαλλοντικές επιπτώσεις ή καθυστερήσεις.

Μελλοντικά, προτείνεται η περαιτέρω ενίσχυση του μοντέλου μέσω εμπλουτισμού του συνόλου δεδομένων, ενσωμάτωσης ανεπιθύμητων εκροών, και αξιοποίησης εξελιγμένων τεχνικών τεχνητής νοημοσύνης, όπως γενετικοί αλγόριθμοι ή ενισχυτική μάθηση (reinforcement learning). Επιπλέον, το συνδυαστικό μοντέλο θα μπορούσε να εφαρμοστεί και σε άλλους τομείς μεταφορών ή δημόσιων υπηρεσιών, προσφέροντας πολύτιμο εργαλείο για την αξιολόγηση και στρατηγική βελτίωση της αποδοτικότητας.

Συνοψίζοντας, η εργασία αυτή ανέδειξε τη σημασία της χρήσης καινοτόμων υποδειγμάτων για την αξιολόγηση και τη βελτιστοποίηση της αποδοτικότητας σε κρίσιμους τομείς, όπως οι σιδηροδρομικές εμπορευματικές μεταφορές. Ο συνδυασμός DEA και ANN αποδείχθηκε ιδιαίτερα αποτελεσματικός, όχι μόνο για την αποτύπωση της υφιστάμενης κατάστασης, αλλά και για την παροχή προτάσεων προσαρμοσμένων στις πραγματικές ανάγκες των οργανισμών. Η προσέγγιση αυτή ενισχύει τη δυνατότητα των υπευθύνων λήψης αποφάσεων να σχεδιάζουν στοχευμένες στρατηγικές βελτίωσης και συμβάλλει, μακροπρόθεσμα, στην επίτευξη βιώσιμων και ανταγωνιστικών σιδηροδρομικών δικτύων στην Ευρώπη και τον κόσμο.

Περιεχόμενα

1. Εισαγωγή	14
1.1 Γενικά στοιχεία	14
1.2 Σκοπός της διπλωματικής εργασίας	16
1.3 Διάρθρωση της διπλωματικής εργασίας	17
2. Βιβλιογραφική ανασκόπηση	19
2.1 Εισαγωγή	19
2.2 Εφαρμογή του μοντέλου DEA στις σιδηροδρομικές μεταφορές	19
2.3 Διεθνή πλαίσια για εκτίμηση της αποδοτικότητας	22
2.4 Συνδυαστική εφαρμογή μοντέλων DEA και AI	24
2.5 Συμπεράσματα βιβλιογραφίας	27
3. Μεθοδολογική προσέγγιση	29
3.1 Περιγραφή διαδικασίας	29
3.2 Θεωρητικό υπόβαθρο της Περιβάλλουσας Ανάλυσης Δεδομένων (Data Envelopment Analysis – DEA)	30
3.2.1 Μαθηματική Διατύπωση της DEA με παραδοχή Σταθερών Αποδόσεων Κλίμακας (CRS)	31
3.2.2 Υπολειπόμενες ποσότητες (Slacks) και τεχνική αποδοτικότητα	33
3.2.3 Προσεγγίσεις επίλυσης DEA με υπολογισμό Slacks	34
3.2.4 Υλοποίηση του υποδείγματος DEA στην πράξη (Λογισμικό DEAP)	35
3.2.5 Πλεονεκτήματα και Μειονεκτήματα της Μεθόδου Περιβάλλουσας Ανάλυσης Δεδομένων (DEA)	36
3.3 Θεωρητικό Υπόβαθρο των Τεχνητών Νευρωνικών Δικτύων (ANN)	36
3.3.1 Βασικές Αρχές και Δομή Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks – ANN)	37
3.3.2 Τύποι Νευρωνικών Δικτύων και Επιλογή Κατάλληλου Μοντέλου	37
3.3.3 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου και Εφαρμογή στη DEA	38
3.3.4 Πλεονεκτήματα και Περιορισμοί της ANN μεθόδου	39
4. Ανάλυση και αποτελέσματα	41
4.1 Βάση δεδομένων	41
4.1.1 Επιλογή οργανισμών-χωρών	41
4.1.2 Επιλογή δεδομένων	42
4.1.3 Επιλογή εισροών και εκροών	42
4.1.4 Παραδοχές-Υπολογισμοί	43
4.2 Σύνθεση Δεδομένων και Περιγραφή Εφαρμογής	49
4.2.1 Πίνακας Δεδομένων	49
4.2.2 Εφαρμογή του Λογισμικού DEAP	52

4.3	Παρουσίαση και Ανάλυση Αποτελεσμάτων	55
4.3.1	Τεχνική Αποδοτικότητα (Output: Ton-kilometres)	55
4.3.2	Οικονομική Αποδοτικότητα (Output: Revenue)	56
4.3.3	Προτεινόμενες Βελτιώσεις από το μοντέλο DEA.....	57
4.3.4	Προτεινόμενα Επίπεδα Εισροών για Βελτίωση της Τεχνικής και της Οικονομικής Αποδοτικότητας.....	57
4.4	Τεχνητά Νευρωνικά Δίκτυα (ANN).....	60
4.4.1	Δημιουργία Συνόλου Δεδομένων	60
4.4.2	Υπολογισμός της DEA Αποδοτικότητας (θ).....	62
4.4.3	Εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου (ANN).....	64
4.4.4	Βελτιστοποίηση και Επαλήθευση μέσω ANN.....	65
5.	Συμπεράσματα	77
5.1	Γενικά	77
5.2	Βασικά συμπεράσματα	77
5.3	Προτάσεις για περαιτέρω έρευνα	79
6.	Βιβλιογραφία	80
7.	Παραρτήματα.....	87

Ευρετήριο πινάκων

Πίνακας 4.1 Οργανισμοί που επιλέχθηκαν και οι αντίστοιχες χώρες	42
Πίνακας 4.2 Βάση δεδομένων	50
Πίνακας 4.3 Στατιστικά των Μεταβλητών	51
Πίνακας 4.4 Τεχνική Αποδοτικότητα με την μέθοδο DEA	56
Πίνακας 4.5 Οικονομική Αποδοτικότητα με την μέθοδο DEA	57
Πίνακας 4.6 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Τεχνικής Αποδοτικότητας DEAP	58
Πίνακας 4.7 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Οικονομικής Αποδοτικότητας DEAP	59
Πίνακας 4.8 Αποτελέσματα της εφαρμογής του μοντέλου ANN για Τεχνική Αποδοτικότητα	69
Πίνακας 4.9 Αποτελέσματα της εφαρμογής του μοντέλου ANN για Οικονομική Αποδοτικότητα	69
Πίνακας 4.10 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Τεχνικής Αποδοτικότητας ANN	70
Πίνακας 4.11 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Οικονομικής Αποδοτικότητας ANN	71
Πίνακας 4.12 Ποσοστιαία διαφορά μεταβολής μεταβλητών DEA-ANN για Τεχνική Αποδοτικότητα	73
Πίνακας 4.13 Ποσοστιαία διαφορά μεταβολής μεταβλητών DEA-ANN για Οικονομική Αποδοτικότητα	74

Ευρετήριο εικόνων

Εικόνα 3.1 Διάγραμμα ροής της μεθοδολογικής διαδικασίας	29
Εικόνα 3.2 Απεικόνιση αποδοτικού μετώπου και περιβάλλοντος συνόλου σε περίπτωση δύο εισροών και μίας εκροής (Πηγή: Farrell, M.J. (1957))	31
Εικόνα 4.1 Αρχείο κειμένου με εκροή τον όγκο μεταφοράς	53
Εικόνα 4.2 Αρχείο κειμένου με εκροή τον τζίρο	53
Εικόνα 4.3 Αρχείο κειμένου οδηγιών (instruction file)	54
Εικόνα 4.4 systematic_dataset_all_firms_sampled_half.xlsx	61
Εικόνα 4.5 calculate_theta_sampled_half.py	63
Εικόνα 4.6 train_ann_with_theta.py	65
Εικόνα 4.7 eval_ann_with_theta_model_REVENUE_FINAL.py	68

1. Εισαγωγή

1.1 Γενικά στοιχεία

Η αποδοτικότητα των σιδηροδρομικών εμπορευματικών μεταφορών αποτελεί καθοριστικό παράγοντα για την οικονομική και περιβαλλοντικά βιώσιμη ανάπτυξη των εφοδιαστικών δικτύων ανά την Ευρώπη και τον κόσμο. Σε μία εποχή όπου η ανάγκη για μείωση των εκπομπών διοξειδίου του άνθρακα και η ενίσχυση των σιδηροδρόμων έναντι των οδικών μεταφορών καθίσταται στρατηγική προτεραιότητα, η ανάγκη για εμπειριστατωμένη αξιολόγηση και βελτιστοποίηση των επιχειρησιακών πρακτικών των σιδηροδρομικών φορέων γίνεται επιτακτική. Η στροφή προς πιο φιλικά προς το περιβάλλον μεταφορικά μέσα, σε συνδυασμό με τον εντεινόμενο ανταγωνισμό και τις τεχνολογικές εξελίξεις, δημιουργεί ένα περιβάλλον στο οποίο η μελέτη της αποδοτικότητας των σιδηροδρομικών εμπορευματικών οργανισμών μεταφορών αποκτά ιδιαίτερη σημασία τόσο για τη λήψη αποφάσεων όσο και για τον στρατηγικό σχεδιασμό πολιτικής.

Η έννοια της αποδοτικότητας συνδέεται στενά με την ικανότητα ενός οργανισμού να αξιοποιεί τους διαθέσιμους πόρους του με τον καλύτερο δυνατό τρόπο για την παραγωγή αποτελεσμάτων. Στον χώρο των σιδηροδρομικών μεταφορών, η αποδοτικότητα αποκτά ιδιαίτερη βαρύτητα, καθώς οι οργανισμοί καλούνται να διασφαλίσουν υψηλού επιπέδου υπηρεσίες με περιορισμένους πόρους, αυστηρά χρονοδιαγράμματα και σημαντικούς περιβαλλοντικούς περιορισμούς. Παράλληλα, οι επενδύσεις σε υποδομές, τροχαίο υλικό και ανθρώπινο δυναμικό απαιτούν διαρκή αποτίμηση του βαθμού αξιοποίησής τους. Για τον λόγο αυτό, προκύπτει η ανάγκη χρήσης κατάλληλων μεθοδολογιών εκτίμησης της αποδοτικότητας, οι οποίες να είναι ικανές να λάβουν υπόψη τη φύση των δεδομένων, την πολυπλοκότητα του συστήματος και την πολλαπλότητα των εισροών και εκροών.

Μια από τις πιο διαδεδομένες μεθόδους σε αυτό το πλαίσιο αποτελεί η Μέθοδος της Περιβάλλουσας Ανάλυσης Δεδομένων (DEA), η οποία παρέχει τη δυνατότητα συγκριτικής αξιολόγησης ομοειδών μονάδων λήψης αποφάσεων (DMUs) χωρίς την ανάγκη προσδιορισμού παραμετρικής μορφής παραγωγής. Η μη παραμετρική φύση της DEA καθώς και η ικανότητά της να χειρίζεται πολλαπλές εισροές και εκροές συγχρόνως την καθιστούν ιδανική για εφαρμογές στον τομέα των μεταφορών και, ειδικότερα, στον κλάδο των σιδηροδρομικών εμπορευματικών οργανισμών.

Η DEA έχει καθιερωθεί ως μια σημαντική μη παραμετρική μέθοδος για την εκτίμηση της αποδοτικότητας ομοειδών μονάδων. Η καινοτομία της μεθόδου έγκειται στο γεγονός ότι δεν απαιτεί την εκ των προτέρων γνώση της μορφής της παραγωγικής συνάρτησης, πράγμα το οποίο συμβαίνει με τις παραμετρικές προσεγγίσεις, όπως η Στοχαστική Ανάλυση Αποδοτικότητας (Stochastic Frontier Analysis – SFA). Η DEA στηρίζεται σε γραμμικό προγραμματισμό και προσδιορίζει το αποδοτικό όριο (efficient frontier) βάσει εμπειρικών δεδομένων, με σκοπό τη μέτρηση της σχετικής αποδοτικότητας κάθε μονάδας σε σχέση με τις βέλτιστες παρατηρήσεις του δείγματος.

Σε αντίθεση, η SFA εισάγει έναν στοχαστικό όρο σφάλματος και προϋποθέτει συγκεκριμένη μορφή συνάρτησης (π.χ. Cobb-Douglas ή Translog), γεγονός που την καθιστά περισσότερο επιρρεπή σε υποκειμενικές υποθέσεις. Επιπλέον, μεθοδολογίες όπως οι αναλύσεις κόστους-οφέλους, οι δείκτες παραγωγικότητας και οι συμβατικές τεχνικές benchmarking, αν και χρήσιμες σε ορισμένα πλαίσια, αδυνατούν να ανταποκριθούν στην πολυδιάστατη φύση της αποδοτικότητας, ιδίως όταν εμπλέκονται πολλαπλές εισροές και εκροές.

Η DEA προσφέρεται ως ένα εργαλείο που ενσωματώνει διαφορετικές ποσοτικές μεταβλητές χωρίς την ανάγκη στάθμισης ή μετατροπής τους σε κοινή μονάδα μέτρησης. Το γεγονός αυτό την καθιστά κατάλληλη για την εκτίμηση της αποδοτικότητας στον σιδηροδρομικό τομέα, όπου οι παράγοντες που επηρεάζουν την απόδοση (όπως για παράδειγμα μήκος δικτύου, ανθρώπινο δυναμικό, ενεργειακή κατανάλωση και τροχαίο υλικό) είναι δύσκολα συγκρίσιμοι μεταξύ τους με άλλες παραδοσιακές τεχνικές.

Παρά τα σημαντικά πλεονεκτήματά της, η Μέθοδος της Περιβάλλουσας Ανάλυσης Δεδομένων παρουσιάζει ορισμένους περιορισμούς που θέτουν όρια στην εφαρμογή της σε πιο σύνθετα ή δυναμικά περιβάλλοντα. Αρχικά, η DEA είναι εξ ορισμού μια συγκριτική μέθοδος, γεγονός που σημαίνει ότι η αποδοτικότητα κάθε μονάδας προσδιορίζεται σε σχέση με τις υπόλοιπες μονάδες του δείγματος. Κατά συνέπεια, το αποτέλεσμα είναι άρρηκτα εξαρτημένο από τη σύνθεση του δείγματος και δεν παρέχει απόλυτη μέτρηση αποδοτικότητας. Επιπλέον, η DEA αν και εντοπίζει ποιοι συνδυασμοί εισροών είναι αποδοτικοί, δεν μπορεί να γενικεύσει ή να προτείνει εισροές πέρα από το υπάρχον δείγμα, ούτε να προβλέψει τη συμπεριφορά μιας νέας μονάδας με διαφορετικά χαρακτηριστικά. Ακόμη, είναι ιδιαίτερα ευαίσθητη σε ακραίες τιμές (outliers), καθώς αυτές μπορούν να επηρεάσουν την τοποθέτηση του αποδοτικού μετώπου και επομένως να αλλοιώσουν την αξιολόγηση των υπολοίπων μονάδων.

Τα τελευταία χρόνια, η ραγδαία εξέλιξη της τεχνητής νοημοσύνης και των αλγορίθμων μηχανικής μάθησης έχει φέρει σημαντικές αλλαγές στον τρόπο με τον οποίο αναλύονται και αξιοποιούνται εκτός των άλλων και στο πεδίο των μεταφορών. Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN), τα οποία εμπνέονται από τη λειτουργία του ανθρώπινου εγκεφάλου και επιτρέπουν την προσομοίωση πολύπλοκων συσχετίσεων μεταξύ μεταβλητών. Τα ANN έχουν τη δυνατότητα να "μαθαίνουν" από τα δεδομένα μέσω μιας διαδικασίας εποπτευόμενης εκπαίδευσης, κατά την οποία αναγνωρίζουν πρότυπα και βελτιώνουν σταδιακά την απόδοσή τους. Η ευελιξία τους αυτή τα έχει καταστήσει εξαιρετικά χρήσιμα σε προβλήματα πρόβλεψης, ταξινόμησης και αναγνώρισης προτύπων, ενώ βρίσκουν εφαρμογή σε πλήθος τομέων, από την ιατρική διάγνωση μέχρι τη βελτιστοποίηση παραγωγικών συστημάτων.

Σε αντίθεση με τις παραδοσιακές μεθόδους στατιστικής ή γραμμικού προγραμματισμού, τα Τεχνητά Νευρωνικά Δίκτυα διακρίνονται για την ικανότητά τους να αντιμετωπίζουν προβλήματα υψηλής πολυπλοκότητας, όπου οι σχέσεις μεταξύ των μεταβλητών δεν είναι απαραίτητα γραμμικές, ούτε εκ των προτέρων γνωστές. Τα ANN αποτελούνται από διασυνδεδεμένους κόμβους (νευρώνες) διατεταγμένους σε επίπεδα (layers), μέσα από τα οποία τα δεδομένα υφίστανται διαδοχικούς μετασχηματισμούς. Μέσω αυτής της δομής, ένα νευρωνικό δίκτυο μπορεί να αναγνωρίσει πρότυπα, να εξάγει κρυμμένες συσχετίσεις και να διαχειριστεί ακόμα και δεδομένα με υψηλό βαθμό θορύβου ή αστάθειας.

Επιπλέον, ένα καλά εκπαιδευμένο ANN μπορεί να γενικεύσει τη γνώση του και να προβλέψει την απόδοση ή τη συμπεριφορά νέων παρατηρήσεων, ακόμη και αν αυτές δεν υπήρχαν στο αρχικό σύνολο εκπαίδευσης. Αυτή η δυνατότητα το καθιστά ιδανικό εργαλείο για προβλήματα όπου στόχος δεν είναι μόνο η κατανόηση του παρελθόντος, αλλά και η πρόβλεψη ή η προσομοίωση μελλοντικών σεναρίων. Για τον λόγο αυτό, τα ANN έχουν ενταχθεί δυναμικά στο πεδίο της επιχειρησιακής έρευνας, της βιομηχανίας, των χρηματοοικονομικών, της ενέργειας, και οπωσδήποτε - τα τελευταία χρόνια - στον τομέα των μεταφορών και της αξιολόγησης της αποδοτικότητας.

Η συμπληρωματική χρήση των Τεχνητών Νευρωνικών Δικτύων με τη Μέθοδο της Περιβάλλουσας Ανάλυσης Δεδομένων προκύπτει ως μία φυσική εξέλιξη των σύγχρονων αναγκών για την εκτίμηση της αποδοτικότητας και προσαρμοστικότητα στις μεταβαλλόμενες συνθήκες του πραγματικού κόσμου. Όπως προαναφέρθηκε, η DEA, αν και ισχυρό διαγνωστικό εργαλείο, παρέχει αποκλειστικά στατικά αποτελέσματα βασισμένα στα δεδομένα υπό παρατήρηση, χωρίς προβλεπτική ικανότητα. Επιπλέον, οι εισροές που προτείνει σε μη αποδοτικές μονάδες συχνά είναι μη ρεαλιστικές ή απαιτητικές, καθώς βασίζονται σε υποθετική σύγκλιση με τις πλέον αποδοτικές μονάδες του δείγματος, χωρίς να λαμβάνεται υπόψη η πραγματικότητα ή οι φυσικοί περιορισμοί κάθε οργανισμού.

Τα ANN έρχονται να καλύψουν αυτό το κενό, παρέχοντας τη δυνατότητα προσομοίωσης της λογικής της DEA, αλλά και την πρόβλεψη εναλλακτικών, πιο εφαρμόσιμων συνδυασμών εισροών, με διατήρηση του επιπέδου εκροής. Μέσω της εκπαίδευσης σε δεδομένα που έχουν παραχθεί από την ίδια τη DEA - όπως βαθμοί αποδοτικότητας και εισροές στόχοι - τα ANN μπορούν να μάθουν τις υποκείμενες σχέσεις και να παράγουν αποδοτικά σενάρια εισροών, ακόμα και για περιπτώσεις που δεν περιλαμβάνονται στο αρχικό δείγμα. Έτσι, ο συνδυασμός των δύο μεθόδων παρέχει ένα πιο ισχυρό εργαλείο: η DEA αξιολογεί και εντοπίζει αδυναμίες, ενώ το ANN προτείνει εφαρμόσιμες λύσεις που μπορούν να επανεξεταστούν εκ νέου μέσω της πρώτης, δημιουργώντας έναν δυναμικό κύκλο βελτιστοποίησης.

Συνοψίζοντας, η ανάγκη για βελτιστοποίηση της αποδοτικότητας των σιδηροδρομικών οργανισμών σε ένα συνεχώς μεταβαλλόμενο, απαιτητικό και ανταγωνιστικό περιβάλλον καθιστά επιτακτική τη χρήση προηγμένων εργαλείων ανάλυσης και υποστήριξης αποφάσεων. Η DEA προσφέρει μια ισχυρή βάση για την κατανόηση των υφιστάμενων επιπέδων αποδοτικότητας και τον εντοπισμό σημείων βελτίωσης, ενώ τα Τεχνητά Νευρωνικά Δίκτυα προσδίδουν επιπλέον δυναμική, επιτρέποντας την πρόβλεψη, την προσαρμογή και τη μοντελοποίηση εναλλακτικών σεναρίων. Ο συνδυασμός αυτών των δύο προσεγγίσεων δημιουργεί ένα ολοκληρωμένο πλαίσιο εκτίμησης και βελτιστοποίησης, το οποίο διερευνάται αναλυτικά στο πλαίσιο της παρούσας διπλωματικής εργασίας.

1.2 Σκοπός της διπλωματικής εργασίας

Η παρούσα διπλωματική εργασία έχει ως βασικό στόχο την εκτίμηση της αποδοτικότητας σιδηροδρομικών οργανισμών εμπορευματικών μεταφορών στην Ευρώπη, καθώς και τη διερεύνηση ρεαλιστικών προσεγγίσεων που αποσκοπούν στη βελτιστοποίηση αυτής, μέσω της εφαρμογής συνδυασμένων μεθοδολογιών επιχειρησιακής έρευνας και τεχνητής νοημοσύνης. Πιο συγκεκριμένα, επιδιώκεται αρχικά η εκτίμηση της τεχνικής και οικονομικής αποδοτικότητας επιλεγμένων σιδηροδρομικών οργανισμών, αρχικά με τη χρήση και εφαρμογή του μοντέλου Data Envelopment Analysis (DEA), υπό την παραδοχή σταθερών αποδόσεων κλίμακας (CRS) και με προσανατολισμό στις εισροές (input-oriented). Η εκτίμηση αυτή βασίζεται σε σύνολα δεδομένων που περιλαμβάνουν βασικές λειτουργικές μεταβλητές, όπως το μήκος του σιδηροδρομικού δικτύου, το τροχαίο υλικό, το ανθρώπινο δυναμικό, οι κεφαλαιουχικές δαπάνες, το κόστος λειτουργίας και η κατανάλωση ενέργειας, σε σχέση με εκροές όπως τα έσοδα και το μεταφορικό έργο.

Σε δεύτερο στάδιο, επιχειρείται η ενίσχυση της DEA μέσω της αξιοποίησης Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks – ANN), τα οποία εκπαιδεύονται με

βάση τα αποτελέσματα της πρώτης μεθόδου, ώστε να προτείνουν ρεαλιστικά σενάρια βελτιστοποίησης των εισροών για τις μη αποδοτικές μονάδες. Η μεθοδολογία αυτή δεν περιορίζεται μόνο στην εκ των υστέρων εκτίμηση, αλλά επιδιώκει την ανάπτυξη ενός προβλεπτικού μηχανισμού υποστήριξης αποφάσεων, που μπορεί να συμβάλει στον επιχειρησιακό σχεδιασμό σιδηροδρομικών φορέων και τη χάραξη πολιτικών. Επιπλέον, εξετάζονται εναλλακτικές προσεγγίσεις κατασκευής dataset, με στόχο την εκπαίδευση και γενίκευση των μοντέλων ANN σε υποθετικά σενάρια.

Μέσα από την εφαρμογή του μοντέλου δύο σταδίων DEA-ANN, η εν λόγω διπλωματική εργασία φιλοδοξεί να συμβάλει στην επιστημονική βιβλιογραφία με μια καινοτόμο εφαρμογή μηχανικής μάθησης σε περιβάλλοντα αξιολόγησης της αποδοτικότητας, αναδεικνύοντας τη δυναμική των συνδυασμένων μεθόδων στην πράξη.

1.3 Διάρθρωση της διπλωματικής εργασίας

Η παρούσα διπλωματική εργασία αποτελείται από έξι κεφάλαια και πλαισιώνεται από βιβλιογραφικές αναφορές και παραρτήματα. Η δομή της ακολουθεί λογική αλληλουχία ώστε να διευκολυνθεί η κατανόηση των εννοιών και της μεθοδολογικής προσέγγισης που ακολουθήθηκε, και να παρουσιαστούν συστηματικά τόσο το θεωρητικό υπόβαθρο όσο και τα αποτελέσματα της ανάλυσης.

Το παρόν κεφάλαιο αποτελεί την εισαγωγή στην εργασία. Εδώ παρουσιάζονται το γενικό πλαίσιο και η σημασία του υπό μελέτη αντικειμένου, με αναφορά στη σπουδαιότητα της αποδοτικότητας στον τομέα των σιδηροδρομικών εμπορευματικών μεταφορών. Γίνεται εκτενής αναφορά στις υπάρχουσες μεθοδολογίες καθώς και του περιορισμούς αυτών και αναδεικνύεται η αναγκαιότητα αξιοποίησης συνδυαστικών προσεγγίσεων, όπως ο συνδυασμός μεθόδων DEA και ANN. Έπειτα, περιγράφεται ο σκοπός της διπλωματικής, και τέλος, παρουσιάζεται, όπως ακολουθεί, η συνοπτική διάρθρωση του περιεχομένου της.

Στο δεύτερο κεφάλαιο παρατίθεται η βιβλιογραφική ανασκόπηση. Καταγράφονται προηγούμενες μελέτες και άρθρα που σχετίζονται με την εκτίμηση της αποδοτικότητας στις μεταφορές και παρουσιάζονται διαφορετικές μεθοδολογικές προσεγγίσεις που έχουν χρησιμοποιηθεί για τον ίδιο σκοπό. Ειδική έμφαση δίνεται σε εφαρμογές της DEA και σε περιπτώσεις όπου συνδυάζεται με τη τεχνητή νοημοσύνη, καθώς και σε σύγχρονες ερευνητικές προσπάθειες που χρησιμοποιούν τα ANN στον τομέα των επιχειρησιακών προβλημάτων.

Το τρίτο κεφάλαιο αφορά τη μεθοδολογική προσέγγιση που υιοθετήθηκε. Αρχικά περιγράφεται η μαθηματική διατύπωση της DEA με παραδοχή σταθερών αποδόσεων κλίμακας και προσανατολισμό στις εισροές, ενώ παρουσιάζεται και το θεωρητικό υπόβαθρο των τεχνητών νευρωνικών δικτύων. Εν συνεχεία εξηγείται η λογική πίσω από το συνδυασμό των δύο μεθόδων, η προσέγγιση δημιουργίας συστηματικού συνόλου δεδομένων και η διαδικασία εκπαίδευσης των μοντέλων ANN. Το κεφάλαιο ολοκληρώνεται με περιγραφή των λογισμικών εργαλείων και των παραμέτρων που χρησιμοποιήθηκαν.

Στο τέταρτο κεφάλαιο παρουσιάζεται η εφαρμογή της μεθοδολογίας στα δεδομένα των σιδηροδρομικών οργανισμών. Αρχικά, περιγράφεται ο τρόπος συλλογής και

καθαρισμού των δεδομένων, η κατασκευή των αρχείων εισόδου και η διαδικασία εξαγωγής των αποτελεσμάτων μέσω χρήσης λογισμικού. Έπειτα εξηγείται η διαδικασία κατασκευής των dataset για τα ANN, η εκπαίδευση των μοντέλων, καθώς και τα βήματα που εφαρμόστηκαν για τη χρήση τους ως εργαλεία πρόβλεψης βελτιστοποιημένων εισροών. Τέλος, γίνεται σύγκριση μεταξύ των αποτελεσμάτων DEA και ANN.

Το πέμπτο κεφάλαιο περιλαμβάνει τη σύνθεση των βασικών ευρημάτων και τη διατύπωση των συμπερασμάτων της παρούσας μελέτης. Αρχικά παρουσιάζεται μια συνολική αποτίμηση της έρευνας και των μεθοδολογικών επιλογών που ακολουθήθηκαν. Στη συνέχεια αναλύονται τα κύρια συμπεράσματα που προέκυψαν από τη συγκριτική εφαρμογή των μεθόδων DEA και ANN, με έμφαση στις πρακτικές προεκτάσεις τους στον κλάδο των σιδηροδρομικών εμπορευματικών μεταφορών. Τέλος, διατυπώνονται προτάσεις για μελλοντική έρευνα και διεύρυνση της παρούσας προσέγγισης.

Στο έκτο κεφάλαιο περιλαμβάνεται η βιβλιογραφία που αξιοποιήθηκε κατά τη συγγραφή της εν λόγω διπλωματικής εργασίας, με στόχο τη θεμελίωση του θεωρητικού πλαισίου και την τεκμηρίωση των μεθοδολογικών και ερευνητικών επιλογών.

Τέλος, στο έβδομο και τελευταίο κεφάλαιο παρατίθενται παραρτήματα με τον πηγαίο κώδικα που χρησιμοποιήθηκε για την υλοποίηση της μεθοδολογίας DEA–ANN, παρέχοντας τη δυνατότητα αναπαραγωγής των αποτελεσμάτων και περαιτέρω διερεύνησης της προτεινόμενης προσέγγισης.

2. Βιβλιογραφική ανασκόπηση

2.1 Εισαγωγή

Η εκτίμηση της αποδοτικότητας των σιδηροδρομικών οργανισμών εμπορευματικών μεταφορών αποτελεί αντικείμενο εκτενούς μελέτης στη διεθνή βιβλιογραφία, κυρίως λόγω της πολυπλοκότητας του τομέα, της στρατηγικής του σημασίας για τις εθνικές οικονομίες και της ανάγκης για ενίσχυση της βιωσιμότητας των μεταφορών. Στο πλαίσιο αυτό, έχουν αναπτυχθεί και εφαρμοστεί πλήθος μεθοδολογιών, τόσο παραδοσιακών όσο και καινοτόμων, για την αποτίμηση της αποτελεσματικότητας των μεταφορικών συστημάτων και την υποστήριξη της στρατηγικής λήψης αποφάσεων.

Η παρούσα βιβλιογραφική ανασκόπηση εστιάζει στις κυριότερες επιστημονικές προσεγγίσεις που έχουν εφαρμοστεί για την εκτίμηση της αποδοτικότητας στον τομέα των σιδηροδρομικών μεταφορών. Ιδιαίτερη έμφαση δίνεται στην Περιβάλλουσα Ανάλυση Δεδομένων (DEA) και τις επεκτάσεις αυτής, καθώς και τις υβριδικές μεθόδους που συνδυάζουν την DEA με αλγορίθμους Τεχνητής Νοημοσύνης (AI), όπως τα Τεχνητά Νευρωνικά Δίκτυα (ANN). Επιπλέον, καταγράφονται παραδείγματα εφαρμογής των παραπάνω μεθοδολογιών σε ποικίλα γεωγραφικά και λειτουργικά πλαίσια, επισημαίνοντας τόσο τις δυνατότητές όσο και τους περιορισμούς τους.

Σκοπός του κεφαλαίου είναι να χαρτογραφήσει το υπάρχον ερευνητικό πεδίο, να εντοπίσει τα βασικά ευρήματα, τις τάσεις, αλλά και τα υφιστάμενα ερευνητικά κενά, ώστε να τεκμηριωθεί η μεθοδολογική προσέγγιση που ακολουθείται στη συνέχεια.

2.2 Εφαρμογή του μοντέλου DEA στις σιδηροδρομικές μεταφορές

Ο McKinnon (2015) διερευνά τη σημασία και τις προκλήσεις της μέτρησης της αποτελεσματικότητας στις εμπορευματικές μεταφορές. Στο πλαίσιο της έρευνας, προσδιορίζονται έξι βασικά κριτήρια απόδοσης: η ένταση των μεταφορών, η κατανομή των μέσων μεταφοράς, η ποικιλομορφία της αγοράς, η λειτουργική αποτελεσματικότητα, η ποιότητα υπηρεσιών και ο περιβαλλοντικός αντίκτυπος. Οι εισροές περιλαμβάνουν τη χωρητικότητα των οχημάτων, την κατανάλωση καυσίμων και τις λειτουργίες logistics, ενώ οι εκροές επικεντρώνονται σε δείκτες λειτουργικής αποτελεσματικότητας όπως τα τονοχιλιόμετρα, οι εκπομπές ρύπων και η ποιότητα υπηρεσιών. Ο συγγραφέας τονίζει τη σημασία της συλλογής δεδομένων για την ανάπτυξη δημόσιων πολιτικών με στόχο τη βελτίωση των συστημάτων μεταφορών, καθώς και την ανάγκη για πολυδιάστατη μέτρηση της αποτελεσματικότητας. Τέλος, η μελέτη υποστηρίζει ότι οι λανθασμένες μετρήσεις δύνανται να προκαλέσουν στρεβλώσεις στη βιομηχανία και καταδεικνύει τις αδυναμίες των σημερινών πρακτικών μέτρησης της αποτελεσματικότητας.

Οι Cowie και Riddington (1996) εξετάζουν μεθόδους και μοντέλα μέτρησης της αποδοτικότητας των ευρωπαϊκών σιδηροδρόμων. Τα μοντέλα που διερευνούν έχουν ως μεταβλητές εισόδου δεδομένα που σχετίζονται με τη γη, την εργασία και το κεφάλαιο και ως δεδομένα εξόδου τα τρενοχιλιόμετρα των επιβατικών μεταφορών και το επίπεδο εξυπηρέτησης. Τα αποτελέσματα της ανάλυσης δείχνουν ότι οι πιο αξιόπιστη μέθοδος για τη μέτρηση της αποδοτικότητας των σιδηροδρομικών δικτύων είναι ο συνδυασμός των μοντέλων Displaced OLS (DOLS) και Unobserved component (UNCOM).

Ο Mazanec (2018) διεξάγει έρευνα για τη μέτρηση της αποδοτικότητας των εταιρειών σιδηροδρομικών εμπορευματικών μεταφορών στη Σλοβακία χρησιμοποιώντας το μοντέλο Data Envelopment Analysis (DEA). Το μοντέλο εφαρμόζεται σε 28 εταιρείες με δεδομένα εισόδου οικονομοτεχνικά στοιχεία αυτών από τη βάση δεδομένων Finstat. Από την ανάλυση των παραπάνω προκύπτουν μέσοι όροι, μέγιστες και ελάχιστες τιμές (περιγραφικά στατιστικά), τα προβλήματα καθώς και οι τομείς που χρήζουν βελτίωσης στην κάθε εταιρεία. Τα αποτελέσματα καταλήγουν στο ότι πλήθος εταιρειών σιδηροδρομικών εμπορευματικών μεταφορών δεν είναι αποτελεσματικές λόγω του μεγάλου ανταγωνισμού από εταιρείες οδικών εμπορευμάτων μεταφορών καθώς και στο ότι εταιρείες με κυρίαρχη θέση στην αγορά δεν παρουσιάζουν καλύτερη αποτελεσματικότητα από τις άλλες.

Οι Yu και Lin (2007) προτείνουν ένα νέο μοντέλο για τη μέτρηση της αποδοτικότητας των σιδηροδρόμων που επιτρέπει την αναπαράσταση των επιχειρηματικών και παραγωγικών τεχνολογιών σε ένα ενιαίο πλαίσιο. Η μέθοδος που ακολουθείται ονομάζεται Multi-activity network DEA (MNDEA) model και σε σχέση με την (ευρέως διαδεδομένη μέθοδο) DEA έχει το προτέρημα να λαμβάνει υπόψη τόσο το μη αποθηκεύσιμο χαρακτηριστικό όσο και τα χαρακτηριστικά πολλαπλών δραστηριοτήτων παραγωγής. Επιπλέον, κατατάσσει τους σιδηροδρόμους με βάση την τεχνική αποδοτικότητα των επιβατικών και εμπορευματικών σιδηροδρόμων (PTE και FTE αντίστοιχα), την Υπηρεσιακή Αποδοτικότητα (SV) και τη Συνολική Αποδοτικότητα (TV). Οι συγγραφείς εφαρμόζουν το μοντέλο με δεδομένα εισόδου το μήκος σιδηροδρομική γραμμή, τον αριθμό υπαλλήλων και βαγονιών και δεδομένα εξόδου τα τρένοχιλιόμετρα επιβατών και εμπορευμάτων. Τα αποτελέσματα επιβεβαιώνουν ότι για να βελτιώσουν την αποδοτικότητά τους, οι σιδηρόδρομοι θα πρέπει να είναι σε θέση να αναγνωρίζουν τις πηγές χαμηλής απόδοσης.

Οι Maltseva et al. (2020) εξετάζουν την αποτελεσματικότητα πέντε μεγάλων ρωσικών εταιρειών σιδηροδρομικών μεταφορών εμπορευμάτων την περίοδο 2013 έως 2017, χρησιμοποιώντας το μοντέλο Super Slack-Based Measurement (SBM) DEA. Η μελέτη αξιολογεί τόσο τη λειτουργική όσο και την οικονομική αποτελεσματικότητα. Για την πρώτη (λειτουργική αποτελεσματικότητα), λαμβάνονται ως εισροές ο αριθμός των σιδηροδρομικών βαγονιών και των εργαζομένων, ενώ ως εκροή ο εμπορευματικός κύκλος εργασιών (δισεκατομμύρια τόνοι-χιλιόμετρα). Για τη δεύτερη (οικονομική αποτελεσματικότητα), εισροές αποτελούν τα ετήσια πάγια περιουσιακά στοιχεία και τα λειτουργικά έξοδα, ενώ εκροές τα έσοδα και τα καθαρά κέρδη. Τα ευρήματα της ανάλυσης υποδηλώνουν ότι οι ιδιωτικοί φορείς επιτυγχάνουν καλύτερες επιδόσεις σε σχέση με τους κρατικούς. Οι εταιρείες που διαχειρίζονται διάφορους τύπους τροχαίου υλικού εμφανίζουν αυξημένη λειτουργική αποδοτικότητα, σε αντίθεση με εκείνες που επικεντρώνονται σε εξειδικευμένο τροχαίο υλικό και παρουσιάζουν καλύτερη οικονομική αποδοτικότητα. Τέλος, οι συγγραφείς προτείνουν την ιδιωτικοποίηση για τη βελτίωση της αποδοτικότητας και την προτεραιότητα του τροχαίου υλικού για καλύτερα λειτουργικά αποτελέσματα.

Ο Doomernik (2014) εξετάζει ένα μοντέλο για τη σύγκριση της αποδοτικότητας των σιδηροδρόμων υψηλών ταχυτήτων στην Ευρώπη και την Ασία. Το μοντέλο λαμβάνει υπόψιν αποκλειστικά το σιδηροδρομικό δίκτυο και τα περιουσιακά στοιχεία των εταιρειών. Ως μεταβλητές εισόδου χρησιμοποιεί το μήκος του δικτύου και τη χωρητικότητα του στόλου για να εξάγει την επιβατική χρήση και τις ταξιδιωτικές επιδόσεις. Ο συγγραφέας εφαρμόζει το μοντέλο σε τέσσερις ευρωπαϊκούς και τέσσερις ασιατικούς σιδηροδρόμους με δεδομένα από το 2007 μέχρι το 2012. Η μέθοδος που χρησιμοποιεί είναι η two stage Network DEA (NDEA) και κάθε σιδηρόδρομος κατατάσσεται με βάση την αποτελεσματικότητα παραγωγής, υπηρεσίας και συστήματος. Τα αποτελέσματα μαρτυρούν ότι το δεδομένο χρονικό διάστημα η Ασία

πέτυχε αύξηση της παραγωγικότητας ενώ η Ευρώπη δεν παρουσίασε κάποια βελτίωση και ότι υπάρχει αρνητική συσχέτιση μεταξύ αποτελεσματικότητας της παραγωγής και των παρεχόμενων υπηρεσιών.

Οι Zong et al. (2023) εφαρμόζουν τα μοντέλα multi-stage NDEA και Tobit για να αξιολογήσουν την αποτελεσματικότητα των σιδηροδρομικών δικτύων υψηλής ταχύτητας (HSR) σε εννέα χώρες, κατά την περίοδο 2009-2018. Αρχικά, αναλύουν τα στάδια παραγωγής και υπηρεσιών, χρησιμοποιώντας ως μεταβλητές εισόδου, για το πρώτο στάδιο, το μήκος της σιδηροδρομικής γραμμής, το εργατικό δυναμικό και τον αριθμό των συρμών. Τα δεδομένα αυτά παράγουν τη χιλιομετρική κάλυψη των συρμών που χρησιμοποιείται ως μεταβλητή εισόδου στο δεύτερο στάδιο για να εξάγει, τελικά, τον αριθμό των επιβατών και την απόσταση σε χιλιόμετρα που καλύπτεται από τους ίδιους τους επιβάτες. Η μελέτη καταλήγει στο συμπέρασμα ότι τα συστήματα HSR στην Ασία, και ιδιαίτερα την Κίνα, παρουσιάζουν υψηλότερη αποτελεσματικότητα συγκριτικά με τα αντίστοιχα ευρωπαϊκά, ιδίως σε ό,τι αφορά το στάδιο των υπηρεσιών, ενώ παρέχει συστάσεις για τη βελτίωση της αποτελεσματικότητας των παραπάνω δικτύων.

Οι Sameni et al. (2016) αξιολογούν την αποτελεσματικότητα των επιβατικών σιδηροδρομικών σταθμών στη Μεγάλη Βρετανία χρησιμοποιώντας την προσέγγιση DEA, και προτείνουν ένα μοντέλο για την τεχνική αποτελεσματικότητα και ένα για την εκείνη των υπηρεσιών. Το πρώτο μοντέλο εξετάζει το πόσο αποτελεσματικά χρησιμοποιούνται οι υποδομές, όπως ο αριθμός των αποβάθρων και το ποσοστό γραμμών διέλευσης, για την εξυπηρέτηση των στάσεων των τρένων ενώ ως δεδομένα εξόδου έχουν τον αριθμό στάσεων των τρένων. Το δεύτερο μοντέλο λαμβάνει ως εισροές τον αριθμό των στάσεων αυτών, τον πληθυσμό της περιοχής κάλυψης και τις θέσεις εργασίας και ως εκροές τις συνολικές επιβιβάσεις και αποβιβάσεις των επιβατών καθώς και τις μετακινήσεις αυτών. Η μελέτη χρησιμοποιεί επιπλέον την παλινδρόμηση Tobit για να εξετάσει την επίδραση της τοποθεσίας και του είδους της κίνησης στην αποτελεσματικότητα των σταθμών. Τα αποτελέσματα καταδεικνύουν ότι οι σταθμοί στο Λονδίνο παρουσιάζουν υψηλότερη αποτελεσματικότητα λόγω καλύτερων συνδέσεων δημόσιας συγκοινωνίας και μεγαλύτερης πυκνότητας πληθυσμού.

Η Ståhlberg (2023) αξιολογεί την αποτελεσματικότητα των Σιδηροδρομικών Διαδρόμων Εμπορευμάτων (Rail Freight Corridors - RFCs) στην Ευρώπη για την περίοδο 2020-2021, χρησιμοποιώντας τη μέθοδο DEA. Η συγγραφέας εφαρμόζει τη μέθοδο αυτή με τον όγκο της προσφερόμενης χωρητικότητας, τον αριθμό των συγκρούσεων, την προγραμματισμένη ταχύτητα και την ακρίβεια στην αναχώρηση ως δεδομένα εισόδου και τον όγκο της ζητούμενης χωρητικότητας, την ακρίβεια στην άφιξη και τον αριθμό των τρένων που διασχίζουν τα σύνορα ως δεδομένα εξόδου. Σύμφωνα με την ίδια, τέσσερις από τους οκτώ διαδρόμους ήταν αποτελεσματικοί, ενώ οι υπόλοιποι παρουσίασαν περιθώρια βελτίωσης. Η μείωση του χρόνου παραμονής στα σύνορα αναγνωρίστηκε ως βασικός παράγοντας βελτίωσης της αποτελεσματικότητας και της ανταγωνιστικότητας.

Οι Zhang et al. (2022) αξιολογούν τη λειτουργική και ενεργειακή αποτελεσματικότητα των γραμμών σε αστικά σιδηροδρομικά δίκτυα σε τέσσερις μεγάλες πόλεις της Κίνας (Πεκίνο, Σανγκάη, Κουαγκτσόου και Σενζέν), χρησιμοποιώντας το μοντέλο Slacks-Based Measure (SBM) DEA. Ως δεδομένα εισόδου επιλέγονται, εδώ, το μήκος γραμμής, ο αριθμός σταθμών και τρένων, καθώς και η κατανάλωση ενέργειας. Ως δεδομένα εξόδου χρησιμοποιούνται ο όγκος μεταφοράς επιβατών, τα έσοδα ανά επιβατοχιλιόμετρο και οι εκπομπές CO₂. Τα αποτελέσματα δείχνουν σημαντικές αποκλίσεις μεταξύ των πόλεων, με την Κουαγκτσόου να επιτυγχάνει την υψηλότερη λειτουργική αποτελεσματικότητα και τη Σαγκάη την υψηλότερη ενεργειακή. Τέλος, οι

σιδηροδρομικές γραμμές που ανήκουν στο κράτος παρουσιάζουν εν γένει υψηλότερη λειτουργική, ενώ οι κοινοπραξίες καταγράφουν καλύτερη ενεργειακή αποτελεσματικότητα.

Οι Ferreira et al. (2020) εξετάζουν τις επιπτώσεις της συγχώνευσης μεταξύ δύο μεγάλων σιδηροδρομικών εταιρειών της Βραζιλίας, της Rumo και της ALL, ως προς τη λειτουργική αποτελεσματικότητα του δικτύου. Μέσω της χρήσης της DEA, η μελέτη αναλύει τις μεταβολές που επήλθαν ως προς την αποτελεσματικότητα μετά τη συγχώνευση, εστιάζοντας στις οικονομίες κλίμακας, στη βελτιστοποίηση του δικτύου και στη διαχείριση των πόρων. Τα ευρήματα δείχνουν ότι η συγχωνευμένη εταιρεία επέτυχε αύξηση της λειτουργικής αποτελεσματικότητας, αλλά με φθίνουσα απόδοση λόγω της υπερβολικής αύξησης των εισροών χωρίς αντίστοιχη αύξηση των εκροών. Η μελέτη προσφέρει χρήσιμες πληροφορίες σχετικά με το πώς οι κάθετες συγχωνεύσεις μπορούν να ενισχύσουν την αποτελεσματικότητα σε μονοπωλιακά περιβάλλοντα μεταφοράς, ειδικά όταν ο ανταγωνισμός είναι περιορισμένος.

Οι Blagojević et al. (2020) αξιολογούν την αποτελεσματικότητα των σιδηροδρομικών επιχειρήσεων εμπορευματικών μεταφορών στα δυτικά Βαλκάνια χρησιμοποιώντας συνδυαστικά την μεθοδολογία της Fuzzy Analytic Hierarchy Process (FAHP) και την DEA. Αυτή η μεθοδολογία ενσωματώνει κριτήρια που σχετίζονται με τους πόρους, τις λειτουργίες, τα χρηματοοικονομικά, την ποιότητα υπηρεσιών και την ασφάλεια. Η μέθοδος FAHP εντοπίζει τα κριτήρια με τη μεγαλύτερη επίδραση σε κάθε κατηγορία, τα οποία στη συνέχεια χρησιμοποιούνται στο μοντέλο DEA για τη μέτρηση της αποτελεσματικότητας. Η ανάλυση καταλήγει σε σύγκριση των επιδόσεων των εταιρειών, αναδεικνύοντας τη Σλοβενία και τη Σερβία ως πρότυπα αποτελεσματικότητας. Η έρευνα ολοκληρώνεται με συστάσεις για την υιοθέτηση πρακτικών που ενισχύουν την αποτελεσματικότητα και ενθαρρύνει τη χρήση μοντέλων προσομοίωσης και βελτιστοποίησης για μελλοντικές έρευνες.

2.3 Διεθνή πλαίσια για εκτίμηση της αποδοτικότητας

Οι Gathon και Perelman (1992) παρουσιάζουν μία ολοκληρωμένη ανάλυση τεχνικής αποτελεσματικότητας στους ευρωπαϊκούς σιδηροδρόμους, χρησιμοποιώντας μία προσέγγιση δεδομένων πάνελ (panel data approach). Οι συγγραφείς εκτιμούν ένα όριο απαιτούμενων συντελεστών (factor requirement frontier), εστιάζοντας στην επίδραση της διοικητικής αυτονομίας των εταιρειών και των ρυθμιστικών πλαισίων στην τεχνική αποτελεσματικότητα. Διαπιστώνουν ότι η τεχνική αποτελεσματικότητα καθορίζεται ενδογενώς και προτείνουν εναλλακτικά μέτρα που δεν επηρεάζονται από εξωτερικές μεταβλητές. Η προσέγγισή τους βασίζεται στις μεθόδους των Schmidt και Sickles (1984) και προσαρμόζουν την τεχνική των Hausman και Taylor (1981) για την αντιμετώπιση ζητημάτων συσχέτισης μεταξύ αποτελεσματικότητας και άλλων μεταβλητών. Τα δεδομένα εισόδου περιλαμβάνουν το εργατικό δυναμικό, τα τρενοχιλιόμετρα, το μέγεθος του δικτύου και την ηλεκτροδότηση, ενώ δεδομένο εξόδου αποτελεί η συνολική απόσταση που διανύθηκε από επιβάτες και από εμπορεύματα. Η μελέτη καταλήγει στο συμπέρασμα ότι η θεσμική διοικητική αυτονομία έχει θετική συσχέτιση με την τεχνική αποτελεσματικότητα, υποδεικνύοντας ότι οι επιχειρήσεις με μεγαλύτερη διοικητική ελευθερία αποδίδουν καλύτερα.

Το Διεθνές Φόρουμ Μεταφορών (ITF) (2019) εξετάζει πώς μπορεί να οριστεί και να μετρηθεί η αποτελεσματικότητα των σιδηροδρομικών λειτουργιών καθώς και η διαχείριση των υποδομών των σιδηροδρομικών δικτύων. Τονίζονται οι προκλήσεις, όπως η πολυπλοκότητα των σιδηροδρομικών συστημάτων, η διαθεσιμότητα δεδομένων και η επίδραση εξωγενών παραγόντων. Η μεθοδολογία που ακολουθείται περιλαμβάνει τη χρήση βασικών δεικτών απόδοσης (KPIs), οικονομετρικών μοντέλων όπως η Stochastic Frontier Analysis (SFA) και η Data Envelopment Analysis (DEA).

Δεδομένα εισόδου στη μελέτη αποτελούν τα λειτουργικά κόστη, οι υποδομές, το ανθρώπινο δυναμικό και οι επιδόσεις των σιδηροδρομικών υπηρεσιών. Τα δεδομένα εξόδου αφορούν την αποτελεσματικότητα στη χρήση πόρων και τις οικονομίες κλίμακας με συγκρίσεις μεταξύ σιδηροδρομικών συστημάτων διαφόρων χωρών. Τα αποτελέσματα υπογραμμίζουν τη σημασία των δεδομένων υψηλής ποιότητας, προτείνουν τα οικονομετρικά μοντέλα ως την πλέον κατάλληλη προσέγγιση για την αξιολόγηση της αποτελεσματικότητας και επισημαίνουν την ανάγκη για μακροπρόθεσμη συνεργασία μεταξύ νομοθετών, ρυθμιστικών αρχών και σιδηροδρομικών εταιρειών, καθώς και τη χρήση προηγμένων μεθοδολογιών για την αντιμετώπιση των προκλήσεων.

Ο Holvad (2020) παρέχει μία εκτενή επισκόπηση των βασικών ζητημάτων στην ανάλυση της αποτελεσματικότητας των σιδηροδρόμων, εξετάζοντας μεθόδους frontier-based, οι οποίες περιλαμβάνουν τόσο παραμετρικές προσεγγίσεις, όπως η Stochastic Frontier Analysis (SFA), όσο και μη παραμετρικές, όπως η DEA. Αυτές αξιολογούν τη μετατροπή των εισροών, όπως το ανθρώπινο δυναμικό, το τροχαίο υλικό, το μήκος των σιδηροτροχιών και την κατανάλωση ενέργειας, σε εκροές, όπως είναι τα επιβατοχιλιόμετρα και τα τονοχιλιόμετρα εμπορευμάτων. Ο συγγραφέας εξετάζει πώς αυτές οι μέθοδοι έχουν εφαρμοστεί στην Ευρώπη και τον διεθνή σιδηροδρομικό τομέα, αναδεικνύοντας τη δυνατότητά τους να εντοπίζουν βέλτιστες πρακτικές και να ποσοτικοποιούν πιθανά οφέλη αποτελεσματικότητας. Παράλληλα, υπογραμμίζει προκλήσεις, όπως η ευαισθησία στην ποιότητα των δεδομένων και η ανάγκη για πιο ολοκληρωμένα σύνολα δεδομένων, τα οποία είναι κρίσιμα για αξιόπιστες αξιολογήσεις αποτελεσματικότητας.

Οι Thompson και Bente (2014) εξετάζουν την πολυπλοκότητα της μέτρησης της αποτελεσματικότητας στους σιδηροδρόμους η οποία απορρέει από την ποικιλία των εκροών, το μεγάλο μέγεθος και εύρος των σιδηροδρομικών δραστηριοτήτων και τους περιορισμούς στα δεδομένα σε παγκόσμιο επίπεδο. Η μελέτη χρησιμοποιεί μια σειρά δεικτών αποτελεσματικότητας με βάση τα διαθέσιμα δεδομένα, συγκρίνοντας την απόδοση διατομεακά και διαχρονικά, λαμβάνοντας υπόψη παράγοντες όπως την ποικιλία των υπηρεσιών, την κλίμακα και τις κοινωνικοοικονομικές επιπτώσεις. Συζητά τις προκλήσεις στην απόκτηση συνεκτικών δεδομένων και την επίδραση των μεταρρυθμίσεων από τις κυβερνήσεις και τους ρυθμιστικούς φορείς, των αλλαγών ιδιοκτησίας και των κινήτρων διαχείρισης στην αποτελεσματικότητα, χρησιμοποιώντας παραδείγματα από σιδηροδρόμους στις Η.Π.Α., τον Καναδά, την Ιαπωνία και την Ε.Ε. Οι συγγραφείς προτείνουν ότι οι βελτιώσεις στην αποτελεσματικότητα συχνά απαιτούν εξατομικευμένες δομικές μεταρρυθμίσεις, διαφάνεια στη δημόσια αναφορά και μια προσέγγιση ισορροπημένης βαθμολογίας επιδόσεων για ολοκληρωμένη αξιολόγηση της απόδοσης.

Οι Goelli et al. (2003) προτείνουν έναν ολοκληρωμένο οδηγό για τους ρυθμιστικούς φορείς στους τομείς των υπηρεσιών κοινής ωφέλειας και των μεταφορών, σχετικά με τη μέτρηση της αποτελεσματικότητας. Τονίζουν τη σημασία της προώθησης και διασφάλισης της αποτελεσματικότητας μέσω πολιτικών που προσφέρουν κίνητρα στις εταιρείες. Οι συγγραφείς εξετάζουν μεθοδολογίες όπως την DEA, την SFA και τους Price-Based Index Numbers (PIN) για τη μέτρηση της Total Factor Productivity (TFP) και της αποτελεσματικότητας των επιχειρήσεων. Κάθε μέθοδος αξιολογείται με βάση τις απαιτήσεις δεδομένων, την εφαρμοσιμότητα της και την ακρίβεια της. Ο οδηγός συζητά τις προκλήσεις που αντιμετωπίζουν οι ρυθμιστικές αρχές λόγω ασυνεπειών στα δεδομένα και της ασυμμετρίας των πληροφοριών. Καταλήγοντας, παρέχει συστάσεις που αφορούν την εφαρμογή ανώτατων ορίων τιμών περιορίζοντας το κόστος για τους καταναλωτές και ενθαρρύνοντας τις εταιρείες να μειώσουν τα λειτουργικά τους κόστη, την ανάλυση συγκριτικής αξιολόγησης και την αξιολόγηση της

ανάπτυξης της παραγωγικότητας. Με τον τρόπο αυτό οι επιχειρήσεις δύνανται να κατανοήσουν εάν πράγματι βελτιώνουν την αποδοτικότητά τους με την πάροδο του χρόνου.

Οι Pazour et al. (2010) προτείνουν τον σχεδιασμό ενός δικτύου σιδηροδρόμων υψηλής ταχύτητας για τη μεταφορά εμπορευμάτων στις Ηνωμένες Πολιτείες. Οι συγγραφείς αναπτύσσουν ένα βελτιστοποιημένο μοντέλο για τον σχεδιασμό ενός δικτύου που ελαχιστοποιεί τους χρόνους μεταφοράς των εμπορευμάτων, μειώνοντας έτσι τη συμφόρηση στους αυτοκινητόδρομους και τις περιβαλλοντικές επιπτώσεις που συνδέονται με τη μεταφορά φορτηγών. Το μοντέλο χρησιμοποιεί μία διαδικασία πολλών σταδίων, ξεκινώντας με τον σχεδιασμό ενός δικτύου χωρίς περιορισμούς χωρητικότητας για να εντοπίσει τους ιδανικούς συνδέσμους σιδηροδρόμων υψηλής ταχύτητας μεταξύ βασικών πόλεων. Στη συνέχεια, ενσωματώνει τους περιορισμούς χωρητικότητας σε ένα δεύτερο στάδιο για να διασφαλίσει τη βιωσιμότητα του σχεδιασμού. Η μελέτη περιλαμβάνει ένα σχέδιο εφαρμογής για σταδιακή ανάπτυξη και παρουσιάζει πιθανά οφέλη, όπως την μείωση των χιλιομέτρων που διανύουν τα φορτηγά και την ελάττωση στους χρόνους μεταφοράς ωστόσο αντιμετωπίζει προκλήσεις όπως περιορισμούς στην χρηματοδότηση, αβεβαιότητες στη ζήτηση και τη διασφάλιση επαρκούς χωρητικότητας. Η έρευνα καταλήγει στο ότι ένα δίκτυο σιδηροδρόμων υψηλής ταχύτητας για τη μεταφορά εμπορευμάτων θα μπορούσε να βελτιώσει σημαντικά την αποτελεσματικότητα των logistics και να μειώσει τη συμφόρηση στους αυτοκινητόδρομους.

Η Karakeke (2020) εστιάζει στην ανάπτυξη ενός εργαλείου με βάση την ανάλυση συνολικού κόστους για την επιλογή μεταξύ σιδηροδρομικών και οδικών εμπορευματικών μεταφορών στην Ευρώπη, διαλέγοντας ως περιοχή μελέτης τον διάδρομο Orient-East Mediterranean (OEM). Χρησιμοποιεί το MATLAB για την ανάλυση δεδομένων και την εφαρμογή συναρτήσεων κόστους, το CargoApps για τον υπολογισμό του κόστους του οδικού δικτύου με βάση τα χαρακτηριστικά των οχημάτων, το Mautex για την ακριβή πρόβλεψη διοδίων στις διαδρομές, το Quotex για τον υπολογισμό συνολικού κόστους, όπως καύσιμα, διόδια, μισθοί οδηγών και συντήρηση οχημάτων, το VECTO για τη μοντελοποίηση εκπομπών CO₂ και κατανάλωσης καυσίμου σε βαρέα οχήματα και το QGIS για τον γεωγραφικό υπολογισμό και την απεικόνιση αποστάσεων στον OEM. Η μέθοδος βασίζεται σε μια πολυκριτηριακή ανάλυση που περιλαμβάνει εσωτερικά, εξωτερικά και χρονικά κόστη, ώστε να παρέχει τη βέλτιστη επιλογή μεταφορικού μέσου όσον αφορά την οικονομική αποτελεσματικότητα και τη βιωσιμότητα του. Μέσω περιπτώσιολογικών μελετών, το εργαλείο βοηθά τους χρήστες να συγκρίνουν τις δύο επιλογές κατά μήκος του διαδρόμου OEM και να επιλέγεται ανάλογα με τις αποστάσεις και τις ποσότητες που μεταφέρονται, η βέλτιστη λύση, προωθώντας έτσι πιο βιώσιμες εμπορευματικές μεταφορές στην Ευρώπη.

2.4 Συνδυαστική εφαρμογή μοντέλων DEA και AI

Ο Kwon (2017) διερευνά τη δυνατότητα συνδυαστικής χρήσης της DEA και των τεχνητών νευρωνικών δικτύων (ANN) για την αποδοτική μοντελοποίηση και πρόβλεψη της απόδοσης σιδηροδρομικών εταιρειών στις ΗΠΑ (Class I Railroads). Αξιοποιείται ένα συνδυαστικό μοντέλο δύο σταδίων: πρώτα εφαρμόζεται το output-oriented μοντέλο DEA τύπου CCR για την αξιολόγηση της τεχνικής αποδοτικότητας κάθε DMU, ενώ στη συνέχεια εκπαιδεύεται ένα νευρωνικό δίκτυο (BPNN) ώστε να προβλέπει τόσο τις βαθμολογίες αποδοτικότητας όσο και τις αναμενόμενες εκροές για κάθε μονάδα. Η ανάλυση βασίστηκε σε δεδομένα χρονοσειρών 8 ετών (2005–2012) για επτά εταιρείες (BNSF, CSX, GTC, KCS, NS, SOO, UP), με συνολικά 56 DMUs. Οι εισροές είναι: ο αριθμός των μηχανών και βαγονιών, το μήκος γραμμής και το προσωπικό. Οι εκροές

είναι οι RTM (Revenue Ton-Miles) και FTON (Freight Tons). Το άρθρο παρουσιάζει δύο νευρωνικά δίκτυα (δομές 5-7-1 και 5-8-1) για κάθε μία από τις δύο εκδοχές του μοντέλου (RTM & FTON), ενώ επιβεβαιώνεται η υψηλή ακρίβεια πρόβλεψης με $R^2 > 0.97$ σε όλα τα σετ δοκιμής. Κεντρική συνεισφορά της μελέτης είναι η εφαρμογή του ANN όχι μόνο για πρόβλεψη του DEA score, αλλά και για την αναπαραγωγή της επιφάνειας αποδοτικότητας (efficient frontier) μέσω εκπαίδευσης με τις βέλτιστες προβολές DEA (projected inputs-outputs). Το μοντέλο μπορεί να προβλέψει εναλλακτικά σενάρια για "τιμές εκροών βέλτιστης απόδοσης" υπό διαφορετικά επίπεδα εισροών, προσφέροντας ένα ισχυρό εργαλείο υποστήριξης αποφάσεων και benchmarking. Η μελέτη αποδεικνύει ότι ο συνδυασμός DEA-BPNN μπορεί να υποστηρίξει managers σε "what-if" σενάρια, στη βαθμιαία βελτίωση απόδοσης και στη διαχείριση περιορισμών πόρων. Η εργασία αυτή θεωρείται θεμελιώδης για το συνδυαστικό μοντέλο DEA-ANN στον σιδηροδρομικό τομέα, ενώ προτείνει ως μελλοντική έρευνα την επέκταση του μοντέλου σε multi-stage διαδικασίες και την ενσωμάτωση του κόστους ως επιπλέον μεταβλητή.

Οι Amirteimoori et al. (2023) προτείνουν έναν συνδυασμό DEA και τεχνητής νοημοσύνης για την αξιολόγηση και πρόβλεψη της περιβαλλοντικής αποδοτικότητας, με έμφαση στη μείωση των εκπομπών CO₂ ως ανεπιθύμητης εκροής. Αναπτύσσεται το μοντέλο FUNDEA (Fuzzy Undesirable Non-Discretionary SBM-DEA), το οποίο ενσωματώνει ασαφείς (τριγωνικές) επιθυμητές και ανεπιθύμητες εκροές και μη ελέγξιμες εισροές, σε συνθήκες αβεβαιότητας. Εφαρμόζεται σε 24 δασικές μονάδες του Ιράν με εισροές: το κόστος παραγωγής, οδοποιίας και μεταφοράς και το εμβαδόν δασικής περιοχής (ND). Οι εκροές είναι: τα έσοδα, ο όγκος αποθεμάτων (επιθυμητές) και οι εκπομπές CO₂ (ανεπιθύμητη). Αρχικά, το μοντέλο εντοπίζει 17 μη αποδοτικές μονάδες και ακολουθεί βελτιωμένη διαδικασία στόχων μέσω FUNDEA και χρήση τριών μοντέλων πρόβλεψης: ANN (MLP δίκτυο), GEP (γενετικός αλγόριθμος) και AIS (τεχνητό ανοσοποιητικό σύστημα). Οι τρεις προσεγγίσεις συγκρίνονται ως προς την ακρίβεια πρόβλεψης των βέλτιστων τιμών CO₂, με το FUNDEA-AIS να υπερέχει ($R^2 = 0.9999$, $MSE = 0.00034$), ενώ το ANN παρουσιάζει τη χαμηλότερη επίδοση. Η μελέτη καταλήγει στο ότι μόνο το FUNDEA-AIS κατόρθωσε να καταστήσει όλες τις μονάδες πλήρως αποδοτικές και αναδεικνύει τη σημασία των υβριδικών μοντέλων DEA-AI σε προβλήματα με ασαφή και μη ποσοτικοποιήσιμα δεδομένα.

Ο Çakır (2024) αναπτύσσει ένα υβριδικό πλαίσιο για την πρόβλεψη της αποδοτικότητας και την επίτευξη των βέλτιστων εκροών για 37 σιδηροδρομικές εταιρείες του ΟΟΣΑ (Οργανισμός Οικονομικής Συνεργασίας και Ανάπτυξης), συνδυάζοντας το context-dependent DEA με μηχανική μάθηση (ANFIS και SVM). Εφαρμόζεται, εδώ, output-oriented stratified DEA για τον υπολογισμό σκορ αποδοτικότητας και βέλτιστων επιπέδων εκροών, με πιο ρεαλιστικούς στόχους από τα κλασικά μοντέλα. Ως εισροές χρησιμοποιούνται: ο αριθμός εργαζομένων, το μήκος γραμμής και το τροχαίο υλικό· και ως μοναδική εκροή τα λειτουργικά έσοδα (σε όρους PPP). Στο επόμενο στάδιο, τα ANN μοντέλα (ANFIS και SVR) εκπαιδεύτηκαν με τις προβολές DEA, για να προβλέψουν τόσο τα σκορ αποδοτικότητας όσο και τις εκροές για μελλοντικές ή μη παρατηρούμενες μονάδες (DMUs). Το ANFIS εμφάνισε καλύτερη ακρίβεια ($R^2 = 0.963$, $MAPE = 2.97\%$) από το SVR, τόσο σε πρόβλεψη σκορ DEA όσο και σε στόχους εκροών. Επίσης, επιβεβαιώθηκε ότι η χρήση προεπεξεργασμένων (monotonic) δεδομένων από τη DEA ενίσχυσε σημαντικά την ακρίβεια των προβλέψεων. Η μελέτη συμβάλλει στην πρόβλεψη επιδόσεων και στόχων με σεβασμό στη μονοτονία, υποστηρίζοντας σταδιακή βελτίωση και benchmarking. Το προτεινόμενο μοντέλο DEA-ANN μπορεί να εφαρμοστεί ευρύτερα σε κάθε κλάδο για πρόβλεψη αποδοτικότητας και λήψη στρατηγικών αποφάσεων.

Οι Azadeh et al. (2008) προτείνουν ένα υβριδικό μοντέλο βελτιστοποίησης σιδηροδρομικών δικτύων μέσω ενσωμάτωσης DEA, analytical hierarchy process (AHP) και προσομοίωσης, το οποίο εφαρμόζεται σε σιδηροδρομική γραμμή 800 χλμ. με 50 σταθμούς. Στόχος είναι η αξιολόγηση 22 σεναρίων δρομολόγησης επιβατικών και εμπορευματικών τρένων, με έμφαση στην αξιοπιστία, τη μείωση καθυστερήσεων και τη λειτουργική βελτίωση. Η προσομοίωση δημιουργήθηκε σε Visual SLAM και παρείχε ποσοτικά δεδομένα, η AHP μετασχημάτισε ποιοτικές εισροές (ανθρώπινα λάθη, κόστος), και η DEA (input-oriented, VRS) αξιολόγησε την αποδοτικότητα κάθε σεναρίου. Οι εισροές περιλαμβάνουν, εδώ, τους χρόνους μετακίνησης, τις καθυστερήσεις, τα σφάλματα και τα κόστη, ενώ οι εκροές την αξιοπιστία δρομολογίων και τα λειτουργικά οφέλη. Τα σενάρια 7 και 11 παρουσίασαν τη μεγαλύτερη αποδοτικότητα (DEA = 1.22 και 1.16), ενώ το υφιστάμενο σύστημα κρίθηκε μη αποδοτικό (DEA = 0.89). Οι στόχοι εισροών για τα μη αποδοτικά σενάρια υπολογίστηκαν μέσω slack-based DEA. Η μελέτη επιβεβαιώνει τη χρησιμότητα της ενσωμάτωσης ποιοτικών και ποσοτικών δεδομένων για την αποδοτική λήψη αποφάσεων σε πολύπλοκα σιδηροδρομικά συστήματα.

Οι Costa και Markellos (1997) συγκρίνουν τρεις μεθόδους μέτρησης αποδοτικότητας - DEA, COLS και MLP - εφαρμόζοντάς τις στο μετρό του Λονδίνου για την περίοδο 1970–1994, με στόχο την αποτίμηση της τεχνικής αποδοτικότητας και την διερεύνηση εναλλακτικών μεθόδων μέτρησης της παραγωγικότητας στις δημόσιες συγκοινωνίες. Εισροές αποτελούν ο αριθμός τρένων και εργαζομένων και μοναδική εκροή τα συνολικά χιλιόμετρα λειτουργίας ανά έτος. Τα MLP (Multilayer Perceptron) εφαρμόζονται σε δύο μορφές: στοχαστική (ECN-MLP) και περιβάλλουσα (ENV-MLP), εκπαιδευμένα σε συνθετικά δεδομένα με οπισθοδρόμηση (backpropagation) και επιλογή βέλτιστου μοντέλου βάσει SIC. Η DEA, η οποία εφαρμόστηκε τόσο με σταθερές (CRS) όσο και με μεταβλητές αποδόσεις κλίμακας (VRS) και τα MLP, παρήγαγαν συγκρίσιμα αποτελέσματα, αφού και οι δύο προσεγγίσεις έδειξαν αύξηση της αποδοτικότητας μετά το 1984. Το μοντέλο ENV-MLP έδωσε εκτιμήσεις κοντά σε αυτές της DEA, ενώ το ECN-MLP παρήγαγε πιο εξομαλυμένα και στοχαστικά αποτελέσματα. Επιπλέον, το MLP μπόρεσε να αναγνωρίσει σχέσεις ανάμεσα στις εισροές και τις εκροές που δεν ήταν απαραίτητα γραμμικές. Σε ορισμένες περιπτώσεις, εντοπίστηκαν σημεία συμφόρησης ή εξαρτήσεις μεταξύ μεταβλητών, όπως για παράδειγμα στο ότι η αποδοτικότητα του στόλου φαινόταν να σχετίζεται με τον αριθμό προσωπικού. Στην ανάλυση περιλήφθηκαν επίσης υπολογισμοί στόχων εισροών και κάποια σενάρια μελλοντικής λειτουργίας, μαζί με εκτιμήσεις για τα όρια της αποδοτικότητας. Συνολικά, το MLP αποδείχθηκε ένα αρκετά ευέλικτο εργαλείο, που δύναται να βοηθήσει τόσο σε προβλέψεις όσο και στον σχεδιασμό παρεμβάσεων, σε αντίθεση με πιο αυστηρές μεθόδους όπως η DEA ή η COLS.

Οι Pendharkar και Rodger (2003) προτείνουν μια μέθοδο επιλογής δεδομένων εκπαίδευσης για τεχνητά νευρωνικά δίκτυα (ANN), βασισμένη στην τεχνική αποδοτικότητα μέσω της DEA, με στόχο την ενίσχυση της ακρίβειας πρόβλεψης υπό την υπόθεση μονοτονίας. Εφαρμόζεται σε 175 νοσοκομεία της Πενσυλβάνια για πρόβλεψη του αριθμού εργαζομένων βάσει των: εισαγωγών, κλινών, επισκέψεων επειγόντων, ιδιοκτησιακού καθεστώτος και κύριας νοσολογικής κατηγορίας (DRG). Ως εισροές για τη DEA λαμβάνονται οι τρεις πρώτες, και ως εκροή ο αριθμός εργαζομένων. Τα δεδομένα χωρίζονται με βάση τα DEA efficiency scores σε “αποδοτικά” και “μη αποδοτικά” σύνολα, ενώ το ANN εκπαιδεύεται με χρήση backpropagation σε δύο σενάρια: (α) με τα 50 πιο αποδοτικά νοσοκομεία (efficiency $\geq$ 53%) και (β) με τα 50 λιγότερο αποδοτικά. Ακολούθως, γίνεται πρόβλεψη σε σύνολο 75 νοσοκομείων. Το μοντέλο που εκπαιδεύτηκε με αποδοτικές μονάδες εμφάνισε

σημαντικά υψηλότερη ακρίβεια ($RMS = 0.08$, accuracy 85.3%) έναντι του μοντέλου με μη αποδοτικές ($RMS = 0.11$, accuracy 54.7%). Η μελέτη καταλήγει στο ότι η προεπιλογή παραδειγμάτων με βάση την αποδοτικότητα βελτιώνει την ακρίβεια και τη μονοτονία στην πρόβλεψη ANN, ενώ μειώνει τον κίνδυνο overfitting. Η προσέγγιση DEA-ANN συνιστάται εδώ όταν η προβλεπτική σχέση οφείλει να διατηρεί μονοτονικό χαρακτήρα, όπως π.χ. στη διαχείριση ανθρωπίνου δυναμικού ή σε άλλες εφαρμογές με αυστηρές επιχειρησιακές σχέσεις.

Η Wu (2009) προτείνει ένα υβριδικό μοντέλο αξιολόγησης και επιλογής προμηθευτών, συνδυάζοντας DEA, δέντρα αποφάσεων (DT) και τεχνητά νευρωνικά δίκτυα (ANN). Το μοντέλο λειτουργεί σε δύο στάδια: πρώτα εφαρμόζεται DEA για την κατάταξη των προμηθευτών σε αποδοτικούς και μη, και στη συνέχεια εκπαιδεύονται τα DT και ANN μοντέλα για την πρόβλεψη ή ταξινόμηση μελλοντικών υποψηφίων. Στόχος είναι η επιλογή προμηθευτών με βάση τόσο ποσοτικά όσο και ποιοτικά κριτήρια, ακόμη και σε συνθήκες μικρών ή ελλιπών datasets. Η εφαρμογή έγινε σε δείγμα 23 προμηθευτών πολυεθνικής τηλεπικοινωνιακής εταιρείας. Οι εισροές περιλαμβάνουν: την ποιότητα συστημάτων, την αυτο-αξιολόγηση, την παραγωγική δυνατότητα, τη διοικητική ικανότητα, την ικανότητα σχεδιασμού και την ικανότητα μείωσης κόστους. Απαναντίας, οι εκροές είναι: η ποιότητα, η τιμή, η παράδοση, η απόδοση στη μείωση κόστους και άλλα. Επιπλέον, ενσωματώθηκαν ποιοτικές μεταβλητές (οικονομικό περιβάλλον και τοποθεσία) και ελλιπείς τιμές, με χειρισμό μέσω κατηγοριοποίησης και αντικατάστασης μέσων όρων. Τα μοντέλα ANN (με backpropagation) και DT εκπαιδεύτηκαν με δεδομένα DEA (CCR και BCC) και απέδωσαν υψηλά ποσοστά πρόβλεψης, με τα ANN να υπερτερούν στην ακρίβεια και σταθερότητα. Το τελικό CCR-ANN μοντέλο πρόέβλεψε με ακρίβεια (>90%) τους καλύτερους προμηθευτές, ταυτίζοντας πλήρως τις επιλογές του με το προηγμένο μοντέλο TN (Talluri & Narasimhan, 2005). Επιβεβαιώθηκε ότι η χρήση DEA ως φίλτρο βελτιώνει την εκμάθηση και τη διατήρηση της μονοτονικότητας στις προβλέψεις ANN. Η μελέτη επιβεβαιώνει τη συμπληρωματικότητα DEA-AI σε σύνθετες αποφάσεις προμηθευτών και προτείνει μελλοντική έρευνα για επαλήθευση σε μεγαλύτερα σύνολα και σύγκριση με άλλες τεχνικές εξόρυξης γνώσης (DM).

2.5 Συμπεράσματα βιβλιογραφίας

Η βιβλιογραφική ανασκόπηση ανέδειξε την εξέλιξη και την πολυπλοκότητα της εκτίμησης της αποδοτικότητας στις σιδηροδρομικές μεταφορές, με κύρια μεθοδολογική βάση την Περιβάλλουσα Ανάλυση Δεδομένων (Data Envelopment Analysis). Η DEA εφαρμόστηκε τόσο σε τεχνικό όσο και οικονομικό επίπεδο, με σταθερές ή μεταβλητές αποδόσεις κλίμακας, ενώ σημαντική έμφαση δόθηκε σε μοντέλα πολλαπλών σταδίων (multi-stage DEA) και σε επεκτάσεις όπως το Super-SBM, MNDEA και το context-dependent stratified DEA, που προσφέρουν μεγαλύτερη ευελιξία και ρεαλισμό στη μοντελοποίηση.

Στην πλειοψηφία τους οι εφαρμογές και εργασίες επικεντρώνονται σε σιδηροδρομικές εταιρείες, σταθμούς, δίκτυα υψηλών ταχυτήτων και διαδρόμους εμπορευματικών μεταφορών, λαμβάνοντας ως εισροές μεταβλητές όπως το ανθρώπινο δυναμικό, το τροχαίο υλικό, τις επενδύσεις και την ενέργεια, και εκροές τα τονοχιλιόμετρα, τα έσοδα και τις περιβαλλοντικές παραμέτρους. Παράλληλα, η εισαγωγή τεχνητής νοημοσύνης (ANN, SVM, ANFIS, AIS) και αλγορίθμων πρόβλεψης έχει ενισχύσει σημαντικά την ικανότητα των μοντέλων να παράγουν ρεαλιστικούς στόχους και να προβλέπουν αποδοτικούς συνδυασμούς εισροών και εκροών, ξεπερνώντας τους περιορισμούς των παραδοσιακών μεθόδων.

Συγκεκριμένα, τα υβριδικά μοντέλα DEA–ANN χρησιμοποιούνται για την πρόβλεψη επιδόσεων, την εκμάθηση της έννοιας της αποδοτικότητας και την εξαγωγή στόχων, αξιοποιώντας δεδομένα που έχουν πρώτα υποστεί επεξεργασία μέσω DEA. Τέτοιες προσεγγίσεις εντοπίζονται σε εφαρμογές μεταφορών, ενέργειας, τραπεζών και εφοδιασμού. Η χρήση μοντέλων όπως τα MLP, GEP, decision trees, ακόμα και τεχνητών ανοσοποιητικών συστημάτων (AIS), επιβεβαιώνει ότι οι συνδυαστικές τεχνικές βελτιώνουν τη σταθερότητα και την ακρίβεια της πρόβλεψης, ιδίως σε προβλήματα με πολυδιάστατα, ποιοτικά ή ασαφή δεδομένα. Παράλληλα, επισημαίνεται η σημασία της μονοτονικότητας στην πρόβλεψη, καθώς και η ανάγκη προεπιλογής αποδοτικών παραδειγμάτων κατά την εκπαίδευση των ANN. Αρκετές μελέτες αξιοποιούν την DEA όχι μόνο για την αξιολόγηση, αλλά και ως φίλτρο/εργαλείο ενίσχυσης της ακρίβειας και γενίκευσης των ANN. Επιπλέον, η σύγκριση της DEA με παραδοσιακές μεθόδους (π.χ. COLS, ARIMA) αναδεικνύει τη δυναμική των ANN στη διαχείριση μη γραμμικών σχέσεων και την παραγωγή ρεαλιστικών σεναρίων για χάραξη πολιτικής.

Συνολικά, η βιβλιογραφία τεκμηριώνει τη δυνατότητα συνδυασμού DEA–AI για την επίτευξη πρόβλεψης, benchmarking, ιεράρχησης και στρατηγικής βελτιστοποίησης σε διαφορετικούς κλάδους ανάμεσα τους και αυτός των μεταφορών. Ωστόσο, εντοπίζονται ακόμη ερευνητικά κενά, όπως η ανάγκη για μεγαλύτερα datasets, η έλλειψη συγκριτικής αξιολόγησης διαφορετικών τεχνικών AI, και η περιορισμένη εφαρμογή σε περιβάλλοντα με περιορισμένους ή μη ελέγξιμους πόρους. Η παρούσα εργασία φιλοδοξεί να καλύψει ένα μέρος αυτού του κενού, εφαρμόζοντας συνδυαστική προσέγγιση DEA–ANN στον σιδηροδρομικό κλάδο, με σκοπό την πρόβλεψη αποδοτικών εισροών σε πραγματικά δεδομένα οργανισμών που δραστηριοποιούνται στην Ευρώπη.

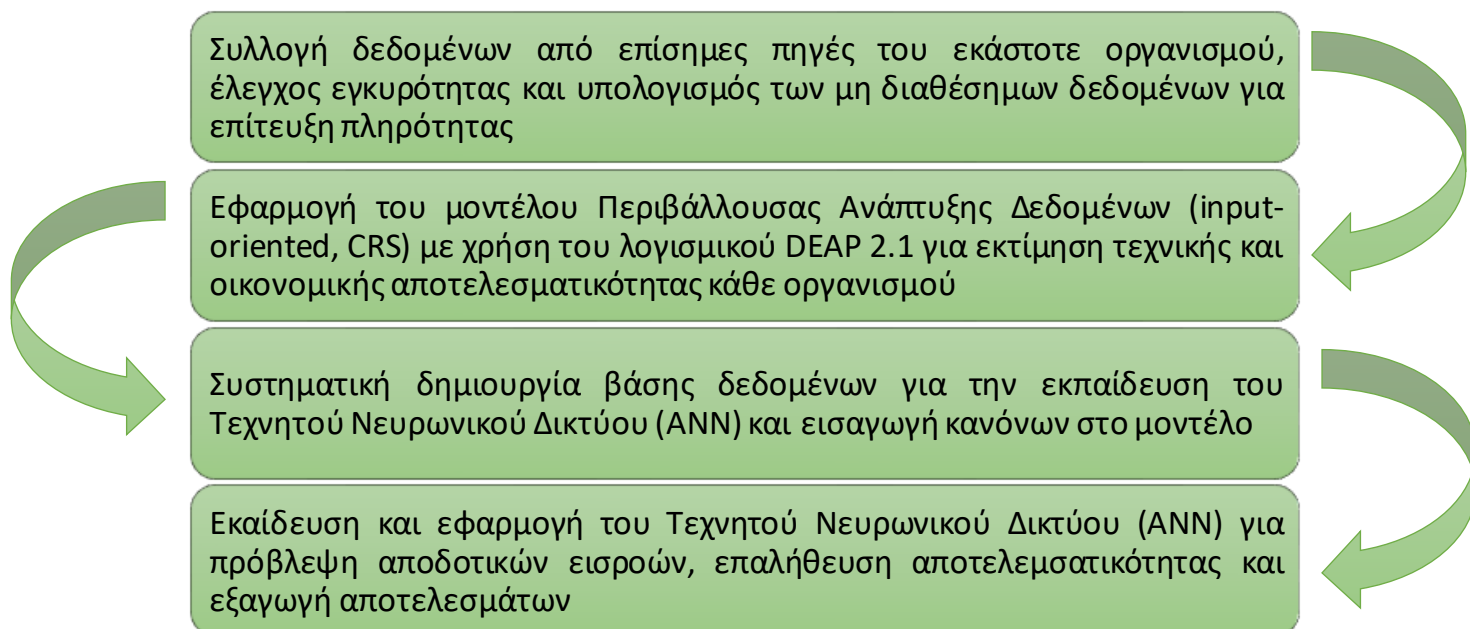
3. Μεθοδολογική προσέγγιση

3.1 Περιγραφή διαδικασίας

Στην παρούσα διπλωματική εργασία επιλέχθηκε μια μεθοδολογική προσέγγιση η οποία συνδυάζει δύο μεθόδους: α) την Περιβάλλουσα Ανάλυση Δεδομένων (Data Envelopment Analysis – DEA) μια μη παραμετρική μέθοδο βασισμένη στον γραμμικό προγραμματισμό, και β) τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN). Αυτό αποσκοπεί στη μελέτη δεκατριών σιδηροδρομικών εμπορευματικών οργανισμών στην Ευρώπη, προκειμένου να αξιολογηθεί η υφιστάμενη αποδοτικότητα τους και να διερευνηθούν πιθανές βελτιώσεις αυτής.

Για τη συλλογή των απαιτούμενων δεδομένων, τόσο εισροών όσο και εκροών, χρησιμοποιήθηκαν επίσημες πηγές, όπως ετήσιες εκθέσεις, ισολογισμοί και δημοσιευμένα άρθρα των ίδιων των οργανισμών. Σε περιπτώσεις όπου υπήρχαν κενά ή ασυνέπειες στα δεδομένα, πραγματοποιήθηκαν εκτιμήσεις και έγιναν υπολογισμοί με βάση συγκριτικά μεγέθη από αντίστοιχες εταιρείες του κλάδου. Ο τελικός πίνακας δεδομένων περιλάμβανε έξι μεταβλητές εισροών και μία μεταβλητή εκροής για κάθε οργανισμό, και χρησιμοποιήθηκε για την εκτέλεση της Περιβάλλουσας ανάλυσης δεδομένων.

Η εφαρμογή της Περιβάλλουσας ανάλυσης δεδομένων πραγματοποιήθηκε με το λογισμικό DEAP 2.1, λαμβάνοντας υπόψη σταθερές αποδόσεις κλίμακας (CRS) με προσανατολισμό προς τις εισροές (input-oriented). Η αποδοτικότητα εξετάστηκε σε δύο επίπεδα: αφενός την τεχνική αποδοτικότητα (με εκροή τα ton-kilometres) και αφετέρου την οικονομική (με εκροή τα έσοδα).



Εικόνα 3.1 Διάγραμμα ροής της μεθοδολογικής διαδικασίας

Παρότι η Περιβάλλουσα ανάλυσης δεδομένων παρέχει σημαντικές πληροφορίες για τη σχετική αποδοτικότητα των μονάδων, παρουσιάζει περιορισμούς, καθώς ενδέχεται να προτείνει θεωρητικά μεν αλλά πρακτικά ανέφικτα σενάρια βελτίωσης, ειδικά όσον

αφορά τις μειώσεις εισροών. Για να καλυφθεί αυτό το κενό, αξιοποιήθηκαν Τεχνητά Νευρωνικά Δίκτυα, τα οποία εκπαιδεύτηκαν με βάση τα αποτελέσματα της Περιβάλλουσας ανάλυσης δεδομένων.

Δημιουργήθηκαν εμπλουτισμένα σύνολα δεδομένων με σταθερή εκροή και ποικίλους συνδυασμούς εισροών, για τους οποίους υπολογίστηκε η αντίστοιχη αποδοτικότητα. Αυτά τα δεδομένα αποτέλεσαν τη βάση για την εκπαίδευση δύο Τεχνητών Νευρωνικών Δικτύων – ένα για κάθε τύπο εκροής – με σκοπό την πρόβλεψη εισροών που οδηγούν σε μέγιστη αποδοτικότητα καθώς και προτροπή για βελτίωση των οργανισμών με ρεαλιστικές λύσεις. Κατά την εκπαίδευση εφαρμόστηκαν ειδικοί κανόνες, ώστε τα αποτελέσματα να είναι συμβατά με τους στόχους της μελέτης.

Η τελική υβριδική προσέγγιση DEA–ANN αξιοποιεί τα πλεονεκτήματα και των δύο μεθόδων: την αναλυτική σαφήνεια της DEA και τη δυναμική προσαρμοστικότητα των νευρωνικών δικτύων. Έτσι, παρέχει ένα πιο ρεαλιστικό και εφαρμόσιμο πλαίσιο για την υποστήριξη στρατηγικών βελτιστοποίησης της αποδοτικότητας.

3.2 Θεωρητικό υπόβαθρο της Περιβάλλουσας Ανάλυσης Δεδομένων (Data Envelopment Analysis – DEA)

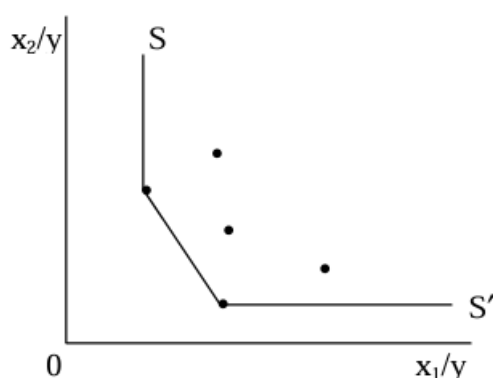
Η Μέθοδος της Περιβάλλουσας Ανάλυσης Δεδομένων (Data Envelopment Analysis - DEA) είναι μια μη παραμετρική τεχνική γραμμικού προγραμματισμού που χρησιμοποιείται για την εκτίμηση της σχετικής αποδοτικότητας ομοειδών μονάδων λήψης αποφάσεων (Decision Making Units - DMUs), οι οποίες αξιοποιούν πολλαπλές εισροές για την παραγωγή πολλαπλών εκροών. Η μονάδα λήψης απόφασης (DMU) μπορεί να είναι μια οποιαδήποτε οντότητα που χρησιμοποιεί πόρους για την παραγωγή αποτελεσμάτων· ενδεικτικά παραδείγματα αποτελούν τα εργοστάσια, τα σχολεία, τα νοσοκομεία ή – στην περίπτωση της παρούσας μελέτης – οι σιδηροδρομικοί οργανισμοί. Οι εισροές (inputs) της μεθόδου αφορούν τους πόρους που χρησιμοποιούνται, όπως είναι το προσωπικό, ο εξοπλισμός, η κατανάλωση ενέργειας και οι επενδυτικές δαπάνες, ενώ οι εκροές (outputs) εκφράζουν την τελική απόδοση ή παραγωγή της μονάδας, όπως είναι ο μεταφορικός όγκος ή τα έσοδα.

Η Μέθοδος της Περιβάλλουσας Ανάλυσης Δεδομένων, η οποία αναπτύχθηκε από τους Charnes, Cooper και Rhodes (1978) και βασίζεται στο έργο του Farrell (1957), έχει το πλεονέκτημα ότι δεν απαιτεί να γνωρίζουμε εκ των προτέρων τη μαθηματική σχέση που συνδέει τις εισροές με τις εκροές μιας παραγωγικής μονάδας. Αντιθέτως, βασίζεται σε εμπειρικά δεδομένα και προσδιορίζει το αποδοτικό μέτωπο (efficient frontier) μέσω της παρατήρησης των καλύτερων επιδόσεων του δείγματος. Η αποδοτικότητα (efficiency) κάθε DMU υπολογίζεται συγκριτικά και εκφράζεται με έναν συντελεστή $\theta \in (0,1]$, όπου η τιμή 1 υποδηλώνει πλήρη αποδοτικότητα, ενώ τιμές μικρότερες του 1 δηλώνουν περιθώριο βελτίωσης. Η DEA κατασκευάζει ένα περιβάλλον σύνολο (envelopment surface), το οποίο περιλαμβάνει τις αποδοτικές μονάδες, ενώ οι λιγότερο αποδοτικές τοποθετούνται εντός αυτού και αξιολογούνται ως προς την απόστασή τους από το σύνολο.

Για λόγους απλοποίησης και ενίσχυσης της κατανόησης της μεθόδου DEA, παρουσιάζεται ενδεικτικά ένα υποθετικό παράδειγμα με δύο εισροές (π.χ. εργατοώρες και κατανάλωση ενέργειας) και μία εκροή (παραχθείσα ποσότητα προϊόντος). Σε αυτό το παράδειγμα, κάθε μονάδα λήψης απόφασης (DMU) απεικονίζεται ως σημείο σε δισδιάστατο διάγραμμα, όπου οι συντεταγμένες αντιστοιχούν στα επίπεδα των δύο εισροών ανά μονάδα εκροής.

Η γραμμή που σχηματίζεται από τις μονάδες με την υψηλότερη δυνατή απόδοση ονομάζεται αποδοτικό μέτωπο (efficient frontier). Οι μονάδες που βρίσκονται εντός της περιοχής που ορίζεται από αυτό το μέτωπο συγκροτούν το λεγόμενο περιβάλλον σύνολο (envelopment surface). Σε αυτό το σύνολο εντάσσονται οι λιγότερο αποδοτικές μονάδες, οι οποίες αξιολογούνται σε σύγκριση με τις πιο αποδοτικές. Η απόσταση κάθε μη αποδοτικής μονάδας από το αποδοτικό μέτωπο λειτουργεί ως μέτρο της τεχνικής της αποδοτικότητας. Η απόδοση αυτή εκφράζεται ποσοτικά μέσω του συντελεστή θ (theta).

Το ενδεικτικό σχήμα που ακολουθεί οπτικοποιεί τα παραπάνω, αναδεικνύοντας τη σημασία του efficient frontier στον εντοπισμό τόσο των αποδοτικών όσο και των μη αποδοτικών μονάδων.



Εικόνα 3.2 Απεικόνιση αποδοτικού μετώπου και περιβάλλοντος συνόλου σε περίπτωση δύο εισροών και μίας εκροής (Πηγή: Farrell, M.J. (1957))

3.2.1 Μαθηματική Διατύπωση της DEA με παραδοχή Σταθερών Αποδόσεων Κλίμακας (CRS)

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, επιλέγεται η εφαρμογή της μεθόδου Περιβάλλουσας Ανάλυσης Δεδομένων (DEA) με παραδοχή σταθερών αποδόσεων κλίμακας (Constant Returns to Scale – CRS) και προσανατολισμό προς τις εισροές (input-oriented). Η επιλογή αυτή βασίζεται στη φύση των δεδομένων που εξετάζονται και στην επιθυμία να διερευνηθεί η τεχνική και οικονομική αποδοτικότητα των σιδηροδρομικών οργανισμών, υπό την παραδοχή ότι το μέγεθος της επιχείρησης δεν επηρεάζει την απόδοση – δηλαδή ότι οι αποδόσεις παραμένουν σταθερές ανεξαρτήτως κλίμακας λειτουργίας.

Για την πληρέστερη κατανόηση της μεθοδολογίας της Περιβάλλουσας Ανάλυσης Δεδομένων με υπόθεση σταθερών αποδόσεων κλίμακας, παρουσιάζεται στη συνέχεια η μαθηματική της διατύπωση. Η διατύπωση αυτή επιτρέπει τον υπολογισμό της σχετικής αποδοτικότητας κάθε μονάδας λήψης απόφασης (DMU) μέσα από ένα πρόβλημα γραμμικού προγραμματισμού, το οποίο επιλύεται επαναληπτικά για κάθε μονάδα του δείγματος. Αρχικά παρουσιάζεται η βασική αναλογική μορφή (ratio form), και εν συνεχεία ακολουθούν η μορφή των πολλαπλασιαστών (multiplier form) και η μορφή περιβάλλοντος συνόλου (envelopment form), η οποία είναι και αυτή που εφαρμόστηκε στο λογισμικό DEAP για τις ανάγκες της παρούσας μελέτης.

Η πρώτη μαθηματική διατύπωση του υποδείγματος της Περιβάλλουσας Ανάλυσης Δεδομένων με παραδοχή σταθερών αποδόσεων κλίμακας δίνεται μέσω της αναλογικής μορφής (ratio form), όπως προτάθηκε από τους Charnes, Cooper και Rhodes (1978), γνωστή και ως CCR μοντέλο. Σε αυτή τη διατύπωση, για κάθε υπό εξέταση μονάδα (μονάδα-στόχο) j_0 , επιχειρείται η μεγιστοποίηση του λόγου της σταθμισμένης εκροής προς τη σταθμισμένη εισροή, με κατάλληλη επιλογή των συντελεστών βαρύτητας u_r (για τις εκροές) και v_i (για τις εισροές).

Ο μαθηματικός τύπος της αποδοτικότητας (efficiency) της μονάδας j_0 , με m εισροές και s εκροές, δίνεται ως εξής:

$$\max_{u,v} \frac{\sum_{r=1}^s u_r y_{rj_0}}{\sum_{i=1}^m v_i x_{ij_0}}$$

Υπό τους περιορισμούς:

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, \text{ για κάθε } j = 1, 2, \dots, n$$

$$u_r, v_i \geq 0, \quad \forall r, i$$

Η μορφή αυτή έχει το πλεονέκτημα της διαισθητικής κατανόησης του σκοπού της DEA: να βρεθεί το μέγιστο ποσοστό «εκροής προς εισροή» για κάθε μονάδα-στόχο, υπό την προϋπόθεση ότι καμία άλλη μονάδα στο δείγμα δεν ξεπερνά αυτή την αναλογία.

Ωστόσο, ένα σημαντικό πρόβλημα αυτής της προσέγγισης είναι ότι επιτρέπει άπειρες λύσεις (multiplicative scaling), δηλαδή διαφορετικά σύνολα u, v μπορούν να δώσουν το ίδιο αποτέλεσμα. Για την αποφυγή του προβλήματος αυτού, προστίθεται ο περιορισμός:

$$\sum_{i=1}^m v_i x_{ij_0} = 1$$

και η μορφή γίνεται:

$$\max_{u,v} \sum_{r=1}^s u_r y_{rj_0}$$

Υπό τους περιορισμούς:

$$\sum_{i=1}^m u_r y_{rj} - \sum_{r=1}^s v_i x_{ij} \leq 0, \quad \forall j$$

Η εναλλακτική και πιο υπολογιστικά αποδοτική διατύπωση του μοντέλου DEA είναι η μορφή περιβάλλοντος συνόλου (envelopment form), η οποία βασίζεται στη δυϊκότητα της γραμμικής βελτιστοποίησης. Στη μορφή αυτή, αντί να αναζητούνται τα βέλτιστα βάρη για εισροές και εκροές (όπως στη μορφή αναλογίας - multiplier form), αναζητείται ο συντελεστής τεχνικής αποδοτικότητας θ , ο οποίος δείχνει σε ποιο βαθμό μπορούν να μειωθούν αναλογικά όλες οι εισροές μιας μονάδας, χωρίς να μειωθεί η εκροή της.

Η μαθηματική διατύπωση του προβλήματος DEA σε envelopment form για μια μονάδα-στόχο j_0 έχει ως εξής:

$$\min_{u,v} \theta$$

Υπό τους περιορισμούς:

$$-y_{rj_0} + \sum_{j=1}^n \lambda_j y_{rj} \geq 0 \quad \forall r = 1, \dots, s$$

$$\theta x_{ij_0} + \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad \forall i = 1, \dots, m$$

$$\lambda_j \geq 0 \quad \forall j$$

Όπου:

- x_{ij} : η τιμή της εισροής i για τη μονάδα j
- y_{rj} : η τιμή της εκροής r για τη μονάδα j
- λ_j : βάρη που χρησιμοποιούνται για τον προσδιορισμό της μονάδας αναφοράς
- θ : συντελεστής τεχνικής αποδοτικότητας της μονάδας j_0 , με $\theta \in (0,1]$
- n : αριθμός μονάδων λήψης απόφασης (DMUs)
- m : αριθμός εισροών
- s : αριθμός εκροών

Η τιμή του θ υποδηλώνει το βαθμό στον οποίο μπορούν να μειωθούν οι εισροές μιας μονάδας χωρίς να επηρεαστεί το παραγόμενο αποτέλεσμα (εκροές). Αν $\theta = 1$, η μονάδα θεωρείται τεχνικά αποδοτική, καθώς δεν μπορεί να επιτύχει την ίδια παραγωγή με λιγότερες εισροές. Αν $\theta < 1$, υπάρχει περιθώριο βελτίωσης, δηλαδή η μονάδα λειτουργεί κάτω από το αποδοτικό μέτωπο.

Η μορφή αυτή είναι προτιμότερη για υπολογιστικούς λόγους, καθώς εμπεριέχει λιγότερους περιορισμούς από τη μορφή αναλογίας ($K + M < N + 1$), και χρησιμοποιείται κυρίως στα λογισμικά πακέτα DEA όπως το DEAP (Coelli, 1996).

3.2.2 Υπολειπόμενες ποσότητες (Slacks) και τεχνική αποδοτικότητα

Αν και η τιμή του συντελεστή θ στον input-oriented DEA-CCR model αποτελεί το βασικό μέτρο τεχνικής αποδοτικότητας, από μόνη της δεν εγγυάται πλήρη αποδοτικότητα. Για να θεωρηθεί μια μονάδα πλήρως αποδοτική (efficient), θα πρέπει επιπλέον να μην παρουσιάζει υπολειπόμενες ποσότητες (slacks).

Οι υπολειπόμενες ποσότητες εμφανίζονται όταν, ακόμη και μετά την αναλογική μείωση των εισροών (μέσω του θ), η μονάδα δεν "προβάλλεται" ακριβώς πάνω στο αποδοτικό μέτωπο αλλά παραμένει εντός του περιβάλλοντος συνόλου. Οι slacks δείχνουν πόση επιπλέον μείωση εισροών (ή αύξηση εκροών) μπορεί να επιτευχθεί χωρίς να αλλάξει η αναλογία μεταξύ εισροών και εκροών.

Υπάρχουν δύο είδη slacks:

- Input slack: υποδεικνύει ότι κάποιες εισροές μπορούν να μειωθούν επιπλέον της μείωσης θ , χωρίς να επηρεαστεί η εκροή.

- Output slack: υποδεικνύει ότι η εκροή θα μπορούσε να αυξηθεί, διατηρώντας τις εισροές σταθερές.

Για την ακριβή μέτρηση των slacks, απαιτείται η επίλυση ενός δεύτερου προβλήματος γραμμικού προγραμματισμού, με σκοπό να ελαχιστοποιηθεί το σύνολο των υπολειπομένων ποσοτήτων, δηλαδή:

$$\min_{OS, IS} \theta (M'_1 OS + K'_1 IS)$$

Υπό τους περιορισμούς:

$$\begin{aligned} -y_i + Y\lambda - OS &= 0 \\ -\theta x_i + X\lambda - IS &= 0 \\ \lambda &\geq 0, \quad OS \geq 0, \quad IS \geq 0 \end{aligned}$$

Όπου:

- OS : output slacks (διανυσματική ποσότητα για κάθε εκροή),
- IS : input slacks (διανυσματική ποσότητα για κάθε εισροή),
- M_1, K_1 : διανύσματα μονάδων (όλων με τιμή 1).

Η παραπάνω διαδικασία αναφέρεται στη βιβλιογραφία ως second-stage DEA και προϋποθέτει ότι έχουν ήδη υπολογιστεί τα θ και λ από την πρώτη φάση. Το μοντέλο αυτό επιδιώκει να βρει το πλησιέστερο σημείο στο αποδοτικό μέτωπο μέσω της ελαχιστοποίησης των slacks (βλ. Coelli et al., 2005).

3.2.3 Προσεγγίσεις επίλυσης DEA με υπολογισμό Slacks

Η ανάλυση των υπολειπομένων ποσοτήτων (slacks) αποτελεί κρίσιμο συμπλήρωμα στον υπολογισμό της τεχνικής αποδοτικότητας μέσω της DEA. Αν και η τιμή του συντελεστή αποδοτικότητας θ προσδιορίζει το βαθμό ακτινικής (radial) βελτίωσης των εισροών, αυτή από μόνη της δεν εγγυάται πλήρη αποδοτικότητα, καθώς ενδέχεται να υπάρχουν περιθώρια περαιτέρω μη-ακτινικών βελτιώσεων στις εισροές ή/και στις εκροές. Για τον λόγο αυτό, η πλήρης αξιολόγηση μιας μονάδας προϋποθέτει τον προσδιορισμό και των slacks.

Στη βιβλιογραφία έχουν διαμορφωθεί τρεις βασικές προσεγγίσεις για την εκτίμηση της αποδοτικότητας και των slacks, κάθε μία με τα δικά της χαρακτηριστικά:

- Μονοσταδιακή (One-Stage DEA): Η πιο διαδεδομένη μορφή, όπου επιλύεται ένα μόνο πρόβλημα γραμμικού προγραμματισμού για την εκτίμηση του θ , και τα slacks υπολογίζονται δευτερευόντως, χωρίς να διασφαλίζεται η πλήρης εξάλειψή τους.
- Δισταδιακή (Two-Stage DEA): Σε αυτή την παραλλαγή, μετά την αρχική ανάλυση DEA, ακολουθεί ένα δεύτερο βήμα βελτιστοποίησης που στοχεύει στη μείωση των slacks (υπολοίπων εισροών ή εκροών). Έτσι, επιτυγχάνεται πιο ακριβής εκτίμηση της αποδοτικότητας κάθε μονάδας.
- Πολυσταδιακή (Multi-Stage DEA): Πρόκειται για μια πιο αναλυτική και ολοκληρωμένη μεθοδολογία η οποία περιλαμβάνει διαδοχικές προβολές των μονάδων στο αποδοτικό μέτωπο, τόσο ακτινικές (radial) όσο και μη ακτινικές (non-radial), με στόχο να εντοπιστεί το πλησιέστερο πλήρως αποδοτικό σημείο. Η συγκεκριμένη προσέγγιση δεν περιορίζεται μόνο στη μέτρηση της

ακτινικής αποδοτικότητας (Farrell efficiency), αλλά επεκτείνεται και στην Koormans-efficiency, και έτσι εξασφαλίζει ότι δεν υπάρχει σπατάλη εισροών ούτε ανεπάρκεια στις εκροές.

Στο πλαίσιο της παρούσας διπλωματικής εργασίας επιλέχθηκε η πολυσταδιακή προσέγγιση, όπως αυτή υλοποιείται στο λογισμικό DEAP (Coelli, 1996), καθώς παρέχει πιο πλήρη αξιολόγηση της αποδοτικότητας και επιτρέπει τον εντοπισμό των στοχευόμενων επιπέδων εισροών που θα καθιστούσαν κάθε μονάδα πλήρως αποδοτική. Η προσέγγιση αυτή κρίνεται κατάλληλη για τον χαρακτήρα της παρούσας ανάλυσης, η οποία συνδυάζει υπολογισμό slacks, στόχων και συγκριτική αξιολόγηση.

3.2.4 Υλοποίηση του υποδείγματος DEA στην πράξη (Λογισμικό DEAP)

Για την πρακτική υλοποίηση της μεθόδου Περιβάλλουσας Ανάλυσης Δεδομένων (DEA) με παραδοχή σταθερών αποδόσεων κλίμακας (CRS) και προσανατολισμό προς τις εισροές (input-oriented), χρησιμοποιήθηκε το εξειδικευμένο λογισμικό DEAP (Data Envelopment Analysis Program), το οποίο αναπτύχθηκε από τον Tim Coelli στο Πανεπιστήμιο του Queensland (Coelli, 1996).

Το λογισμικό DEAP προσφέρει ένα ευέλικτο και εύχρηστο περιβάλλον εργασίας για την εκτίμηση της σχετικής αποδοτικότητας ενός συνόλου μονάδων λήψης αποφάσεων (DMUs), βασισμένο στη μεθοδολογία των Charnes, Cooper και Rhodes (1978). Το λογισμικό υποστηρίζει τόσο μοντέλα με CRS όσο και με VRS (μεταβλητές αποδόσεις κλίμακας), καθώς και προσανατολισμό είτε προς τις εισροές είτε προς τις εκροές. Επιπλέον, επιλύει αυτόματα το δεύτερο στάδιο της ανάλυσης για τον εντοπισμό των υπολειπομένων ποσοτήτων (slacks) και των βέλτιστων επιπέδων εισροών και εκροών.

Η διαδικασία υλοποίησης περιλαμβάνει:

- Δημιουργία αρχείου δεδομένων (data file): Ο ερευνητής ορίζει σε αρχείο τύπου .txt τις τιμές των εκροών και των εισροών για κάθε μονάδα. Το αρχείο πρέπει να είναι διατεταγμένο κατά τρόπο ώστε στην πρώτη στήλη να περιλαμβάνονται οι εκροές (outputs) και στις επόμενες οι εισροές (inputs), με κάθε γραμμή να αντιστοιχεί σε μία μονάδα λήψης αποφάσεων. Στην παρούσα μελέτη εξετάστηκαν 13 σιδηροδρομικοί οργανισμοί, με μία εκροή (είτε το μεταφορικό έργο σε τονοχιλιόμετρα είτε τα έσοδα) και έξι εισροές.
- Καθορισμός των παραμέτρων μοντέλου: Μέσω του αρχείου εντολών (instruction file), καθορίζονται οι βασικές παράμετροι για την εκτέλεση του μοντέλου DEA, όπως ο αριθμός εισροών και εκροών, ο προσανατολισμός του μοντέλου (input-oriented) και η μορφή των αποδόσεων κλίμακας (Constant Returns to Scale – CRS). Παράλληλα, καθορίζονται τα αρχεία εισόδου και εξόδου της ανάλυσης και επιλέγεται εάν θα υπολογιστούν επιπλέον παράμετροι όπως τα slacks ή οι στόχοι (targets). Η παραμετροποίηση αυτή είναι απαραίτητη ώστε το λογισμικό να ευθυγραμμιστεί πλήρως με τη μεθοδολογική προσέγγιση που υιοθετείται στην παρούσα εργασία.
- Εκτέλεση ανάλυσης: Το DEAP επιλύει το πρόβλημα βελτιστοποίησης για κάθε μονάδα ξεχωριστά και επιστρέφει:
 - Τον συντελεστή αποδοτικότητας θ ,
 - Τα slacks εισροών και εκροών,
 - Τα στοχευόμενα επίπεδα εισροών και εκροών (targets),
 - Τις μονάδες αναφοράς (peers) για κάθε DMU.
- Ανάλυση αποτελεσμάτων: Τα αποτελέσματα της DEA χρησιμοποιούνται ως είσοδος σε μεταγενέστερα στάδια της παρούσας εργασίας, περιλαμβανομένης

της εκπαίδευσης τεχνητού νευρωνικού δικτύου με σκοπό την πρόβλεψη αποδοτικών συνδυασμών εισροών.

Η επιλογή του DEAP βασίστηκε στην ευρεία αποδοχή του στη διεθνή βιβλιογραφία, στην αξιοπιστία των αποτελεσμάτων του και στη δυνατότητα υλοποίησης πολυσταδιακής DEA, όπως απαιτείται για την πλήρη εξάλειψη των slacks και τον προσδιορισμό αποδοτικών στόχων

3.2.5 Πλεονεκτήματα και Μειονεκτήματα της Μεθόδου Περιβάλλουσας Ανάλυσης Δεδομένων (DEA)

Η μέθοδος της Περιβάλλουσας Ανάλυσης Δεδομένων (DEA) αποτελεί ένα ισχυρό εργαλείο για την αξιολόγηση της σχετικής αποδοτικότητας ομοειδών μονάδων, ιδιαίτερα σε περιβάλλοντα όπου συμμετέχουν πολλαπλές εισροές και εκροές χωρίς την ανάγκη καθορισμού εκ των προτέρων της λειτουργικής σχέσης μεταξύ τους. Μεταξύ των σημαντικότερων πλεονεκτημάτων συγκαταλέγεται η μη παραμετρική της φύση, καθώς δεν απαιτεί την υπόθεση κάποιας συγκεκριμένης μορφής συνάρτησης παραγωγής, η ευελιξία στην εφαρμογή της σε διαφορετικούς τομείς (δημόσιος και ιδιωτικός), καθώς και η ικανότητα εντοπισμού τόσο της τεχνικής αποδοτικότητας όσο και των επιμέρους πηγών αναποτελεσματικότητας μέσω των slacks.

Ωστόσο, η DEA παρουσιάζει και ορισμένα μειονεκτήματα, τα οποία χρήζουν ιδιαίτερης προσοχής. Πρώτον, η ευαισθησία της μεθόδου στην παρουσία ακραίων τιμών (outliers) ενδέχεται να επηρεάσει σημαντικά τον προσδιορισμό του αποδοτικού μετώπου. Επιπλέον, η σύγκριση των μονάδων γίνεται μόνο εντός του δείγματος, γεγονός που καθιστά τα αποτελέσματα αμιγώς σχετικά και όχι απόλυτα. Ένα ακόμη περιοριστικό στοιχείο αποτελεί η αδυναμία της DEA να προβλέψει αποδοτικούς συνδυασμούς εισροών για νέες μονάδες ή για μεταβολές στο επιχειρησιακό περιβάλλον, καθώς και η μη δυνατότητα γενίκευσης των αποτελεσμάτων σε άλλες συνθήκες εκτός δείγματος. Σημαντικό επίσης είναι το γεγονός ότι οι λύσεις βελτιστοποίησης που προτείνονται για τις μη αποδοτικές μονάδες – όπως οι προτεινόμενες μειώσεις εισροών ή αυξήσεις εκροών – βασίζονται σε γραμμικές προβολές επί του αποδοτικού μετώπου, και συχνά κρίνονται μη ρεαλιστικές ή επιχειρησιακά ανεφάρμοστες, ιδίως όταν προτείνουν δραστικές μεταβολές σε πολλαπλές εισροές ταυτόχρονα αλλά και σε εισροές που είναι ανέφικτες να μεταβληθούν.

Τα παραπάνω μειονεκτήματα, και ειδικότερα η αδυναμία πρόβλεψης και ρεαλιστικής στοχοθεσίας, οδήγησαν στη διερεύνηση συμπληρωματικών τεχνικών, όπως τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN), τα οποία παρέχουν τη δυνατότητα μάθησης από τα δεδομένα και πρόβλεψης αποδοτικών σεναρίων με βάση τη δομή του προβλήματος.

3.3 Θεωρητικό Υπόβαθρο των Τεχνητών Νευρωνικών Δικτύων (ANN)

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN) είναι μια τεχνική της μηχανικής μάθησης που, σε γενικές γραμμές, προσπαθεί να μιμηθεί τον τρόπο που επεξεργάζεται πληροφορίες ο ανθρώπινος εγκέφαλος. Δεν είναι εύκολο να εξηγηθεί με ακρίβεια, αλλά μπορούμε να πούμε ότι "μαθαίνουν" από δεδομένα, βρίσκουν συσχετίσεις και τελικά μπορούν να κάνουν προβλέψεις — ακόμα και χωρίς να ξέρουν εξ αρχής πώς σχετίζονται οι μεταβλητές. Αυτή η ιδιότητα τα κάνει χρήσιμα όταν έχουμε προβλήματα όπου εμπλέκονται πολυδιάστατες και μη γραμμικές σχέσεις, όπως συχνά

συμβαίνει με πραγματικά επιχειρησιακά σενάρια. Για παράδειγμα, μπορεί δύο εισροές να σχετίζονται με την εκροή με τρόπο που δεν μπορεί να μοντελοποιηθεί απλά.

Η χρήση των Τεχνητών Νευρωνικών Δικτύων στη παρούσα μελέτη αποσκοπεί στην αντιμετώπιση βασικών περιορισμών της Περιβάλλουσας Ανάλυσης Δεδομένων, και πιο συγκεκριμένα στην ανάγκη πρόβλεψης νέων αποδοτικών συνδυασμών εισροών, πέρα από όσους περιλαμβάνονται στο αρχικό δείγμα. Τα ANN μπορούν να εκπαιδευτούν βάσει των αποτελεσμάτων της DEA και να λειτουργήσουν ως ένα εργαλείο εξαγωγής σεναρίων, το οποίο λαμβάνει υπόψη τόσο τα πρότυπα αποδοτικότητας όσο και την πολυπλοκότητα των αλληλεπιδράσεων μεταξύ εισροών και εκροών. Στο υπόλοιπο του κεφαλαίου παρουσιάζεται το θεωρητικό υπόβαθρο των ANN, η εσωτερική τους δομή, ο τρόπος λειτουργίας τους, καθώς και η σύνδεσή τους με τη συγκεκριμένη εφαρμογή που υλοποιείται στο πλαίσιο της παρούσας διπλωματικής.

3.3.1 Βασικές Αρχές και Δομή Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks – ANN)

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN) αποτελούν ένα σύνολο υπολογιστικών μοντέλων εμπνευσμένων από τη λειτουργία του ανθρώπινου εγκεφάλου όπως προαναφέρθηκε. Στη βάση τους βρίσκεται ο τεχνητός νευρώνας: μια μονάδα που παίρνει αριθμητικές τιμές (inputs), τις πολλαπλασιάζει με κάποια βάρη (weights) και περνάει το αποτέλεσμα μέσα από μια λεγόμενη ενεργοποιητική συνάρτηση (activation function). Το αποτέλεσμα αυτό μεταφέρεται στους επόμενους νευρώνες και συνεχίζεται η διαδικασία. Συνήθως, ένα Τεχνητό Νευρωνικό Δίκτυο περιλαμβάνει τρία είδη επιπέδων (layers): Την είσοδο (input layer), που αποτελείται από τους νευρώνες οι οποίοι λαμβάνουν τα δεδομένα, τις κρυφές στοιβάδες (hidden layers), όπου γίνεται όλη η επεξεργασία — είναι σαν τα ενδιάμεσα βήματα σκέψης, και την έξοδο (output layer), που μας δίνει το τελικό αποτέλεσμα: π.χ. αν μια είσοδος ανήκει σε μια κατηγορία ή τι τιμή θα προβλεφθεί.

Όσο πιο πολλά επίπεδα και νευρώνες έχει το δίκτυο, τόσο καλύτερα μπορεί να χειριστεί δύσκολες σχέσεις. Δηλαδή, αντί να βλέπει μόνο απλές συνδέσεις, μπορεί να μάθει και περίπλοκα μοτίβα. Αυτό γίνεται επειδή οι νευρώνες στις κρυφές στοιβάδες χρησιμοποιούν συναρτήσεις όπως sigmoid, ReLU ή tanh, οι οποίες δεν είναι γραμμικές.

Κατά τη φάση της εκπαίδευσης, το δίκτυο χρησιμοποιεί έναν αλγόριθμο οπισθοδιάδοσης (backpropagation) ώστε να μειώσει το σφάλμα πρόβλεψης. Αυτό γίνεται με τη βοήθεια μιας συνάρτησης κόστους (loss function) και μιας διαδικασίας βαθμωτής βελτιστοποίησης (optimization) όπως η μέθοδος βαθμίδωσης (gradient descent). Το δίκτυο προσαρμόζει σταδιακά τα βάρη του, έτσι ώστε να μειώνει το σφάλμα μεταξύ προβλεπόμενων και πραγματικών τιμών.

Η ικανότητα των ANN να προσεγγίζουν μη γραμμικές και σύνθετες συναρτήσεις, καθιστά τα δίκτυα αυτά ιδιαίτερα αποτελεσματικά σε προβλήματα πρόβλεψης, αναγνώρισης προτύπων και, στην περίπτωση της παρούσας εργασίας, προσομοίωσης της λογικής της DEA για την πρόβλεψη αποδοτικών συνδυασμών εισροών.

3.3.2 Τύποι Νευρωνικών Δικτύων και Επιλογή Κατάλληλου Μοντέλου

Υπάρχουν διάφοροι τύποι τεχνητών νευρωνικών δικτύων, καθένας εκ των οποίων είναι κατάλληλος για την επίλυση διαφορετικών προβλημάτων. Η επιλογή του

κατάλληλου τύπου δικτύου εξαρτάται από τη φύση των δεδομένων, την πολυπλοκότητα των συσχετίσεων και τον στόχο του εκάστοτε προβλήματος.

1. Πλήρως συνδεδεμένα προωθητικά δίκτυα (Feedforward Neural Networks-FNNs): Τα FNNs είναι από τις απλούστερες μορφές τεχνητών νευρωνικών δικτύων. Σε αυτά, τα δεδομένα ρέουν μονόδρομα: ξεκινούν από το επίπεδο εισόδου, περνούν μέσα από ένα ή περισσότερα κρυφά επίπεδα και καταλήγουν στο επίπεδο εξόδου. Δεν υπάρχουν βρόχοι ή ανατροφοδοτήσεις (feedback). Αυτή η αρχιτεκτονική είναι κατάλληλη για προβλήματα όπου οι παρατηρήσεις είναι ανεξάρτητες μεταξύ τους, όπως η ταξινόμηση (classification) ή η παλινδρόμηση (regression) .

2. Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNNs): Τα RNNs έχουν σχεδιαστεί για την επεξεργασία ακολουθιακών δεδομένων, όπου η σειρά των στοιχείων έχει σημασία. Διαθέτουν εσωτερική μνήμη που τους επιτρέπει να "θυμούνται" προηγούμενες εισόδους και να τις χρησιμοποιούν για την επεξεργασία των επόμενων. Αυτό τα καθιστά ιδανικά για εφαρμογές όπως η αναγνώριση φωνής, η επεξεργασία φυσικής γλώσσας και η πρόβλεψη χρονοσειρών. Δεν ενδείκνυνται όμως για στατικές εισροές, όπως στην περίπτωση της παρούσας ανάλυσης.

3. Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNNs): Τα CNNs είναι εξειδικευμένα για την επεξεργασία δεδομένων με χωρική δομή, όπως εικόνες. Χρησιμοποιούν φίλτρα (kernels) που "συνελίσσονται" πάνω στα δεδομένα εισόδου για την εξαγωγή χαρακτηριστικών, όπως άκρες ή υφές. Αυτή η προσέγγιση τα καθιστά ιδιαίτερα αποτελεσματικά σε εφαρμογές αναγνώρισης εικόνων, ανίχνευσης αντικειμένων και ιατρικής απεικόνισης .

4. Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks – DNNs): Αποτελούν εξέλιξη των FNNs με πολλές κρυφές στοιβάδες. Είναι ικανά να μοντελοποιήσουν εξαιρετικά σύνθετες και αφηρημένες σχέσεις, αλλά απαιτούν μεγάλο όγκο δεδομένων και υψηλή υπολογιστική ισχύ. Είναι ιδιαίτερα χρήσιμα σε εφαρμογές όπως η αναγνώριση φωνής, η επεξεργασία φυσικής γλώσσας και η ανάλυση εικόνων .

Επιλογή κατάλληλου μοντέλου για την παρούσα εργασία: Για τις ανάγκες της διπλωματικής αυτής, όπου στόχος είναι η πρόβλεψη πολλαπλών εισροών X_2 – X_6 με δεδομένα X_1 και Y_2 , και η μοντελοποίηση της αποδοτικής σχέσης που αναδεικνύει η DEA, επιλέγεται ένα fully connected feedforward neural network. Ο τύπος αυτός προσφέρεται για στατικά αριθμητικά δεδομένα και έχει αποδειχθεί κατάλληλος σε αντίστοιχες εφαρμογές προσομοίωσης DEA (βλ. Kwon et al., 2017).

Το επιλεγμένο μοντέλο αναμένεται να προσεγγίσει τη λογική του DEA ως προς την πρόβλεψη των βέλτιστων εισροών, προσαρμόζοντας τα βάρη του έτσι ώστε να «μάθει» από τα πρότυπα των αποδοτικών μονάδων και να τα γενικεύσει σε νέες περιπτώσεις.

3.3.3 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου και Εφαρμογή στη DEA

Η εκπαίδευση του τεχνητού νευρωνικού δικτύου (ANN) στην παρούσα εργασία πραγματοποιείται με στόχο την προσομοίωση της λογικής της DEA και, συγκεκριμένα, την πρόβλεψη αποδοτικών τιμών εισροών (X_2 – X_6), δεδομένων της πρώτης εισροής (X_1) η οποία πρακτικά δεν μπορεί να μεταβληθεί, της εκροής (Y_2) και του συντελεστή τεχνικής αποδοτικότητας (θ) όπως αυτός υπολογίστηκε μέσω της DEA. Το ANN εκπαιδεύεται επομένως να "μάθει" το πρότυπο των αποδοτικών μονάδων και να προτείνει νέους, βελτιστοποιημένους συνδυασμούς εισροών που προσεγγίζουν ή επιτυγχάνουν πλήρη αποδοτικότητα.

Η αρχιτεκτονική που επιλέγεται είναι πλήρως συνδεδεμένο προωθητικό (fully connected feedforward) νευρωνικό δίκτυο με δύο κρυφές στοιβάδες των 64 νευρώνων έκαστη και συνάρτηση ενεργοποίησης ReLU, που είναι κατάλληλη για την εκμάθηση μη γραμμικών σχέσεων. Η έξοδος αποτελείται από πέντε συνεχείς μεταβλητές, χωρίς συνάρτηση ενεργοποίησης, όπως επιβάλλεται σε προβλήματα παλινδρόμησης (regression). Η εκπαίδευση βασίζεται στη χρήση του αλγορίθμου Adam για βελτιστοποίηση και του μέσου τετραγωνικού σφάλματος (MSE) ως συνάρτησης κόστους, με σκοπό τη σταδιακή ελαχιστοποίηση της διαφοράς μεταξύ προβλεπόμενων και στόχων εισροών.

Το σύνολο των δεδομένων που χρησιμοποιείται για την εκπαίδευση του ANN έχει δημιουργηθεί συστηματικά και περιλαμβάνει χιλιάδες παρατηρήσεις με ελαφρώς διαφοροποιημένες εισροές, εκροή Y_2 (Revenue ή BTK) και τιμές αποδοτικότητας θ , όπως υπολογίστηκαν μέσω του μοντέλου DEA. Η προσέγγιση αυτή βασίζεται στην πρακτική της έμμεσης εκμάθησης του DEA: τα αποτελέσματα της DEA λειτουργούν ως «δάσκαλος» και τα ANN εκπαιδεύονται ώστε να προσομοιώσουν τον μηχανισμό της.

Με τον τρόπο αυτό, το μοντέλο είναι σε θέση να προβλέπει νέες τιμές εισροών οι οποίες, εφόσον επανεισαχθούν σε DEA, αποφέρουν σημαντικά βελτιωμένες τιμές αποδοτικότητας για τις μονάδες που υστερούν, καλύπτοντας έτσι ένα σημαντικό μειονέκτημα της παραδοσιακής DEA: την αδυναμία πρόβλεψης και γενίκευσης.

3.3.4 Πλεονεκτήματα και Περιορισμοί της ANN μεθόδου

Σε αυτή τη μελέτη, χρησιμοποιούνται τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN) για να προβλέψουμε με μεγαλύτερη εγκυρότητα και ρεαλισμό, ποιοι συνδυασμοί εισροών είναι πιο αποτελεσματικοί. Τα ANN είναι ιδιαίτερα χρήσιμα γιατί μπορούν να χειριστούν πολύπλοκες σχέσεις που δεν είναι πάντα γραμμικές. Δηλαδή, όταν η σύνδεση ανάμεσα στις εισροές και τις εκροές δεν είναι απλή ή άμεσα προβλέψιμη, τα δίκτυα αυτά μπορούν να βοηθήσουν περισσότερο από άλλες, πιο “άκαμπτες” μεθόδους. Επιπλέον, το ANN έχει την ικανότητα γενικεύει τη λογική της DEA, μαθαίνοντας από το σύνολο των αποτελεσμάτων και παρέχοντας προβλέψεις ακόμη και για σενάρια που δεν υπήρχαν στο αρχικό δείγμα.

Ένα επιπλέον πλεονέκτημα είναι η ικανότητα πρόβλεψης του μοντέλου. Ενώ η DEA προσδιορίζει μόνο τη σχετική αποδοτικότητα βάσει υπάρχοντος δείγματος, το ANN μπορεί να προτείνει νέους συνδυασμούς εισροών που βελτιώνουν την αποδοτικότητα, προσφέροντας ουσιαστική καθοδήγηση για τη λήψη αποφάσεων και την ανακατανομή πόρων στον εκάστοτε οργανισμό. Επίσης, σε αντίθεση με τη DEA, τα ANN είναι πιο αποτελεσματικά στις ακραίες τιμές, έπειτα από κατάλληλη εκπαίδευση.

Ωστόσο, τα ANN παρουσιάζουν και ορισμένους περιορισμούς. Ένας από τους βασικότερους είναι η απαιτητικότητα σε δεδομένα. Η αποτελεσματική εκπαίδευση ενός ANN απαιτεί επαρκές πλήθος ποιοτικών δεδομένων, καθώς και προσεκτική επιλογή χαρακτηριστικών και παραμέτρων (όπως learning rate, epochs, αριθμός νευρώνων). Επιπλέον, στα ANN τα βάρη και οι εσωτερικοί τους μετασχηματισμοί δεν είναι πάντα προσιτά και ερμηνεύσιμα.

Τέλος, στην εκπαίδευση είναι πολύ σημαντικό να αποφεύγεται η υπερβολική προσαρμογή (overfitting) του μοντέλου, καθώς αυτό θα μειώσει την ικανότητά του να γενικεύσει σε νέα δεδομένα. Για την αντιμετώπιση αυτού του προβλήματος, εφαρμόζονται διάφορες τεχνικές, όπως η απλοποίηση της αρχιτεκτονικής του δικτύου,

η χρήση μεθόδων κανονικοποίησης (regularization), η τεχνική του dropout και η διακοπή της εκπαίδευσης σε πρώιμο στάδιο (early stopping).

Παρά τους περιορισμούς αυτούς, όταν σχεδιάζονται και εκπαιδεύονται κατάλληλα, τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να αποτελέσουν ισχυρά συμπληρωματικά εργαλεία σε μεθοδολογίες αξιολόγησης αποτελεσματικότητας, ενισχύοντας τη λήψη τεκμηριωμένων αποφάσεων και προσφέροντας ρεαλιστικότερες και πιο προσιτές λύσεις.

Μετά την παρουσίαση των θεωρητικών αρχών και της μεθοδολογίας που εφαρμόστηκε σε αυτή τη μελέτη, το επόμενο κεφάλαιο επικεντρώνεται στην πρακτική υλοποίηση των μοντέλων DEA και ANN. Η αποτελεσματικότητα κάθε μονάδας εξετάζεται μέσω ανάλυσης πραγματικών δεδομένων από σιδηροδρομικούς οργανισμούς και προτείνονται στοχευμένες βελτιώσεις, κάνοντας χρήση τόσο της γραμμικής λογικής του DEA όσο και της προγνωστικής ικανότητας που προσφέρουν τα τεχνητά νευρωνικά δίκτυα.

4. Ανάλυση και αποτελέσματα

4.1 Βάση δεδομένων

Η αξιοπιστία της ανάλυσης αποδοτικότητας των σιδηροδρομικών οργανισμών εξαρτάται σε μεγάλο βαθμό από την ποιότητα και την ακρίβεια των δεδομένων που χρησιμοποιούνται. Για τον λόγο αυτό, η συλλογή των δεδομένων πραγματοποιήθηκε από επίσημες πηγές, όπως ετήσιες και ενδιάμεσες εκθέσεις και επίσημους ιστότοπους των ίδιων των οργανισμών καθώς και εξειδικευμένες βάσεις δεδομένων. Η επιλογή των δεδομένων έγινε με γνώμονα τη διαθεσιμότητα, την πληρότητα και τη δυνατότητα σύγκρισης μεταξύ των εξεταζόμενων οργανισμών.

Στην παρούσα διπλωματική, δίνεται ιδιαίτερη έμφαση στην ομοιομορφία των δεδομένων ώστε να διασφαλιστεί η εγκυρότητα των αποτελεσμάτων. Η επιλογή των οργανισμών βασίστηκε σε συγκεκριμένα κριτήρια, τα οποία διασφαλίζουν ότι η ανάλυση καλύπτει ένα ευρύ φάσμα του ευρωπαϊκού σιδηροδρομικού τομέα, ενώ η επιλογή των μεταβλητών έγινε με στόχο την ακριβή αποτύπωση της παραγωγικής και οικονομικής αποδοτικότητας των οργανισμών.

4.1.1 Επιλογή οργανισμών-χωρών

Η επιλογή των χωρών που περιλαμβάνονται στην παρούσα μελέτη βασίστηκε στη σημασία των σιδηροδρομικών τους δικτύων στη διακίνηση εμπορευμάτων στην Ευρώπη. Οι επιλεγμένες χώρες διαθέτουν ανεπτυγμένα και στρατηγικά κρίσιμα σιδηροδρομικά συστήματα, τα οποία λειτουργούν ως βασικοί άξονες της διεθνούς εφοδιαστικής αλυσίδας. Το μέγεθος του σιδηροδρομικού δικτύου, η διασύνδεση με άλλες ευρωπαϊκές αγορές και η συμβολή τους στη συνολική εμπορευματική κίνηση αποτέλεσαν καθοριστικά κριτήρια για την επιλογή τους. Παράλληλα, η διαθεσιμότητα αξιόπιστων δεδομένων και στατιστικών στοιχείων επέτρεψε τη συγκριτική ανάλυση μεταξύ διαφορετικών χωρών, ενισχύοντας την εγκυρότητα της μελέτης.

Όσον αφορά την επιλογή των σιδηροδρομικών οργανισμών, κρίθηκαν καταλληλότερες, εταιρείες που κατέχουν κυρίαρχη θέση στην αγορά της αντίστοιχης χώρας, διασφαλίζοντας ότι διαχειρίζονται σημαντικούς όγκους εμπορευματικών μεταφορών. Η παρουσία τους στην αγορά και η έκταση των δραστηριοτήτων τους εξασφαλίζουν μια ολοκληρωμένη προσέγγιση στην ανάλυση της αποτελεσματικότητας. Παράλληλα, η διαθεσιμότητα δεδομένων ήταν ένας επιπλέον παράγοντας επιλογής, καθώς οι μεγαλύτεροι οργανισμοί δημοσιεύουν λεπτομερείς εκθέσεις, διευκολύνοντας την ποσοτική ανάλυση. Εξαίρεση αποτελεί η εταιρεία Hupac (Ελβετία), η οποία, αν και μικρότερου μεγέθους, εντάχθηκε στη μελέτη λόγω της ευκολίας εύρεσης δεδομένων. Στον παρακάτω πίνακα παρουσιάζονται οι επιλεγμένοι οργανισμοί και οι αντίστοιχες χώρες που δραστηριοποιούνται.

Πίνακας 4.1 Οργανισμοί που επιλέχθηκαν και οι αντίστοιχες χώρες

α/α	ΟΡΓΑΝΙΣΜΟΣ	ΧΩΡΑ
1	DB Cargo	Γερμανία
2	Xexafret (πρώην SNCF Fret)	Γαλλία
3	Renfe Mercancías	Ισπανία
4	CD Cargo	Τσεχία
5	OBB Rail Cargo Group	Αυστρία
6	ZSSK Cargo	Σλοβακία
7	Hupac	Ελβετία
8	Green Cargo	Σουηδία
9	PKP Cargo	Πολωνία
10	VR Transpoint	Φινλανδία
11	Rail Cargo Hungaria Zrt	Ουγγαρία
12	Mercitalia Rail	Ιταλία
13	DB Cargo Netherlands	Ολλανδία

4.1.2 Επιλογή δεδομένων

Η συλλογή των δεδομένων αποτέλεσε μια εξαιρετικά δύσκολη διαδικασία καθώς οι περισσότεροι οργανισμοί δεν δημοσιεύουν επαρκή δεδομένα σχετικά με τον εξοπλισμό, τον περιβαλλοντικό αντίκτυπο και πολλά από τα οικονομικά τους στοιχεία και έτσι, επιλέχθηκαν στοχευμένα, οργανισμοί από χώρες με τη μεγαλύτερη ευχέρεια εύρεσης δεδομένων. Παρόλα αυτά, πολλά από τα δεδομένα ακόμα και για αυτούς τους οργανισμούς δεν ήταν δημοσιευμένα και για αυτόν τον λόγο, για την εξασφάλιση της πληρότητας των δεδομένων, εφαρμόστηκαν κάποιες εύλογες παραδοχές οι οποίες αναλύονται στο Κεφάλαιο 4.1.4. Έχοντας υπόψιν τα παραπάνω, οι οργανισμοί που επιλέχθηκαν είναι συγκεκριμένοι προκειμένου να διασφαλιστεί η αξιοπιστία και η εγκυρότητα των δεδομένων. Για τον ίδιο λόγο επιλέχθηκαν και συγκεκριμένα δεδομένα εισόδου και εξόδου.

4.1.3 Επιλογή εισροών και εκροών

Η επιλογή των κατάλληλων εισροών και των εκροών αποτελεί θεμελιώδες στάδιο στην ανάλυση της αποδοτικότητας των σιδηροδρομικών οργανισμών. Η ορθολογική διαμόρφωση αυτού του συνόλου μεταβλητών διασφαλίζει ότι η αξιολόγηση αντικατοπτρίζει με ακρίβεια τόσο τους πόρους που χρησιμοποιούνται όσο και τα αποτελέσματα που επιτυγχάνονται. Ο καθορισμός των εισροών πραγματοποιείται με βάση τη συνολική λειτουργία και απόδοση των οργανισμών, ενώ οι εκροές επιλέγονται με γνώμονα την αποτύπωση της παραγωγικής και οικονομικής τους απόδοσης. Στην παρούσα διπλωματική, η επιλογή των μεταβλητών πραγματοποιείται με κριτήρια που

εξασφαλίζουν μια ολοκληρωμένη και αντικειμενική αξιολόγηση της αποδοτικότητας, γεγονός που επιτρέπει τη συγκριτική ανάλυση μεταξύ διαφορετικών οργανισμών. Παρακάτω παρουσιάζονται οι εισροές και εκροές που επιλέχθηκαν:

Εισροές

- Μήκος δικτύου (network length): Αφορά την έκταση της σιδηροδρομικής υποδομής και επηρεάζει την ικανότητα μεταφοράς φορτίων σε μεγάλες αποστάσεις.
- Αριθμός τροχαίου υλικού (rolling stock): Περιλαμβάνει τις διαθέσιμες αμαξοστοιχίες και βαγόνια του κάθε οργανισμού, καθορίζοντας τη μεταφορική ικανότητα και την ευελιξία στην εξυπηρέτηση των εμπορευμάτων.
- Αριθμός προσωπικού (staff number): Αναφέρεται στο ανθρώπινο δυναμικό που απαιτείται για τη λειτουργία, τη συντήρηση και τη διοικητική υποστήριξη του σιδηροδρομικού δικτύου.
- Κεφαλαιουχικές δαπάνες (CAPEX): Αφορούν επενδύσεις τόσο σε υφιστάμενες υποδομές όσο και σε νέα έργα. Σκοπός τους είναι να ενισχύσουν τη συνολική απόδοση, τη μεταφορική ικανότητα και την ανταγωνιστικότητα των σιδηροδρομικών εμπορευματικών υπηρεσιών.
- Λειτουργικά έξοδα (operational costs): Περιλαμβάνουν το σύνολο των εξόδων που σχετίζονται με την καθημερινή λειτουργία του οργανισμού, όπως η συντήρηση, οι μισθοί και οι δαπάνες διαχείρισης. Αποτελούν καθοριστικό παράγοντα για την οικονομική αποδοτικότητα.
- Ενεργειακή κατανάλωση (energy consumption): Αντιστοιχεί στην ενέργεια που καταναλώνεται για τη λειτουργία του τροχαίου υλικού. Παίζει σημαντικό ρόλο τόσο στο λειτουργικό κόστος όσο και στις περιβαλλοντικές επιπτώσεις της μεταφοράς.

Εκροές

- Όγκος μεταφοράς (volume transported): Αποτυπώνει τη συνολική ποσότητα εμπορευμάτων που μεταφέρθηκαν, λαμβάνοντας υπόψη τόσο το βάρος όσο και την απόσταση μεταφοράς, αποτελώντας βασικό δείκτη της επιχειρησιακής απόδοσης.
- Τζίρος εταιρείας (revenue): Αντικατοπτρίζει τα συνολικά έσοδα που προκύπτουν από τις σιδηροδρομικές εμπορευματικές μεταφορές, παρέχοντας μια οικονομική διάσταση στην αξιολόγηση της αποδοτικότητας των οργανισμών.

Οι μεταβλητές αυτές συλλέχθηκαν από τις επίσημες ιστοσελίδες του κάθε οργανισμού, ετήσιες εκθέσεις αυτών, θεσμικούς φορείς και εξειδικευμένες βάσεις δεδομένων. Ωστόσο κάποιες από τις μεταβλητές δεν ήταν δημοσιευμένες και για τον λόγω αυτό στο παρακάτω κεφάλαιο παρουσιάζονται όλες οι παραδοχές που έγιναν.

4.1.4 Παραδοχές-Υπολογισμοί

Στο παρόν κεφάλαιο παρουσιάζονται οι παραδοχές και οι υπολογισμοί που πραγματοποιήθηκαν, με στόχο τη διασφάλιση της πληρότητας αλλά και της εγκυρότητας των δεδομένων. Παρότι οι παραδοχές βασίζονται σε επίσημα δεδομένα και συγκριτικούς υπολογισμούς από άλλες Ευρωπαϊκές εταιρείες, είναι εύλογο να θεωρηθεί αποδεκτή μια απόκλιση της τάξης του 5%-10%, η οποία θα ληφθεί υπόψη στην τελική ανάλυση.

Στη συνέχεια, εξετάζονται αποκλειστικά τα δεδομένα που αφορούν εταιρείες για τις οποίες δεν υπήρχε άμεση πρόσβαση σε πρωτογενή στοιχεία. Σε αυτές τις περιπτώσεις, οι υπολογισμοί βασίστηκαν σε δημοσιευμένα δεδομένα, συγκριτικές αναλύσεις και έγκυρες εκτιμήσεις, με σκοπό τη διατήρηση της συγκρισιμότητας και της συνοχής των αποτελεσμάτων. Τυχόν αποκλίσεις και σφάλματα που ενδέχεται να προκύψουν λόγω της έλλειψης άμεσων και ακριβή στοιχείων λαμβάνονται υπόψη στη διαμόρφωση των τελικών συμπερασμάτων.

1. Κεφαλαιουχικές Δαπάνες (CAPEX)

- Το 2023, η **Fret SNCF** βρισκόταν σε αναδιάρθρωση, γεγονός που πιθανότατα επηρέασε τις επενδύσεις της. Υποθέτοντας μια συσχέτιση μεταξύ Κεφαλαιουχικών Δαπανών και Τζίρου μεταξύ 10% και 20%, η εκτίμηση για τις κεφαλαιουχικές της δαπάνες κυμαίνεται από 70 εκατ. ευρώ (συντηρητική εκτίμηση, 10% των εσόδων) έως 143 εκατ. ευρώ (επιθετική εκτίμηση, 20% των εσόδων), με έναν μέσο όρο στα 107 εκατ. ευρώ (15% των εσόδων). Συγκριτικά, για την DB Cargo, οι Κεφαλαιουχικές της Δαπάνες το 2023 αποτελούσαν περίπου το 5,7% των εσόδων της, ενώ αν μια εταιρεία βρίσκεται υπό αναδιάρθρωση το ποσοστό αυτό μπορεί να είναι και υψηλότερο. Λαμβάνοντας υπόψη όλους αυτούς τους παράγοντες, μια εύλογη εκτίμηση για τις κεφαλαιουχικές δαπάνες της Fret SNCF το 2023 είναι περίπου 105 εκατ. ευρώ.
- Σύμφωνα με την ενδιάμεση οικονομική έκθεση για το Α' εξάμηνο του 2023, το συνολικό CAPEX του Ομίλου České Dráhy ανερχόταν σε 7,343 εκατ. CZK (~€300 εκατ.). Υποθέτοντας ότι το ετήσιο CAPEX του Ομίλου παρέμεινε στα ίδια επίπεδα, η εκτίμηση για το σύνολο του 2023 ανέρχεται σε €600 εκατ. Η **CD Cargo**, ως θυγατρική εμπορευματικών μεταφορών του Ομίλου, συνήθως αντιπροσωπεύει το 25-35% των συνολικών Κεφαλαιουχικών Δαπανών, με βάση τις οικονομικές αναφορές από παλαιότερα έτη. Με την υπόθεση ότι το ίδιο ακριβώς ποσοστό διατηρήθηκε και το 2023, όπου είναι αρκετά ακριβής, τότε η εκτίμηση για τις Κεφαλαιουχικές Δαπάνες της ČD Cargo κυμαίνονται μεταξύ €150-210 εκατ., με μια μέση τιμή στα €180 εκατ..
- Το σύνολο των Κεφαλαιουχικών Δαπανών της **ZSSK Cargo** για το 2023 ανήλθε σε €69,822 εκατ., βάσει των επιμέρους επενδυτικών κατηγοριών που παρουσιάζονται παρακάτω. Η εταιρεία διέθεσε €1,181 εκατ. σε άυλα περιουσιακά στοιχεία, όπως λογισμικό και ψηφιακές πλατφόρμες, €30,815 εκατ. σε υλικά πάγια, περιλαμβάνοντας την αγορά και αναβάθμιση μηχανών, βαγονιών και εγκαταστάσεων συντήρησης και €37,826 εκατ. σε δικαιώματα χρήσης, κυρίως μέσω μακροχρόνιων συμβολαίων μίσθωσης για τροχαίο υλικό και σχετικές υποδομές. Το άθροισμα αυτών των ποσών οδηγεί στο συνολικό ύψος επενδύσεων, με σημαντική εστίαση στην αναβάθμιση του τροχαίου υλικού και την ενίσχυση της τεχνολογικής υποδομής. Αξιοσημείωτο είναι ότι το 91,5% των επενδύσεων ευθυγραμμίζεται με την EU Taxonomy, γεγονός που αποδεικνύει τη στρατηγική στόχευση της εταιρείας στη βιωσιμότητα και την ενεργειακή αποδοτικότητα. Οι παραπάνω ενέργειες αντανakλούν τον γενικότερο σχεδιασμό της ZSSK Cargo για εκσυγχρονισμό, καλύτερη αποδοτικότητα και προσαρμογή στις σύγχρονες τεχνολογικές και περιβαλλοντικές απαιτήσεις που ισχύουν στον τομέα των σιδηροδρομικών εμπορευματικών μεταφορών.

- Οι Κεφαλαιουχικές Δαπάνες της **Green Cargo** για το 2023 δεν έχουν δημοσιευθεί επισήμως, ωστόσο μπορούν να εκτιμηθούν βάσει ιστορικών δεδομένων και συγκριτικής ανάλυσης με άλλες σιδηροδρομικές εμπορευματικές εταιρείες. Η εταιρεία κατέγραψε έσοδα ύψους 4,2 δισ. SEK και λειτουργική ζημία 44 εκατ. SEK, ενώ το μεταφορικό της έργο ανήλθε σε 10,2 δισ. τονοχιλιόμετρα. Ιστορικά, η Green Cargo επενδύει 7-10% των εσόδων της σε κεφαλαιουχικές δαπάνες, ενώ ο μέσος όρος του κλάδου κυμαίνεται μεταξύ 10-15%.
Υποθέτοντας ότι οι Κεφαλαιουχικές δαπάνες της Green Cargo το 2023 βρίσκονταν μεταξύ 8-12%, μια εκτίμηση αρκετά ρεαλιστική, τότε καταλήγουμε στα 336 - 504 εκατ. με μέσο όρο τα 420 εκατ. SEK ή 36,5 εκατ. EUR, με βάση την ισοτιμία SEK/EUR του 2023. Το ποσό αυτό αφορά κυρίως επενδύσεις σε τροχάιο υλικό, ενεργειακά αποδοτικές λύσεις και υποδομές. Η εκτίμηση αυτή βασίζεται σε διαθέσιμα οικονομικά στοιχεία και μέσους δείκτες του κλάδου.
- Για την εκτίμηση των Κεφαλαιουχικών Δαπανών της **Rail Cargo Hungaria** δεν υπάρχουν δημοσιευμένα στοιχεία. Υπάρχουν όμως για τον όμιλο στον οποίο ανήκει. Έτσι για το 2023, μπορούμε να χρησιμοποιήσουμε την αναλογία του μεταφορικού της όγκου σε σχέση με όμιλο Rail Cargo Austria. Σύμφωνα με τα διαθέσιμα δεδομένα, η Rail Cargo Hungaria διαχειρίστηκε περίπου 19,5 εκατ. τόνους, από το συνολικό φορτίο της Rail Cargo Austria, το οποίο ανήλθε σε 78,5 εκατ. τόνους, αντιπροσωπεύοντας έτσι περίπου 24,8% του συνολικού μεταφορικού έργου. Δεδομένου ότι το συνολικό CAPEX της Rail Cargo Austria για το 2023 ανήλθε στα 170,5 εκατ. ευρώ, η ίδια αναλογία μπορεί να εφαρμοστεί για την εκτίμηση της επενδυτικής δραστηριότητας της Rail Cargo Hungaria. Με αυτή τη μεθοδολογία, το εκτιμώμενο CAPEX της Rail Cargo Hungaria ανέρχεται περίπου σε 42,3 εκατ. ευρώ (170,5 εκατ. ευρώ × 24,8%). Αυτή η εκτίμηση βασίζεται στη σχέση των μεταφορικών όγκων και αποτελεί έναν εύλογο δείκτη των επενδύσεων της Rail Cargo Hungaria εντός του ομίλου Rail Cargo Austria.
- Ο υπολογισμός των Κεφαλαιουχικών Δαπανών της **DB Cargo Netherlands** βασίστηκε, όπως και προηγουμένως, σε αναλογική προσέγγιση, καθώς δεν υπήρχαν άμεσα διαθέσιμα στοιχεία για τη συγκεκριμένη θυγατρική εταιρεία. Για την εκτίμηση, χρησιμοποιήθηκαν τα συνολικά στοιχεία του Ενοποιημένου Ετήσιου Απολογισμού της Deutsche Bahn για το 2023, όπου οι συνολικές μικτές κεφαλαιουχικές δαπάνες ανήλθαν σε 16.867 εκατ. ευρώ. Στη συνέχεια, υπολογίστηκε το ποσοστό συμμετοχής της DB Cargo Netherlands στις συνολικές δραστηριότητες του ομίλου, με βάση τη σχέση των εσόδων της (137 εκατ. ευρώ) προς τα συνολικά έσοδα του ομίλου (45.191 εκατ. ευρώ). Εφαρμόζοντας αυτό το ποσοστό στις συνολικές κεφαλαιουχικές δαπάνες, το εκτιμώμενο ποσό της DB Cargo Netherlands ανέρχεται σε περίπου 51,1 εκατ. ευρώ. Η συγκεκριμένη μεθοδολογία βασίζεται στην υπόθεση ότι οι επενδύσεις που γίνονται στον όμιλο, κατανέμονται αναλογικά με τα έσοδα των επιμέρους θυγατρικών.

2. Λειτουργικά έξοδα

Η εκτίμηση των λειτουργικών εξόδων πραγματοποιήθηκε βάσει του τύπου Έσοδα – ΚΠΤΦ (Κέρδη προ Τόκων και Φόρων), Revenue – EBIT (Earnings Before Interest and Taxes). Η επιλογή του συγκεκριμένου τύπου διασφαλίζει ότι το προκύπτον αποτέλεσμα αποτυπώνει το σύνολο των λειτουργικών

εξόδων των υπό εξέταση οργανισμών, περιλαμβάνοντας τόσο τα χρηματικά έξοδα (λειτουργικές δαπάνες) όσο και τα μη χρηματικά έξοδα, όπως αποσβέσεις και απομειώσεις. Κατ' αυτόν τον τρόπο, υπολογίζεται το συνολικό κόστος λειτουργίας των οργανισμών, παρέχοντας μια ολοκληρωμένη εικόνα των δαπανών που απαιτούνται για τη δραστηριότητά τους.

3. Ενεργειακή κατανάλωση

- Η **Fret SNCF** δεν δημοσιεύει επίσημα στοιχεία για τις εκπομπές CO₂e ανά τονοχιλιόμετρο (gCO₂e/tkm). Ωστόσο, δεδομένου ότι περίπου το 90% των εμπορευματικών της μεταφορών πραγματοποιείται με ηλεκτρικές αμαξοστοιχίες, οι εκπομπές της εκτιμάται ότι είναι παρόμοιες ή χαμηλότερες από τον εθνικό μέσο όρο. Σύμφωνα με το Γαλλικό Υπουργείο Οικολογικής Μετάβασης, οι μέσες εκπομπές CO₂ για τις σιδηροδρομικές εμπορευματικές μεταφορές στη Γαλλία ανέρχονται σε 1,99 gCO₂e/tkm, λόγω του μεγάλου ποσοστού ηλεκτροκίνησης του σιδηροδρομικού δικτύου της χώρας. Συνεπώς, δεδομένης της έλλειψης άμεσων δεδομένων και της χαμηλής έντασης άνθρακα του γαλλικού σιδηροδρόμου, η παρούσα μελέτη υιοθετεί το 1,99 gCO₂e/tkm ως αντιπροσωπευτική εκτίμηση των εκπομπών CO₂e της Fret SNCF.
- Η **Renfe Mercancías**, ο εμπορευματικός κλάδος του ομίλου Renfe, έχει δεσμευτεί στη βιώσιμη μεταφορά, και εφαρμόζει υψηλά ποσοστά ηλεκτροκίνησης και ενεργειακής απόδοσης. Περίπου 65% του στόλου της είναι ηλεκτρικός, ενώ η εταιρεία χρησιμοποιεί εξ ολοκλήρου πράσινη ηλεκτρική ενέργεια για τις ηλεκτροκίνητες μεταφορές της. Σύμφωνα με μελέτες, οι εκπομπές CO₂ των σιδηροδρομικών εμπορευματικών μεταφορών στην Ισπανία εκτιμώνται σε 8,3 gCO₂e/tkm, ενώ η Renfe Mercancías έχει καταγράψει επίπεδα εκπομπών έως και 5,54 gCO₂e/tkm. Λαμβάνοντας υπόψη αυτούς τους δείκτες, μια ρεαλιστική εκτίμηση των εκπομπών CO₂ της Renfe Mercancías είναι 6,9 gCO₂e/tkm, βάσει του μέσου όρου των διαθέσιμων τιμών.
- Η **CD Cargo**, ο εμπορευματικός κλάδος των Τσεχικών Σιδηροδρόμων (České dráhy), κατέγραψε συνολικές εκπομπές CO₂ ύψους 137.572,7 τόνων CO₂e το 2023. Με συνολικό μεταφορικό έργο 19,05 δισεκατομμυρίων τονοχιλιομέτρων (tkm). Έτσι, οι εκπομπές CO₂ ανά τονοχιλιόμετρο υπολογίζονται σε 7,22 gCO₂e/tkm. Η τιμή αυτή είναι χαμηλότερη από τον ευρωπαϊκό μέσο όρο, γεγονός που υποδηλώνει ότι η ČD Cargo λειτουργεί με υψηλή ενεργειακή απόδοση και μειωμένο ανθρακικό αποτύπωμα. Αυτή η επίδοση αποδίδεται κυρίως στον υψηλό βαθμό ηλεκτροκίνησης του δικτύου και στη βελτίωση των τεχνολογιών εξοικονόμησης ενέργειας.
- Η **ZSSK Cargo**, που είναι η μεγαλύτερη εταιρεία σιδηροδρομικών εμπορευματικών μεταφορών στη Σλοβακία, δεν δημοσιεύει επίσημα δεδομένα για τις εκπομπές CO₂ ανά τονοχιλιόμετρο (gCO₂e/tkm). Ωστόσο, μια εκτίμηση μπορεί να γίνει βάσει διαθέσιμων στοιχείων και συγκρίσεων με αντίστοιχες εταιρείες της Κεντρικής Ευρώπης. Η ČD Cargo στην Τσεχία, η οποία λειτουργεί σε παρόμοιο περιβάλλον, καταγράφει εκπομπές 7,22 gCO₂e/tkm. Δεδομένου ότι η Σλοβακία διαθέτει καθαρότερο ενεργειακό μείγμα, με υψηλότερη συμμετοχή της πυρηνικής και των ανανεώσιμων πηγών ενέργειας, οι εκπομπές της ZSSK Cargo εκτιμώνται χαμηλότερες από αυτές της ČD Cargo. Λαμβάνοντας υπόψη όλα αυτά τα δεδομένα, μια ρεαλιστική και ακριβής εκτίμηση για τις εκπομπές CO₂ της ZSSK Cargo είναι 6,5 gCO₂e/tkm. Αυτή η εκτίμηση, αντικατοπτρίζει την υψηλή ενεργειακή απόδοση του σλοβακικού σιδηροδρομικού δικτύου και τη χαμηλή εξάρτηση από ορυκτά καύσιμα.

- Η **Hupac**, μία από τις κορυφαίες εταιρείες συνδυασμένων μεταφορών στην Ευρώπη, καταγράφει σημαντικά μειωμένες εκπομπές CO₂e στις σιδηροδρομικές εμπορευματικές της μεταφορές. Σύμφωνα με τα επίσημα δεδομένα της για το 2023, οι συνολικές εκπομπές CO₂e από τις σιδηροδρομικές μεταφορές της ανήλθαν σε 137.542,5 τόνους CO₂e, ενώ το συνολικό μεταφορικό έργο έφτασε τα 13,3 δισεκατομμύρια τονοχιλιόμετρα (tkm). Βάσει αυτών των στοιχείων, οι εκπομπές CO₂e της Hupac υπολογίζονται σε 10,34 gCO₂e/tkm, επιβεβαιώνοντας την υψηλή περιβαλλοντική απόδοση της εταιρείας σε σύγκριση με τις οδικές μεταφορές. Η επίδοση αυτή ενισχύει τη στρατηγική της Hupac για τη μείωση του αποτυπώματος άνθρακα στις εμπορευματικές μεταφορές στην Ευρώπη.
- Η **Rail Cargo Hungaria Zrt.**, την μεγαλύτερη εταιρεία σιδηροδρομικών εμπορευματικών μεταφορών στην Ουγγαρία, καταβάλλει σημαντικές προσπάθειες στη μείωση των εκπομπών CO₂e μέσω ενεργειακά αποδοτικών πρακτικών. Με βάση τα διαθέσιμα στοιχεία, οι συνολικές εκπομπές CO₂e από τις δραστηριότητές της ανήλθαν σε 81.679 τόνους, ενώ το συνολικό μεταφορικό έργο έφτασε τα 5,85 δισεκατομμύρια τονοχιλιόμετρα (tkm). Με αυτά τα δεδομένα, οι εκπομπές CO₂e της Rail Cargo Hungaria υπολογίζονται σε 13,96 gCO₂e/tkm. Επιπλέον, η εταιρεία εφαρμόζει στρατηγικές βελτίωσης της ενεργειακής της απόδοσης, όπως η ανάκτηση ενέργειας κατά την πέδηση, ενισχύοντας τη δέσμευσή της για βιώσιμες μεταφορές στην Ουγγαρία και την ευρύτερη περιοχή.
- Βάσει διαθέσιμων στοιχείων, οι εκπομπές CO₂e ανά τονοχιλιόμετρο (gCO₂e/tkm) της **DB Cargo Netherlands** εκτιμώνται μεταξύ 1-2 gCO₂e/tkm. Αυτό δικαιολογείται λόγω της χρήσης 100% πράσινης ηλεκτρικής ενέργειας για τις περισσότερες μεταφορές της. Ο κλάδος σιδηροδρομικών εμπορευματικών μεταφορών στην Ολλανδία εμφανίζει σημαντικά χαμηλότερες εκπομπές σε σύγκριση με άλλες ευρωπαϊκές χώρες, γεγονός που αποδίδεται στην εκτεταμένη χρήση ηλεκτροκίνησης. Αν και τα ακριβή στοιχεία της εταιρείας δεν είναι διαθέσιμα, η εκτίμηση αυτή ευθυγραμμίζεται με τις γενικές τιμές εκπομπών για τον τομέα των ηλεκτροκίνητων εμπορευματικών σιδηροδρομικών μεταφορών στην Ολλανδία.

4. Όγκος μεταφοράς

- Με βάση τα διαθέσιμα στοιχεία, η **Fret SNCF** φαίνεται πως το 2023 μετέφερε περίπου 16,7 δισεκατομμύρια τονοχιλιόμετρα (tkm). Πρόκειται για πτώση της τάξης του 16,7% σε σχέση με την προηγούμενη χρονιά, όπου η δραστηριότητά της είχε φτάσει περίπου τα 20 δισ. tkm. Αυτή η μείωση αποτυπώνει τις δυσκολίες που αντιμετώπισε η εταιρεία στη διάρκεια του έτους, αλλά ίσως και μια γενικότερη τάση μείωσης στις εμπορευματικές σιδηροδρομικές μεταφορές στη Γαλλία.
- Βάσει των διαθέσιμων δεδομένων, η **CD Cargo** μετέφερε περίπου 59,4 εκατομμύρια τόνους φορτίου το 2023, σημειώνοντας μείωση κατά 4,8 εκατομμύρια τόνους σε σύγκριση με το προηγούμενο έτος. Το 2021, η εταιρεία είχε μεταφέρει περίπου 20,6 δισεκατομμύρια τονοχιλιόμετρα (tkm) εντός της Τσεχίας. Λαμβάνοντας υπόψη όλα αυτά αλλά και τη μείωση της μεταφερόμενης ποσότητας κατά 7,5%, η εκτίμηση για το μεταφορικό έργο του 2023 ανέρχεται σε περίπου 19,05 δισεκατομμύρια tkm.

- Βάσει των διαθέσιμων δεδομένων, η **ÖBB Rail Cargo Group** μετέφερε συνολικά 78,5 εκατομμύρια καθαρούς τόνους φορτίου το 2023. Αν και τα επίσημα τονοχιλιόμετρα (tkm) δεν αναφέρονται ρητά στην έκθεση, μια ακριβής εκτίμηση μπορεί να γίνει με βάση τη μέση απόσταση μεταφοράς. Ιστορικά, η RCA καταγράφει μέση απόσταση μεταφοράς περίπου 264 km ανά τόνο, καθώς δραστηριοποιείται έντονα σε διεθνείς και περιφερειακές μεταφορές. Λαμβάνοντας υπόψη αυτή τη μέση τιμή, η εκτιμώμενη συνολική μεταφορική απόδοση της ÖBB Rail Cargo Group για το 2023 ανέρχεται σε περίπου 20,7 δισεκατομμύρια τονοχιλιόμετρα (tkm).
- Η **Rail Cargo Hungaria Zrt.** μετέφερε το 2023 περίπου 19,5 εκατομμύρια καθαρούς τόνους φορτίου, σύμφωνα με τα διαθέσιμα δεδομένα. Στην επίσημη έκθεση δεν αναφέρεται πόσα ήταν τα τονοχιλιόμετρα, αλλά μπορεί να γίνει μια εκτίμηση με βάση τη μέση απόσταση που καλύπτεται ανά τόνο. Δεδομένου ότι η εταιρεία κάνει κυρίως διεθνείς μεταφορές στην Κεντρική και Ανατολική Ευρώπη, μια λογική εκτίμηση είναι πως κάθε τόνος διανύει γύρω στα 300 χιλιόμετρα. Αν πολλαπλασιάσουμε αυτό με τον όγκο του φορτίου, προκύπτει ότι το μεταφορικό της έργο για το 2023 φτάνει περίπου τα 5,85 δισεκατομμύρια τονοχιλιόμετρα (tkm).
- Η **Hupac** μετέφερε συνολικά 22,16 εκατομμύρια τόνους φορτίου το 2023. Τα επίσημα τονοχιλιόμετρα (tkm) δεν αναφέρονται ρητά στην έκθεση της εταιρείας, όμως μια ακριβής εκτίμηση μπορεί να γίνει με βάση τη μέση απόσταση μεταφοράς που ισχύει για τις διεθνείς διατροπικές μεταφορές. Η Hupac εξειδικεύεται στις διατροπικές μεταφορές μεγάλων αποστάσεων, γεγονός που συνεπάγεται μεγαλύτερη μέση απόσταση μεταφοράς σε σύγκριση με τις παραδοσιακές σιδηροδρομικές εμπορευματικές μεταφορές χύδην φορτίου. Δεδομένων των διαδρομών της στην Ευρώπη, μια μέση απόσταση μεταφοράς 600 km ανά τόνο θεωρείται ρεαλιστική εκτίμηση. Λαμβάνοντας υπόψη αυτή την τιμή, η εκτιμώμενη συνολική μεταφορική απόδοση της Hupac για το 2023 ανέρχεται σε περίπου 13,3 δισεκατομμύρια τονοχιλιόμετρα (tkm).
- Βάσει των διαθέσιμων δεδομένων, η **VR Transpoint** μετέφερε συνολικά 23,4 εκατομμύρια καθαρούς τόνους φορτίου το 2023. Αν και τα επίσημα τονοχιλιόμετρα (tkm) δεν αναφέρονται ρητά στην έκθεση, μια ακριβής εκτίμηση μπορεί να γίνει με βάση τη μέση απόσταση μεταφοράς. Ιστορικά, οι σιδηροδρομικές μεταφορές στη Φινλανδία πραγματοποιούνται σε αποστάσεις που κυμαίνονται μεταξύ 350-400 km ανά τόνο, καθώς το σιδηροδρομικό δίκτυο καλύπτει μεγάλες αποστάσεις εντός της χώρας. Λαμβάνοντας υπόψη έναν μέσο όρο 375 km/τόνο, η εκτιμώμενη συνολική μεταφορική απόδοση της VR Transpoint για το 2023 ανέρχεται σε περίπου 8,8 δισεκατομμύρια τονοχιλιόμετρα (tkm).
- Για να εκτιμηθούν τα τονοχιλιόμετρα που αντιστοιχούν στη **Mercitalia Rail** για το 2023, ακολουθήθηκε μια αναλογική προσέγγιση, βασισμένη σε οικονομικά στοιχεία. Συγκεκριμένα, έγινε η υπόθεση ότι η συμμετοχή της εταιρείας στο συνολικό μεταφορικό έργο του τομέα Logistics του Ομίλου FS είναι ανάλογη με το ποσοστό των εσόδων που της αναλογούν. Το Logistics Segment της FS Group κατέγραψε συνολικά 21,08 δισ. τονοχιλιόμετρα (TKM) για το 2023, με έσοδα ύψους 1.143 εκατ. ευρώ. Από αυτά, η Mercitalia Rail συνεισέφερε περίπου 480,06 εκατ. ευρώ, που αντιστοιχεί σε περίπου 42% των συνολικών εσόδων του τομέα. Αν υποθέσουμε ότι το ίδιο ποσοστό ισχύει και για το μεταφορικό έργο, τότε εκτιμάται ότι η Mercitalia Rail πραγματοποίησε περίπου 8,85 δισ. TKM κατά το 2023. Η εκτίμηση αυτή προσφέρει μια ενδεικτική εικόνα

της δραστηριότητας της εταιρείας στο ιταλικό σιδηροδρομικό δίκτυο και της συνολικής της συμβολής στις εμπορευματικές μεταφορές της χώρας.

Για τα υπόλοιπα δεδομένα τα οποία είναι το μήκος του δικτύου (χλμ), ο αριθμός του τροχαίου υλικού, ο αριθμός του προσωπικού και τα έσοδα (revenue), οι εταιρείες τα δημοσιεύουν στις ετήσιες εκθέσεις τους και έτσι υπήρχαν διαθέσιμες πληροφορίες, επιτρέποντας την άμεση καταγραφή τους χωρίς την ανάγκη για περαιτέρω εκτιμήσεις.

4.2 Σύνθεση Δεδομένων και Περιγραφή Εφαρμογής

4.2.1 Πίνακας Δεδομένων

Η καταγραφή και οργάνωση των δεδομένων αποτελεί κρίσιμο στάδιο της ανάλυσης, καθώς διαμορφώνει τη βάση για την εφαρμογή τόσο της μεθοδολογίας Data Envelopment Analysis (DEA) όσο και των Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks - ANN). Στην παρούσα μελέτη, αξιοποιούνται στοιχεία που αποτυπώνουν τις βασικές παραγωγικές, οικονομικές και περιβαλλοντικές παραμέτρους των επιλεγμένων σιδηροδρομικών οργανισμών. Τα δεδομένα προέρχονται κυρίως από επίσημες πηγές, οικονομικές εκθέσεις και εξειδικευμένες βάσεις δεδομένων, ενώ όπου κρίθηκε αναγκαίο πραγματοποιήθηκαν τεκμηριωμένες εκτιμήσεις που αναφέρονται στο προηγούμενο κεφάλαιο.

Η ομαδοποίησή τους στον παρακάτω πίνακα 2 επιτρέπει την άμεση σύγκριση των οργανισμών και διευκολύνει τη διενέργεια της ανάλυσης αποτελεσματικότητας με χρήση των δύο μεθοδολογιών. Το σύνολο των μεταβλητών εισόδου και εξόδου που παρουσιάζεται αποτελεί τη βάση για την περαιτέρω ποσοτική ανάλυση και την εξαγωγή συμπερασμάτων σχετικά με τη λειτουργική και περιβαλλοντική τους αποτελεσματικότητα.

Ακολουθεί ο πίνακας 3 με τα βασικά περιγραφικά στατιστικά των μεταβλητών εισόδου και εξόδου, ο οποίος αποτυπώνει τη μέση τιμή, την τυπική απόκλιση, καθώς και τα όρια των τιμών του δείγματος.

Πίνακας 4.2 Βάση δεδομένων

DMU	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Volume of ton-kilometers transported (B t-km)	Revenue (MEUR)
α/α	ΟΡΓΑΝΙΣΜΟΣ	ΧΩΡΑ	X1	X2	X3	X4	X5	X6	Y1	Y2
1	DB Cargo	Γερμανία	33,400.00	81,536	31,350	319.0	6,079	20.1	74.5	5,582
2	XeXafet (πρώην SNCF)	Γαλλία	28,000.00	7,677	5,000	105.0	630	1.99	16.7	714
3	Renfe Mercancías	Ισπανία	15,500.00	7,200	945	232.7	210.89	6.9	4.42	198.81
4	CD Cargo	Τσεχία	9,500.00	20,411	6,628	180.0	488.5	7.22	19.05	530
5	ÖBB Rail Cargo Group	Αυστρία	5,000.00	24,973	5,912	170.5	1,920	2.9	20.7	1,910
6	ZSSK Cargo	Σλοβακία	3,600.00	2,239	3,771	69.8	301.8	6.5	5.18	303.6
7	Hupac	Ελβετία	5,300.00	8,500	700	35.0	668.44	10.34	13.3	667
8	Green Cargo	Σουηδία	10,900.00	5,570	1,800	36.5	369.2	2.28	10.2	365.4
9	PKP Cargo	Πολωνία	18,800.00	49,757	14,062	349.2	897.82	18.49	22.3	940.3
10	VR Transpoint	Φινλανδία	9,200.00	8,360	2,800	82.7	349.5	5.5	8.8	339.9
11	Rail Cargo Hungaria Zrt	Ουγγαρία	7,606.00	2,600	1,460	42.3	475.7	13.96	5.85	473
12	Mercitalia Rail	Ιταλία	16,800.00	12,380	3,000	76.1	499.8	10.2	8.85	480
13	DB Cargo Netherlands	Ολλανδία	7,000.00	7,000	620	51.1	149	1.5	6.26	137

Πίνακας 4.3 Στατιστικά των Μεταβλητών

	μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Volume of ton-kilometers transported (B t-km)	Revenue (MEUR)
Mean	13,123.54	18,323.31	6,003.69	134.61	1,003.05	8.30	16.62	972.39
SE Mean	2,535.82	6,353.60	2,339.83	29.88	441.02	1.70	5.12	404.21
StDev	9,143.02	22,908.24	8,436.37	107.73	1,590.12	6.11	18.44	1,457.40
Sum	170,606.00	238,203.00	78,048.00	1,749.94	13,039.65	107.88	216.11	12,641.01
Minimum	3,600.00	2,239.00	620.00	35.00	149.00	1.50	4.42	137.00
Median	9,500.00	8,360.00	3,000.00	82.70	488.50	6.90	10.20	480.00
Maximum	33,400.00	81,536.00	31,350.00	349.24	6,079.00	20.10	74.50	5,582.00
Range	29,800.00	79,297.00	30,730.00	314.24	5,930.00	18.60	70.08	5,445.00

4.2.2 Εφαρμογή του Λογισμικού DEAP

Στην παρούσα ενότητα περιγράφεται η πειραματική διαδικασία που ακολουθήθηκε για την εκτίμηση της αποτελεσματικότητας των σιδηροδρομικών οργανισμών που εξετάζονται. Για την επίτευξη μιας συνολικής και αξιόπιστης ανάλυσης, εφαρμόστηκαν δύο διαφορετικές μεθοδολογικές προσεγγίσεις: το μοντέλο Data Envelopment Analysis (DEA) και τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks - ANN). Η χρήση δύο εναλλακτικών μεθόδων επιτρέπει την πολυδιάστατη αποτίμηση της αποτελεσματικότητας και τη σύγκριση των αποτελεσμάτων από διαφορετικές οπτικές. Αρχικά, πραγματοποιείται η εκτίμηση της τεχνικής αποτελεσματικότητας μέσω του λογισμικού DEAP, υλοποιώντας ένα μοντέλο DEA προσανατολισμένο στις εισροές και με την υπόθεση σταθερών αποδόσεων κλίμακας (CRS). Στη συνέχεια, εφαρμόζεται η προσέγγιση των Τεχνητών Νευρωνικών Δικτύων για την πρόβλεψη και ερμηνεία της απόδοσης των οργανισμών, παρέχοντας μία συμπληρωματική εκτίμηση της αποτελεσματικότητας μέσω μεθόδων τεχνητής νοημοσύνης.

Η πρώτη φάση της ανάλυσης αφορά την εφαρμογή της μεθόδου DEA, χρησιμοποιώντας το λογισμικό DEAP (Data Envelopment Analysis Program), που έχει αναπτυχθεί από τον Tim Coelli και θεωρείται ένα από τα πιο αξιόπιστα εργαλεία για την εκτίμηση αποδοτικότητας. Στην παρούσα εργασία, το μοντέλο DEA εφαρμόστηκε με προσανατολισμό στις εισροές (input-oriented) και με την παραδοχή σταθερών αποδόσεων κλίμακας (CRS).

Η επόμενη ενότητα περιγράφει βήμα-βήμα πώς δημιουργήθηκε το αρχείο δεδομένων, πώς έγινε η εκτέλεση του λογισμικού και με ποιον τρόπο αναγνώστηκαν και ερμηνεύτηκαν τα αποτελέσματα που προέκυψαν.

1. Δομή αρχείου εισόδου (.txt):

Για την εφαρμογή του μοντέλου DEA μέσω του λογισμικού DEAP, απαιτείται η δημιουργία ενός αρχείου κειμένου (.txt) με συγκεκριμένη δομή, το οποίο περιλαμβάνει τα δεδομένα για όλους τους οργανισμούς (DMUs). Αν και το DEAP υποστηρίζει την εισαγωγή περισσότερων της μίας εκροών, στο πλαίσιο της παρούσας εργασίας επιλέχθηκε η προσέγγιση της ξεχωριστής εκτέλεσης του μοντέλου για καθεμία από τις δύο βασικές εκροές:

- (α) Ton-kilometres (tkm) – μεταφορικό έργο
- (β) Revenue – έσοδα

Η απόφαση αυτή λήφθηκε προκειμένου να αναλυθεί η αποτελεσματικότητα των οργανισμών σε δύο διαφορετικές διαστάσεις απόδοσης: λειτουργική και οικονομική.

Η δομή του αρχείου εισόδου περιλαμβάνει 13 γραμμές εκ των οποίων η κάθε γραμμή ξεκινάει με την εκροή (ανάλογα με την εκτέλεση) και ακολουθούν οι έξι εισροές.

Στη δεύτερη εκτέλεση, η πρώτη στήλη αντικαθίσταται με τα έσοδα κάθε DMU, ενώ οι εισροές παραμένουν αμετάβλητες.

74.5	33400	81536	31350	319	6079	20.1
16.7	28000	7677	5000	105	630	1.99
4.42	15500	7200	945	232.7	210.89	6.9
19.05	9500	20411	6628	180	488.5	7.22
20.7	5000	24973	5912	170.5	1920	2.9
5.18	3600	2239	3771	69.8	301.8	6.5
13.3	5300	8500	700	35	668.44	10.34
10.2	10900	5570	1800	36.5	369.2	2.28
22.3	18800	49757	14062	349.24	897.82	18.49
8.8	9200	8360	2800	82.7	349.5	5.5
5.85	7606	2600	1460	42.3	475.7	13.96
8.85	16800	12380	3000	76.1	499.8	10.2
6.26	7000	7000	620	51.1	149	1.5

Εικόνα 4.1 Αρχείο κειμένου με εκροή τον όγκο μεταφοράς

5582	33400	81536	31350	319	6079	20.1
714	28000	7677	5000	105	630	1.99
198.81	15500	7200	945	232.7	210.89	6.9
530	9500	20411	6628	180	488.5	7.22
1910	5000	24973	5912	170.5	1920	2.9
303.6	3600	2239	3771	69.8	301.8	6.5
667	5300	8500	700	35	668.44	10.34
365.4	10900	5570	1800	36.5	369.2	2.28
940.3	18800	49757	14062	349.24	897.82	18.49
339.9	9200	8360	2800	82.7	349.5	5.5
473	7606	2600	1460	42.3	475.7	13.96
480	16800	12380	3000	76.1	499.8	10.2
137	7000	7000	620	51.1	149	1.5

Εικόνα 4.2 Αρχείο κειμένου με εκροή τον τζίρο

2. Δημιουργία αρχείου οδηγιών (instruction file)

Μετά τη δημιουργία του αρχείου δεδομένων, το επόμενο βήμα ήταν η κατασκευή του αρχείου οδηγιών, μέσω του οποίου το λογισμικό DEAP καθοδηγείται ως προς τις παραμέτρους της ανάλυσης. Το αρχείο αυτό είναι ένα απλό αρχείο κειμένου, με κατάληξη .txt και περιλαμβάνει διαδοχικά τις παρακάτω πληροφορίες:

Εικόνα 4.3 Αρχείο κειμένου οδηγιών (instruction file)

```

eg1-dta.txt      DATA FILE NAME
eg1-out.txt      OUTPUT FILE NAME
13              NUMBER OF FIRMS
1               NUMBER OF TIME PERIODS
1               NUMBER OF OUTPUTS
6               NUMBER OF INPUTS
0               0=INPUT AND 1=OUTPUT ORIENTATED
0               0=CRS AND 1=VRS
0               0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)

```

Οι επιλογές που χρησιμοποιήθηκαν ήταν οι εξής:

- Αρχείο δεδομένων: eg1-dta.txt
- Αρχείο αποτελεσμάτων: eg1-out.txt
- Αριθμός οργανισμών (DMUs): 13
- Περίοδος: 1 (στατική ανάλυση)
- Εκροές: 1
- Εισροές: 6
- Προσανατολισμός: 0 → input-oriented
- Αποδόσεις κλίμακας: 0 → CRS (Constant Returns to Scale)
- Είδος DEA: 0 → Standard DEA (multi-stage)

Το αρχείο αυτό αποθηκεύτηκε και εισήχθη στο πρόγραμμα μέσω της εντολής εκκίνησης του DEAP.exe, ορίζοντας ως input το path του instruction file.

3. Εκτέλεση του λογισμικού

Η εκτέλεση του λογισμικού DEAP έγινε μέσω του αρχείου deap.exe, το οποίο ενεργοποιήθηκε από το περιβάλλον των Windows. Κατά την εκκίνηση, το πρόγραμμα ζητά το όνομα του αρχείου οδηγιών (instruction file), στο οποίο έχουν καταχωρηθεί όλες οι βασικές παράμετροι για την ανάλυση. Με την επιλογή του συγκεκριμένου αρχείου, το DEAP εντοπίζει αυτόματα και το αντίστοιχο αρχείο δεδομένων, διαβάζει τις σχετικές πληροφορίες και εκτελεί το μοντέλο.

Στο πλαίσιο της παρούσας μελέτης, η ανάλυση πραγματοποιήθηκε δύο φορές, χρησιμοποιώντας την ίδια διάταξη εισροών και αλλάζοντας μόνο την εκροή κάθε φορά, όπως αναφέρεται και στο πρώτο βήμα.

Κατά την εκτέλεση, το DEAP εφαρμόζει το επιλεγμένο μοντέλο DEA (input-oriented, με CRS) και δημιουργεί αυτόματα ένα αρχείο εξόδου (.out) που περιλαμβάνει:

- Τιμές τεχνικής αποτελεσματικότητας (TE) για κάθε οργανισμό.
- Slacks για κάθε εισροή, υποδεικνύοντας δυναμικό περαιτέρω μείωσης.
- Προβλεπόμενες (target) τιμές εισροών και εκροών.
- Το σύνολο των μονάδων αναφοράς (peers) και των σχετικών βαρών (λ) που χρησιμοποιήθηκαν για την προβολή κάθε μη αποδοτικής μονάδας στο αποδοτικό σύνορο.

Με την ολοκλήρωση της επεξεργασίας, τα αποτελέσματα εξήχθησαν και αποθηκεύτηκαν σε αρχεία .txt, τα οποία αξιοποιήθηκαν στο επόμενο κεφάλαιο για τη συγκριτική ανάλυση της αποτελεσματικότητας των οργανισμών.

4.3 Παρουσίαση και Ανάλυση Αποτελεσμάτων

Στην παρούσα ενότητα παρουσιάζονται και αναλύονται τα αποτελέσματα που προέκυψαν από την εφαρμογή της μεθοδολογίας DEA στους επιλεγμένους σιδηροδρομικούς οργανισμούς. Η ανάλυση που ακολουθεί περιλαμβάνει συγκεντρωτικούς πίνακες αποτελεσμάτων, παρατηρήσεις ανά οργανισμό και συνοπτικά συμπεράσματα.

4.3.1 Τεχνική Αποδοτικότητα (Output: Ton-kilometres)

Η πρώτη εκτέλεση του μοντέλου DEA αφορούσε την τεχνική αποτελεσματικότητα των σιδηροδρομικών οργανισμών, λαμβάνοντας ως εκροή το μεταφορικό έργο σε ton-kilometres. Το μοντέλο ήταν input-oriented με υπόθεση σταθερών αποδόσεων (CRS).

Από τα αποτελέσματα προκύπτουν τα εξής:

- 9 από τους 13 οργανισμούς εμφάνισαν σκορ τεχνικής αποτελεσματικότητας ίσο με 1 (DB Cargo, Xexafret, CD Cargo, OBB Rail Cargo Group, ZSSK Cargo, Hupac, Green Cargo, Rail Cargo Hungaria, DB Cargo Netherlands) άρα θεωρούνται πλήρως αποδοτικοί.
- Οι υπόλοιποι είχαν τιμές <1 , με τις χαμηλότερες αποδοτικότητες στους:
 - Οργανισμός Mercitalia Rail: TE = 0.575
 - Οργανισμός Renfe Mercancías: TE = 0.579
 - Οργανισμός PKP Cargo: TE = 0.634
 - Οργανισμός VR Transpoint: TE = 0.802
- Ο μέσος όρος αποδοτικότητας ήταν 0.892, γεγονός που υποδηλώνει την ύπαρξη περιθωρίων βελτίωσης σε σημαντικό μέρος του δείγματος.

Παρατηρήσεις:

- Οι input slacks ήταν ιδιαίτερα υψηλοί σε μη αποδοτικές μονάδες, όπως ο DMU 9 (7605 στη 2η εισροή και 1482 στην 3η) και ο DMU 12 (π.χ. 778 στη 2η, 369 στην 3η).
- Τα reference sets και τα βάρη (λ) δείχνουν ποιοι DMUs χρησιμοποιήθηκαν ως πρότυπα για τις μη αποδοτικές μονάδες. Π.χ. ο DMU 3 είχε peers τους DMUs 13, 8 και 2 με βάρη 0.490, 0.129 και 0.002 αντίστοιχα.

Τα αποτελέσματα που προέκυψαν παρουσιάζονται στον παρακάτω πίνακα:

Πίνακας 4.4 Τεχνική Αποδοτικότητα με την μέθοδο DEA

DMU	ΟΡΓΑΝΙΣΜΟΣ	ΧΩΡΑ	Τεχνική Αποτελεσματικότητα
1	DB Cargo	Γερμανία	1
2	Hexafret (πρώην SNCF Fret)	Γαλλία	1
4	CD Cargo	Τσεχία	1
5	OBB Rail Cargo Group	Αυστρία	1
6	ZSSK Cargo	Σλοβακία	1
7	Hupac	Ελβετία	1
8	Green Cargo	Σουηδία	1
11	Rail Cargo Hungaria Zrt	Ουγγαρία	1
13	DB Cargo Netherlands	Ολλανδία	1
10	VR Transpoint	Φινλανδία	0.802
9	PKP Cargo	Πολωνία	0.634
3	Renfe Mercancías	Ισπανία	0.579
12	Mercitalia Rail	Ιταλία	0.575

4.3.2 Οικονομική Αποδοτικότητα (Output: Revenue)

Η δεύτερη εκτέλεση είχε ως εκροή τα έσοδα (revenue), εστιάζοντας στην οικονομική αποδοτικότητα των οργανισμών με βάση τα ίδια inputs.

Τα βασικά αποτελέσματα είναι τα εξής:

- 7 οργανισμοί παρουσίασαν $TE = 1.000$ (DMUs: 1, 2, 4, 5, 6, 7, 11).
- Οι υπόλοιποι είχαν σκορ μεταξύ 0.868 και 0.963, με τις χαμηλότερες τιμές:
 - DMU 13: $TE = 0.868$
 - DMU 12: $TE = 0.878$
 - DMU 3: $TE = 0.885$
- Ο μέσος όρος αποτελεσματικότητας ήταν 0.958, υψηλότερος από τον τεχνικό μέσο, γεγονός που ενδέχεται να αντανάκλα διαφοροποιήσεις στην τιμολογιακή πολιτική ή στην εμπορική απόδοση των οργανισμών.

Παρατηρήσεις:

- Παρατηρήθηκαν σημαντικά slacks, κυρίως στους DMUs 3, 8 και 9.
- Οι DMUs 2, 5 και 7 λειτούργησαν συχνά ως πρότυπα για άλλους οργανισμούς.
- Ο DMU 3, παρότι είχε πολύ χαμηλό efficiency score, προσεγγίζει την αποτελεσματικότητα μέσω συνδυασμού προβολών από τους αποδοτικούς DMUs 2 και 7.

Τα αποτελέσματα που προέκυψαν παρουσιάζονται στον παρακάτω πίνακα:

Πίνακας 4.5 Οικονομική Αποδοτικότητα με την μέθοδο DEA

DMU	ΟΡΓΑΝΙΣΜΟΣ	ΧΩΡΑ	Οικονομική Αποτελεσματικότητα
1	DB Cargo	Γερμανία	1
2	Hexafet (πρώην SNCF Fret)	Γαλλία	1
4	CD Cargo	Τσεχία	1
5	OBB Rail Cargo Group	Αυστρία	1
6	ZSSK Cargo	Σλοβακία	1
7	Hupac	Ελβετία	1
11	Rail Cargo Hungaria Zrt	Ουγγαρία	1
9	PKP Cargo	Πολωνία	0.963
8	Green Cargo	Σουηδία	0.958
10	VR Transpoint	Φινλανδία	0.899
3	Renfe Mercancías	Ισπανία	0.885
12	Mercitalia Rail	Ιταλία	0.878
13	DB Cargo Netherlands	Ολλανδία	0.868

4.3.3 Προτεινόμενες Βελτιώσεις από το μοντέλο DEA

Στο πλαίσιο της ανάλυσης DEA, για κάθε οργανισμό που αξιολογήθηκε ως μη αποδοτικός (με $TE < 1.000$), το λογισμικό DEAP υπολόγισε τα επίπεδα εισροών (input targets) τα οποία πρέπει να στοχεύσει ο κάθε ένας από αυτούς έτσι ώστε λειτουργεί ως πλήρως αποτελεσματικός, διατηρώντας σταθερό το επίπεδο εκροής. Οι τιμές αυτές προκύπτουν μέσω ακτινικής μείωσης των εισροών και επιπλέον προσαρμογών μέσω slacks.

4.3.4 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση της Τεχνικής και της Οικονομικής Αποδοτικότητας

Στον παρακάτω πίνακα 6 παρουσιάζονται οι αρχικές τιμές καθώς και οι τιμές στόχοι εισροών για κάθε οργανισμό που προέκυψε μη αποδοτικός ($TE < 1$) καθώς και η ποσοστιαία μεταβολή, όπως προκύπτουν από την ανάλυση DEA με εκροή το μεταφορικό έργο (ton-kilometres).

Αντίστοιχα, ο πίνακας 7 που ακολουθεί αφορά τα προτεινόμενα επίπεδα εισροών στους μη αποδοτικούς οργανισμούς ($TE < 1$) με στόχο τη βελτίωση της οικονομική αποτελεσματικότητας τους, με εκροή τα έσοδα (revenue).

Παρατηρείται ότι, σε αρκετές περιπτώσεις, οι μειώσεις στις εισροές που προτείνει το μοντέλο DEA για την επίτευξη πλήρους αποδοτικότητας είναι ιδιαίτερα σημαντικές, και σε ορισμένες περιπτώσεις μη ρεαλιστικές από ειδικά όσον αφορά κρίσιμες παραμέτρους όπως ο αριθμός προσωπικού και το μήκος του δικτύου.

Η έλλειψη ρεαλιστικών και εφαρμόσιμων λύσεων καθιστά αναγκαία τη χρήση προηγμένων μεθοδολογιών, όπως είναι τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks - ANN). Στην επόμενη ενότητα αναπτύσσεται η εφαρμογή της μεθόδου ANN, με στόχο την πρόβλεψη και την καλύτερη κατανόηση της αποδοτικής συμπεριφοράς των οργανισμών.

Πίνακας 4.6 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Τεχνικής Αποδοτικότητας DEAP

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Volume of ton-kilometers transported (B t-km)
α/α		ΟΡΓΑΝΙΣΜΟΣ		X1	X2	X3	X4	X5	X6	Y1
3	Πραγματικά Δεδομένα	Renfe Mercancías	Ισπανία	15500.0	7200.0	945.0	232.7	210.9	6.9	4.4
	Δεδομένα Στόχος			4896.7	4165.8	546.8	30.0	122.0	1.0	
	Ποσοστιαία μείωση			68.41%	42.14%	42.14%	87.12%	42.14%	85.01%	
9	Πραγματικά Δεδομένα	PKP Cargo	Πολωνία	18800.0	49757.0	14062.0	349.2	897.8	18.5	22.3
	Δεδομένα Στόχος			11924.1	23953.8	7436.0	209.0	569.5	8.3	
	Ποσοστιαία μείωση			36.57%	51.86%	47.12%	40.14%	36.57%	55.27%	
10	Πραγματικά Δεδομένα	VR Transpoint	Φινλανδία	9200.0	8360.0	2800.0	82.7	349.5	5.5	8.8
	Δεδομένα Στόχος			7378.9	6705.2	2158.8	52.6	280.3	2.5	
	Ποσοστιαία μείωση			19.79%	19.79%	22.90%	36.37%	19.79%	54.16%	
12	Πραγματικά Δεδομένα	Mercitalia Rail	Ιταλία	16800.0	12380.0	3000.0	76.1	499.8	10.2	8.9
	Δεδομένα Στόχος			9588.4	6345.0	1357.1	43.8	287.6	2.0	
	Ποσοστιαία μείωση			42.93%	48.75%	54.76%	42.46%	42.46%	80.19%	

Πίνακας 4.7 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Οικονομικής Αποδοτικότητας DEAP

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Revenue (MEUR)
α/α		ΟΡΓΑΝΙΣΜΟΣ		X1	X2	X3	X4	X5	X6	Y2
3	Πραγματικά Δεδομένα	Renfe Mercancías	Ισπανία	15500.0	7200.0	945.0	232.7	210.9	6.9	198.8
	Δεδομένα Στόχος			4876.0	2323.6	836.2	20.4	186.6	1.7	
	Ποσοστιαία μείωση			68.54%	67.73%	11.51%	91.23%	11.51%	74.75%	
8	Πραγματικά Δεδομένα	Green Cargo	Σουηδία	10900.0	5570.0	1800.0	36.5	369.2	2.3	365.4
	Δεδομένα Στόχος			5495.1	4490.2	1336.7	35.0	353.7	2.2	
	Ποσοστιαία μείωση			49.59%	19.39%	25.74%	4.21%	4.21%	4.21%	
9	Πραγματικά Δεδομένα	PKP Cargo	Πολωνία	18800.0	49757.0	14062.0	349.2	897.8	18.5	940.3
	Δεδομένα Στόχος			18099.6	34588.8	11437.2	308.1	864.4	12.2	
	Ποσοστιαία μείωση			3.73%	30.48%	18.67%	11.78%	3.73%	34.15%	
10	Πραγματικά Δεδομένα	VR Transpoint	Φινλανδία	9200.0	8360.0	2800.0	82.7	349.5	5.5	339.9
	Δεδομένα Στόχος			8268.4	6273.5	2516.5	61.7	314.1	1.8	
	Ποσοστιαία μείωση			10.13%	24.96%	10.13%	25.43%	10.13%	67.65%	
12	Πραγματικά Δεδομένα	Mercitalia Rail	Ιταλία	16800.0	12380.0	3000.0	76.1	499.8	10.2	480.0
	Δεδομένα Στόχος			14751.5	5420.2	2634.2	59.1	438.9	2.7	
	Ποσοστιαία μείωση			12.19%	56.22%	12.19%	22.34%	12.19%	73.14%	
13	Πραγματικά Δεδομένα	DB Cargo Netherlands	Ολλανδία	7000.0	7000.0	620.0	51.1	149.0	1.5	137.0
	Δεδομένα Στόχος			3160.6	1613.9	538.3	13.5	129.4	1.3	
	Ποσοστιαία μείωση			54.85%	76.94%	13.18%	73.67%	13.18%	14.60%	

4.4 Τεχνητά Νευρωνικά Δίκτυα (ANN)

Η προσέγγιση των Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks - ANN) εφαρμόζεται στην παρούσα μελέτη ως ένα εργαλείο πρόβλεψης και βελτιστοποίησης της αποδοτικότητας, με στόχο την υπέρβαση ορισμένων εγγενών περιορισμών του μοντέλου DEA. Παρότι το DEA αποτελεί μια ισχυρή και διαφανή μεθοδολογία για την εκτίμηση της σχετικής αποτελεσματικότητας, συχνά προτείνει λύσεις που, στην πράξη, μπορεί να είναι μη ρεαλιστικές ή δύσκολα εφαρμόσιμες, όπως δραστικές μειώσεις σε κρίσιμες εισροές.

Τα ANN, ως μαθησιακά μοντέλα που εμπνέονται από τη λειτουργία του ανθρώπινου εγκεφάλου, έχουν τη δυνατότητα να εντοπίζουν πολύπλοκα και μη γραμμικά πρότυπα στα δεδομένα. Στο πλαίσιο αυτής της μελέτης, τα ANN αξιοποιούνται για να εκπαιδευτούν πάνω σε εκτεταμένα δεδομένα DEA (σενάρια εισροών-εκροών και τιμές αποτελεσματικότητας θ) και να αποκτήσουν τη δυνατότητα πρόβλεψης αποδοτικών εισροών για κάθε οργανισμό.

Η υλοποίηση της προσέγγισης ANN περιλαμβάνει τέσσερα βασικά στάδια:

1. Δημιουργία ενός συνόλου δεδομένων μέσω συστηματικής μεταβολής των εισροών για κάθε οργανισμό.
2. Εκτίμηση της αποδοτικότητας κάθε εγγραφής με το μοντέλο DEA και προσθήκη της αποδοτικότητας (θ) ως ετικέτα.
3. Εκπαίδευση του ANN, ώστε να προβλέπει αποδοτικούς συνδυασμούς εισροών με βάση τις μεταβλητές Y_2 και θ .
4. Πρόβλεψη και επαλήθευση, όπου οι λύσεις που προτείνει το ANN αξιολογούνται εκ νέου με DEA, και επιλέγεται η βέλτιστη ως προς τη βελτίωση της αποδοτικότητας.

Στις ενότητες που ακολουθούν περιγράφονται αναλυτικά τα βήματα αυτά και παρουσιάζονται τα αντίστοιχα αποτελέσματα.

4.4.1 Δημιουργία Συνόλου Δεδομένων

Το πρώτο στάδιο για την υλοποίηση του μοντέλου Τεχνητού Νευρωνικού Δικτύου (ANN) αφορά τη δημιουργία ενός επαρκούς και κατάλληλα δομημένου συνόλου δεδομένων για την εκπαίδευση του μοντέλου. Το σύνολο αυτό δημιουργήθηκε με βάση τα πραγματικά δεδομένα των 13 σιδηροδρομικών οργανισμών και περιλαμβάνει συνδυασμούς εισροών που παραχθήκαν συστηματικά και όχι με τυχαίο τρόπο.

Συγκεκριμένα:

- Για κάθε οργανισμό, οι μεταβλητές εισόδου X_2 έως X_6 μεταβλήθηκαν συστηματικά από το 50% έως και το 100% των αρχικών τους τιμών, με βήματα 5%.
- Η μεταβλητή X_1 (μήκος δικτύου) παρέμεινε σταθερή, καθώς θεωρείται εξωγενής και μη επηρεάσιμη μεταβλητή στο πλαίσιο της ανάλυσης.
- Η διαδικασία εφαρμόστηκε για δύο διαφορετικές εκροές (Y):
 - (α) Y_2 = έσοδα (revenue) για την εκτίμηση οικονομικής αποτελεσματικότητας, και
 - (β) Y_1 = ton-kilometres για την εκτίμηση τεχνικής αποτελεσματικότητας.

Για κάθε περίπτωση, δημιουργήθηκαν εναλλακτικοί συνδυασμοί εισροών μέσω ολικών συνδυασμών (cartesian product) των επιπέδων μεταβολής, οδηγώντας αρχικά σε 7.776 συνδυασμούς ανά οργανισμό.

Η παραπάνω διαδικασία οδήγησε στη δημιουργία ενός αρχικού dataset 100.008 εγγραφών (13 οργανισμοί × 7776 σενάρια). Ωστόσο, για λόγους υπολογιστικής αποδοτικότητας, εφαρμόστηκε δειγματοληψία σε κάθε 20η εγγραφή, με αποτέλεσμα ένα τελικό σύνολο 5.049 εγγραφών για κάθε εκδοχή της ανάλυσης.

Η επιλογή συστηματικής και όχι τυχαίας μεταβολής έγινε ώστε:

- να διασφαλιστεί η ρεαλιστικότητα των παραγόμενων συνδυασμών εισροών,
- να επιτευχθεί ομοιομορφία στην κατανομή των δεδομένων ανά DMU,
- και να εξασφαλιστεί η αναπαραγωγιμότητα της μεθοδολογίας.

```

1 import pandas as pd
2 import numpy as np
3 from itertools import product
4
5 # === STEP 1: Define Original Firm Data ===
6 firms = pd.DataFrame({
7     "DMU": ["DB Cargo", "Xefafet", "Renfe", "CD Cargo", "08B", "ZSSK", "Hupac", "Green Cargo",
8           "PKP Cargo", "VR", "RCH", "Mercitalia", "DB NL"],
9     "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18000, 7606, 7606, 16000, 7000],
10    "X2": [81536, 7677, 7200, 20411, 24973, 4289, 8500, 5170, 49757, 8360, 6000, 12380, 7000],
11    "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460, 3000, 620],
12    "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7, 42.3, 76.1, 51.1],
13    "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82, 349.5, 475.7, 499.8, 149],
14    "X6": [20.1, 1.99, 6.9, 7.22, 2.9, 6.5, 10.34, 2.28, 18.49, 5.5, 13.96, 10.2, 1.5],
15    "Y2": [5582, 714, 198.81, 530, 1910, 303.6, 667, 365.4, 940.3, 339.9, 473, 480, 137]
16 })
17
18 # === STEP 2: Define Scaling Strategy ===
19 scales = np.arange(0.5, 1.01, 0.1) # 50% to 100% in 10% steps (6 values)
20 scale_combinations = list(product(scales, repeat=5)) # for X2-X6 → 6^5 = 7776 rows per firm
21
22 # === STEP 3: Generate Combinations for Each Firm ===
23 all_rows = []
24
25 for _, row in firms.iterrows():
26     base_inputs = np.array([row["X2"], row["X3"], row["X4"], row["X5"], row["X6"]])
27     for scale_set in scale_combinations:
28         scaled_inputs = base_inputs * np.array(scale_set)
29         all_rows.append({
30             "DMU": row["DMU"],
31             "X1": row["X1"],
32             "Y2": row["Y2"],
33             "X2": scaled_inputs[0],
34             "X3": scaled_inputs[1],
35             "X4": scaled_inputs[2],
36             "X5": scaled_inputs[3],
37             "X6": scaled_inputs[4],
38         })
39
40 df_full = pd.DataFrame(all_rows)
41
42 # === STEP 4: Sample Every 20th Row to Reduce Size ===
43 df_sampled_half = df_full.groupby(by="DMU", group_keys=False).apply(lambda x: x.iloc[::20]).reset_index(drop=True)
44
45 # === STEP 5: Export Final Dataset ===
46 df_sampled_half.to_excel(excel_writer="systematic_dataset_all_firms_sampled_half.xlsx", index=False)
47 print("✅ Dataset created and saved.")

```

Εικόνα 4.4 systematic_dataset_all_firms_sampled_half.xlsx

Η παραγωγή των συνόλων δεδομένων πραγματοποιήθηκε με χρήση ειδικά ανεπτυγμένων Python scripts:

- Για την εκδοχή με Revenue (Y2): generate_systematic_sampled_half.py
- Για την εκδοχή με Ton-kilometres (Y1): generate_systematic_sampled_half_BTK.py

Τα αρχεία .xlsx που προέκυψαν και χρησιμοποιήθηκαν στα επόμενα στάδια ήταν:

- systematic_dataset_all_firms_sampled_half.xlsx (για Revenue)
- systematic_dataset_all_firms_sampled_half_BTK.xlsx (για Ton-kilometres)

Και τα δύο αποτέλεσαν τη βάση για την επόμενη φάση: υπολογισμός της αποτελεσματικότητας DEA (θ) για κάθε γραμμή του dataset.

4.4.2 Υπολογισμός της DEA Αποδοτικότητας (θ)

Αφού δημιουργήθηκε το πλήρες σύνολο δεδομένων με συστηματικά παραγόμενους συνδυασμούς εισροών, το επόμενο βήμα περιλάμβανε τον υπολογισμό της τεχνικής αποτελεσματικότητας (θ) για κάθε εγγραφή, μέσω εφαρμογής της μεθοδολογίας Data Envelopment Analysis (DEA).

Η συγκεκριμένη φάση ήταν απαραίτητη προκειμένου να αξιοποιηθεί η αποτελεσματικότητα κάθε παρατήρησης ως ετικέτα (label) για την εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου. Η αποτελεσματικότητα θ προσδιορίζει τον βαθμό στον οποίο ένας συνδυασμός εισροών μπορεί να θεωρηθεί αποδοτικός, σε σχέση με τα υπόλοιπα σημεία του dataset.

DEA Μοντέλο – Προσέγγιση και Ρυθμίσεις

Η ανάλυση πραγματοποιήθηκε με χρήση του κλασικού μοντέλου:

- Input-oriented, το οποίο στοχεύει στη μείωση των εισροών, διατηρώντας σταθερή την εκροή.
- Υπό την υπόθεση σταθερών αποδόσεων ως προς την κλίμακα (CRS - Constant Returns to Scale).

Η DEA μοντελοποίηση επιλύει, για κάθε παρατήρηση (συνδυασμό εισροών), ένα πρόβλημα γραμμικού προγραμματισμού, ελαχιστοποιώντας την τιμή θ υπό συγκεκριμένους περιορισμούς, ώστε η παρατήρηση να συγκριθεί με το αποδοτικό μέτωπο (efficient frontier) το οποίο είναι της μορφής :

$$\min \theta$$

$$\theta x_i \leq \sum_{j=1}^n \lambda_j x_j \geq 0$$

για όλες τις μεταβλητές εισόδου

$$y_i \geq \sum_{j=1}^n \lambda_j y_j$$

για όλες τις μεταβλητές εξόδου

$$\lambda_j \geq 0 \quad \forall j$$

Η διαδικασία επαναλήφθηκε ανεξάρτητα για τις δύο εκδοχές εκροής:

- (α) $Y2$ = Έσοδα (Revenue) για οικονομική αποτελεσματικότητα.
- (β) $Y1$ = Ton-kilometres για τεχνική αποτελεσματικότητα.

Η εφαρμογή πραγματοποιήθηκε με τη χρήση των ακόλουθων Python scripts:

Για την εκδοχή της οικονομικής αποτελεσματικότητας με εκροή τα έσοδα ($Y2$), χρησιμοποιήθηκε το script **calculate_theta_sampled_half.py**, το οποίο υπολόγισε την τιμή θ για κάθε παρατήρηση και παράγαγε το αρχείο εξόδου **systematic_dataset_all_firms_sampled_half_with_theta.xlsx**.

Αντίστοιχα, για την εκδοχή της τεχνικής αποτελεσματικότητας με εκροή τον όγκο μεταφοράς σε ton-kilometres ($Y1$), εφαρμόστηκε το script **calculate_theta_sampled_half_BTK.py**, με αποτέλεσμα το αρχείο **systematic_dataset_all_firms_sampled_half_BTK_with_theta.xlsx**.

Για κάθε συνδυασμό εισροών, το αποτέλεσμα του μοντέλου DEA ήταν μια τιμή $\theta \in (0, 1]$, η οποία εκφράζει πόσο «μακριά» βρίσκεται το συγκεκριμένο σενάριο από το αποδοτικό μέτωπο. Τιμή $\theta = 1$ σημαίνει πλήρη αποδοτικότητα, ενώ $\theta < 1$ υποδηλώνει ότι υπάρχει περιθώριο βελτίωσης μέσω μείωσης των εισροών.

Οι τιμές αυτές ενσωματώθηκαν ως νέα στήλη στο dataset και αποτέλεσαν το βασικό στοιχείο για την εκπαίδευση του ANN μοντέλου που ακολουθεί.

```

1
2 import pandas as pd
3 import numpy as np
4 from scipy.optimize import linprog
5 from tqdm import tqdm
6
7 # Load the reduced dataset
8 df = pd.read_excel("systematic_dataset_all_firms_sampled_half.xlsx")
9
10 # DEA Efficiency Function (Input-oriented, CRS)
11 def dea_efficiency(inputs, outputs, eval_index):
12     n, m_inputs = inputs.shape
13     c = np.zeros(n + 1)
14     c[0] = 1
15     A_ub, b_ub = [], []
16     for j in range(m_inputs):
17         row = np.zeros(n + 1)
18         row[0] = -inputs[eval_index, j]
19         row[1:] = inputs[:, j]
20         A_ub.append(row)
21         b_ub.append(0)
22     A_eq, b_eq = [], []
23     for j in range(outputs.shape[1]):
24         row = np.zeros(n + 1)
25         row[1:] = -outputs[:, j]
26         A_eq.append(row)
27         b_eq.append(-outputs[eval_index, j])
28     bounds = [(0, 1)] + [(0, None)] * n
29     res = linprog(c, A_ub=A_ub, b_ub=b_ub, A_eq=A_eq, b_eq=b_eq, bounds=bounds, method="highs")
30     return res.x[0] if res.success else np.nan
31
32 # Prepare inputs and outputs
33 inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
34 outputs = df[["Y2"]].values
35
36 # Calculate theta with progress bar
37 theta_scores = []
38 for i in tqdm(range(len(df)), desc="Calculating DEA Efficiency (Half Sample)"):
39     theta = dea_efficiency(inputs, outputs, i)
40     theta_scores.append(theta)
41
42 # Save results
43 df["Theta"] = theta_scores
44 df.to_excel("systematic_dataset_all_firms_sampled_half_with_theta.xlsx", index=False)
45 print("DEA efficiency calculation complete and file saved.")

```

Εικόνα 4.5 calculate_theta_sampled_half.py

4.4.3 Εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου (ANN)

Μετά τον υπολογισμό της αποτελεσματικότητας (θ) για κάθε συνδυασμό εισροών, το επόμενο βήμα περιλάμβανε την εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου (ANN). Το μοντέλο ANN αναπτύχθηκε με στόχο να μάθει να προβλέπει αποδοτικούς συνδυασμούς εισροών, αξιοποιώντας ως δεδομένα εισόδου τόσο τα χαρακτηριστικά του κάθε οργανισμού όσο και την πληροφορία σχετικά με την αποτελεσματικότητά του.

Δομή του ANN

Η διαμόρφωση της εισόδου και της εξόδου του ANN βασίστηκε στη λογική της προσέγγισης DEA και στις ανάγκες της πρόβλεψης:

- Είσοδοι (Input Features):
 - X_1 : Μήκος δικτύου (σταθερό)
 - Y (Y_2 ή Y_1): Εκροή (Revenue ή Ton-kilometres)
 - θ : Βαθμός αποδοτικότητας από το DEA
- Έξοδοι (Target Outputs):
 - X_2, X_3, X_4, X_5, X_6 : Προβλεπόμενες αποδοτικές τιμές των εισροών

Η επιλογή αυτής της δομής επιτρέπει στο ANN να προβλέπει προσαρμοσμένα επίπεδα εισροών, λαμβάνοντας υπόψη τόσο τα χαρακτηριστικά λειτουργίας κάθε οργανισμού όσο και το επίπεδο αποτελεσματικότητας στο οποίο βρίσκεται.

Διαδικασία Εκπαίδευσης

Η εκπαίδευση του ANN πραγματοποιήθηκε ξεχωριστά για κάθε εκδοχή εκροής:

- Για το Revenue (Y_2): Το script που χρησιμοποιήθηκε είναι το **train_ann_with_theta.py** και παράγει το μοντέλο **ann_theta_optimizer_model.h5**
- Για το Ton-kilometres (Y_1): Το script που χρησιμοποιήθηκε είναι το **train_ann_with_theta_BTK.py** και παράγει το μοντέλο **ann_theta_optimizer_model_BTK.h5**

Τα datasets που χρησιμοποιήθηκαν περιλάμβαναν τις τιμές X_1, Y, θ ως εισόδους και τις αντίστοιχες αποδοτικές τιμές X_2-X_6 ως στόχους.

Η διαδικασία εκπαίδευσης περιλάμβανε τα εξής βασικά στάδια:

- Κανονικοποίηση των εισόδων και εξόδων.
- Εκπαίδευση σε πολλαπλά epochs για την επίτευξη σταθερής ελαχιστοποίησης του σφάλματος.
- Χρήση λειτουργίας απώλειας τύπου Mean Squared Error (MSE).
- Επιλογή της καλύτερης εκδοχής του μοντέλου βάσει της απόδοσής του στα δεδομένα επικύρωσης (validation set).

Η εκπαίδευση του ANN επέτρεψε τη δημιουργία ενός προγνωστικού μοντέλου, το οποίο μπορεί να προτείνει ρεαλιστικές και αποδοτικές προσαρμογές εισροών, υπερβαίνοντας τους περιορισμούς της παραδοσιακής DEA ανάλυσης.


```

1
2 import pandas as pd
3 import numpy as np
4 from sklearn.model_selection import train_test_split
5 from tensorflow.keras.models import Sequential
6 from tensorflow.keras.layers import Dense
7 from tensorflow.keras.optimizers import Adam
8
9 # Load the DEA-evaluated dataset
10 df = pd.read_excel("systematic_dataset_all_firms_sampled_half_with_theta.xlsx")
11
12 # Drop any rows with missing theta (NaN)
13 df = df.dropna(subset=["Theta"])
14
15 # Define input (X1, Y2, Theta) and output (X2-X6)
16 X = df[["X1", "Y2", "Theta"]].values
17 y = df[["X2", "X3", "X4", "X5", "X6"]].values
18
19 # Split the dataset
20 X_train, X_test, y_train, y_test = train_test_split(*arrays: X, y, test_size=0.2, random_state=42)
21
22 # Build the ANN model
23 model = Sequential()
24 model.add(Dense(64, input_dim=3, activation='relu'))
25 model.add(Dense(64, activation='relu'))
26 model.add(Dense(5)) # 5 outputs for X2-X6
27
28 # Compile the model
29 model.compile(optimizer=Adam(learning_rate=0.001), loss='mse')
30
31 # Train the model
32 model.fit(X_train, y_train, epochs=500, batch_size=16, validation_split=0.1)
33
34 # Save the trained model
35 model.save("ann_theta_optimizer_model.h5")
36
37 # Evaluate the model
38 loss = model.evaluate(X_test, y_test)
39 print(f"✅ Model trained. Test MSE Loss: {loss}")

```

Εικόνα 4.6 train_ann_with_theta.py

4.4.4 Βελτιστοποίηση και Επαλήθευση μέσω ANN

Αφού ολοκληρώθηκε η εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου (ANN) για κάθε εκδοχή εκροής, το τελευταίο στάδιο της διαδικασίας περιλάμβανε τη χρήση του εκπαιδευμένου μοντέλου για τη δημιουργία νέων προτεινόμενων εισροών και την επαλήθευση της αποτελεσματικότητάς τους μέσω νέας ανάλυσης DEA.

Για κάθε οργανισμό που δεν είχε πλήρη αποτελεσματικότητα ($\theta < 1.000$), εφαρμόστηκε η εξής διαδικασία:

1. Δημιουργία πολλαπλών σεναρίων:
 - Για κάθε μη αποδοτικό οργανισμό, το ANN δημιούργησε 10 διαφορετικές προβλέψεις εισροών.
 - Τα σενάρια παραγόντουσαν μεταβάλλοντας το θ από την αρχική του τιμή έως το 1.0.
2. Περιορισμοί στις προτεινόμενες εισροές:
 - Οι προτεινόμενες εισροές περιορίστηκαν ώστε να κινούνται μεταξύ 50% και 100% των αρχικών τιμών κάθε οργανισμού, ώστε να διατηρηθεί ρεαλιστικότητα.

3. Αξιολόγηση μέσω DEA:

- Κάθε πρόβλεψη εισροών αξιολογήθηκε εκ νέου με DEA για να υπολογιστεί η νέα τιμή αποτελεσματικότητας θ .

4. Επιλογή βέλτιστης λύσης:

- Από τις 10 προτάσεις, επιλέχθηκε εκείνη που απέδωσε την υψηλότερη τιμή θ .

Για τους οργανισμούς που ήδη είχαν πλήρη αποδοτικότητα ($\theta = 1.000$), δεν έγινε καμία προσαρμογή.

Η διαδικασία βελτιστοποίησης και επαλήθευσης πραγματοποιήθηκε με τα παρακάτω scripts:

- Για το Revenue (Y2): Το script `eval_ann_with_theta_model_REVENUE_FINAL.py` και τελικά δημιουργήθηκε το αρχείο `ann_theta_prediction_results_REVENUE_FINAL.xlsx`
- Για το Ton-kilometres (Y1): Το script `eval_ann_with_theta_model_BTK_FINAL.py` και τελικά δημιουργήθηκε το αρχείο `ann_theta_prediction_results_BTK_FINAL.xlsx`

Τα αρχεία εξόδου περιέχουν, για κάθε οργανισμό:

- Τις προτεινόμενες νέες τιμές εισροών (X2–X6),
- Τη νέα τιμή αποτελεσματικότητας θ μετά από την επανεκτίμηση DEA,
- Και τη σύγκριση με την αρχική αποτελεσματικότητα.

Η εφαρμογή του ANN μοντέλου σε συνδυασμό με την επαλήθευση μέσω DEA επέτρεψε την πρόβλεψη πιο ρεαλιστικών, βελτιστοποιημένων εισροών, με στόχο τη βελτίωση της αποτελεσματικότητας των σιδηροδρομικών οργανισμών, ξεπερνώντας τους περιορισμούς της παραδοσιακής DEA προσέγγισης.

Στους παρακάτω πίνακες παρουσιάζονται τα αποτελέσματα της εφαρμογής του εκπαιδευμένου ANN μοντέλου και η επαλήθευσή τους μέσω νέου υπολογισμού αποτελεσματικότητας με τη μέθοδο DEA. Για κάθε οργανισμό εμφανίζονται:

- Οι αρχικές τιμές εισροών (X1–X6),
- Η αρχική αποτελεσματικότητα (θ) από το DEA (στήλη `Theta_DEAP`),
- Οι προτεινόμενες εισροές από το ANN (στήλες `X2_best` έως `X6_best`),
- Η νέα αποτελεσματικότητα που επιτεύχθηκε με τις προτεινόμενες εισροές (`Theta_ANN_best`),
- Και η ποσοστιαία βελτίωση (%) της αποτελεσματικότητας (`Improvement %`), σε σχέση με την αρχική κατάσταση.

Η μέθοδος επέτρεψε σε αρκετούς μη αποδοτικούς οργανισμούς να βελτιώσουν σημαντικά το σκορ αποτελεσματικότητας, περιορίζοντας τις εισροές τους σε πιο ρεαλιστικά επίπεδα, σε σχέση με τις συχνά υπερβολικές μειώσεις που προτείνονται από τη DEA.

Οι πίνακες που ακολουθούν αντιστοιχούν στην εκδοχή με εκροή τον όγκο μεταφοράς (Y1 – ton-km) και αντίστοιχα τα έσοδα (Y2 – Revenue) και εκροή. Στους παρακάτω πίνακες(8,9,10) παρουσιάζονται οι αρχικές τιμές καθώς και οι τιμές στόχοι εισροών για κάθε οργανισμό που προέκυψε μη αποδοτικός ($TE < 1$) καθώς και η ποσοστιαία μεταβολή, όπως προκύπτουν από την ανάλυση του μοντέλου ANN με εκροή το

μεταφορικό έργο (ton-kilometres). Αντίστοιχα, ο πίνακας που ακολουθεί αφορά τα προτεινόμενα επίπεδα εισροών στους μη αποδοτικούς οργανισμούς ($TE < 1$) του μοντέλου ANN με στόχο τη βελτίωση της οικονομική αποτελεσματικότητας τους, με εκροή τα έσοδα (revenue).

```

1
2 import pandas as pd
3 import numpy as np
4 from tensorflow.keras.models import load_model
5 from scipy.optimize import linprog
6
7 # DEA Efficiency Function
8 def dea_efficiency(inputs, outputs, eval_index):
9     n, m_inputs = inputs.shape
10    c = np.zeros(n + 1)
11    c[0] = 1
12    A_ub, b_ub = [], []
13    for j in range(m_inputs):
14        row = np.zeros(n + 1)
15        row[0] = -inputs[eval_index, j]
16        row[1:] = inputs[:, j]
17        A_ub.append(row)
18        b_ub.append(0)
19    A_eq, b_eq = [], []
20    for j in range(outputs.shape[1]):
21        row = np.zeros(n + 1)
22        row[1:] = -outputs[:, j]
23        A_eq.append(row)
24        b_eq.append(-outputs[eval_index, j])
25    bounds = [(0, 1)] + [(0, None)] * n
26    res = linprog(c, A_ub, b_ub, A_eq, b_eq, bounds, method="highs")
27    return res.x[0] if res.success else np.nan
28
29 # STEP 1: Load Correct Firm Data (Revenue version with user provided corrections)
30 df = pd.DataFrame({
31     "DMU": [
32         "DB Cargo", "Xefafet", "Renfe", "CD Cargo", "08B", "ZSSK", "Hupac",
33         "Green Cargo", "PKP Cargo", "VR", "Rail Cargo Hungaria", "Mercitalia", "DB NL"
34     ],
35     "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18800, 9200, 7606, 16800, 7000],
36     "X2": [81536, 7677, 7208, 20411, 24973, 2239, 8500, 5570, 49757, 8360, 2600, 12380, 7000],
37     "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460, 3000, 620],
38     "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7, 42.3, 76.1, 51.1],
39     "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82, 349.5, 475.7, 499.8, 149],
40     "X6": [28.1, 1.99, 6.9, 7.22, 2.9, 6.5, 18.34, 2.28, 18.49, 5.5, 13.96, 18.2, 1.5],
41     "Y2": [5582, 714, 198.81, 530, 1910, 303.6, 667, 365.4, 940.3, 339.9, 473, 480, 137],
42     "Theta_DEAP": [1.000, 1.000, 0.885, 1.000, 1.000, 1.000, 1.000, 0.958, 0.963, 0.899, 1.000, 0.878, 0.868]
43 })
44
45 # STEP 2: Load Trained Revenue Model
46 model = load_model("ann_theta_optimizer_model.h5", compile=False)
47
48 X1_vals = df["X1"].values
49 Y2_vals = df["Y2"].values
50 theta_deap = df["Theta_DEAP"].values
51 original_inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
52 outputs = df[["Y2"]].values
53
54 # STEP 3: Predict Optimized Inputs and Evaluate Multi-Try
55 final_inputs = original_inputs.copy()
56 theta_ann_best = []
57
58 for i in range(len(df)):
59     if theta_deap[i] == 1.0:
60         theta_ann_best.append(1.0)
61         continue
62
63     best_theta = -1
64     best_input = None
65     theta_range = np.linspace(theta_deap[i], 1.0, num=10)
66
67     for theta_target in theta_range:
68         input_ann = np.array([X1_vals[i], Y2_vals[i], theta_target])
69         pred = model.predict(input_ann, verbose=0)[0]
70         base = original_inputs[i, 1:]
71         clipped = np.clip(pred, base * 0.5, base)
72         test_input = final_inputs.copy()
73         test_input[i] = np.concatenate([X1_vals[i], clipped])
74         score = dea_efficiency(test_input, outputs, i)
75         if not np.isnan(score) and score > best_theta:
76             best_theta = score
77             best_input = clipped
78
79     if best_input is not None:
80         final_inputs[i] = np.concatenate([X1_vals[i], best_input])
81     theta_ann_best.append(best_theta)
82
83 # STEP 4: Save Final Output
84 df["Theta_ANN_best"] = theta_ann_best
85 for j, col in enumerate(["X2_best", "X3_best", "X4_best", "X5_best", "X6_best"]):
86     df[col] = final_inputs[:, j + 1]
87 df["Improvement_%"] = ((df["Theta_ANN_best"] - df["Theta_DEAP"]) / df["Theta_DEAP"]) * 100
88
89 df.to_excel("excelwriter/ann_theta_prediction_results_REVENUE_FINAL.xlsx", index=False)
90 print("✅ ANN optimization + DEA evaluation (Revenue FINAL) completed and saved.")

```

Εικόνα 4.7 eval_ann_with_theta_model_REVENUE_FINAL.py

Πίνακας 4.8 Αποτελέσματα της εφαρμογής του μοντέλου ANN για Τεχνική Αποδοτικότητα

DMU	ΟΡΓΑΝΙΣΜΟΣ	X1	X2	X3	X4	X5	X6	Y1	Theta_DEAP	Theta_ANN_best	X1	X2_best	X3_best	X4_best	X5_best	X6_best	Improvement_%
1	DB Cargo	33400	81536	31350	319	6079	20.1	74.5	1	1	33400	81536	31350	319	6079	20.1	0
2	Xefafet	28000	7677	5000	105	630	1.99	16.7	1	1	28000	7677	5000	105	630	1.99	0
3	Renfe	15500	7200	945	232.7	210.89	6.9	4.42	0.579	0.808	15500	5394.98	945.00	116.35	145.56	3.45	39.50
4	CD Cargo	9500	20411	6628	180	488.5	7.22	19.05	1	1	9500	20411	6628	180	488.5	7.22	0
5	OBB	5000	24973	5912	170.5	1920	2.9	20.7	1	1	5000	24973	5912	170.5	1920	2.9	0
6	ZSSK	3600	2239	3771	69.8	301.8	6.5	5.18	1	1	3600	2239	3771	69.8	301.8	6.5	0
7	Hupac	5300	8500	700	35	668.44	10.34	13.3	1	1	5300	8500	700	35	668.44	10.34	0
8	Green Cargo	10900	5570	1800	36.5	369.2	2.28	10.2	1	1	10900	5570	1800	36.5	369.2	2.28	0
9	PKP Cargo	18800	49757	14062	349.2	897.82	18.49	22.3	0.634	1	18800	40512.33	10850.67	287.37	449.07	9.245	57.73
10	VR	9200	8360	2800	82.7	349.5	5.5	8.8	0.802	1	9200	4746.60	1479.94	48.83	174.75	2.75	24.69
11	Rail Cargo Hungaria	7606	2600	1460	42.3	475.7	13.96	5.85	1	1	7606	2600	1460	42.3	475.7	13.96	0
12	Mercitalia	16800	12380	3000	76.1	499.8	10.2	8.85	0.575	0.760	16800	6249.69	1995.05	74.80	249.90	5.1	32.12
13	DB NL	7000	7000	620	51.1	149	1.5	6.26	1	1	7000	7000	620	51.1	149	1.5	0

Πίνακας 4.9 Αποτελέσματα της εφαρμογής του μοντέλου ANN για Οικονομική Αποδοτικότητα

DMU	ΟΡΓΑΝΙΣΜΟΣ	X1	X2	X3	X4	X5	X6	Y2	Theta_DEAP	Theta_ANN_best	X1	X2_best	X3_best	X4_best	X5_best	X6_best	Improvement_%
1	DB Cargo	33400	81536	31350	319	6079	20.1	5582	1	1	33400	81536	31350	319	6079	20.1	0
2	Xefafet	28000	7677	5000	105	630	1.99	714	1	1	28000	7677	5000	105	630	1.99	0
3	Renfe	15500	7200	945	232.7	210.89	6.9	198.81	0.885	0.969	15500	6083.34	767.32	214.19	194.42	6.9	9.53
4	CD Cargo	9500	20411	6628	180	488.5	7.22	530	1	1	9500	20411	6628	180	488.5	7.22	0
5	OBB	5000	24973	5912	170.5	1920	2.9	1910	1	1	5000	24973	5912	170.5	1920	2.9	0
6	ZSSK	3600	2239	3771	69.8	301.8	6.5	303.6	1	1	3600	2239	3771	69.8	301.8	6.5	0
7	Hupac	5300	8500	700	35	668.44	10.34	667	1	1	5300	8500	700	35	668.44	10.34	0
8	Green Cargo	10900	5570	1800	36.5	369.2	2.28	365.4	0.958	1	10900	4794.13	1431.41	36.50	283.89	2.28	4.38
9	PKP Cargo	18800	49757	14062	349.2	897.82	18.49	940.3	0.963	1	18800	37450.10	10275.89	283.81	644.47	18.49	3.84
10	VR	9200	8360	2800	82.7	349.5	5.5	339.9	0.899	1	9200	4346.54	1400.00	51.69	266.88	5.5	11.23
11	Rail Cargo Hungaria	7606	2600	1460	42.3	475.7	13.96	473	1	1	7606	2600	1460	42.3	475.7	13.96	0
12	Mercitalia	16800	12380	3000	76.1	499.8	10.2	480	0.878	0.967	16800	7756.07	2457.14	76.10	381.41	10.2	10.13
13	DB NL	7000	7000	620	51.1	149	1.5	137	0.868	0.903	7000	3500	620	51.10	117.58	1.5	3.99

Πίνακας 4.10 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Τεχνικής Αποδοτικότητας ANN

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Volume of ton-kilometers transported (B t-km)
α/α		ΟΡΓΑΝΙΣΜΟΣ		Χ1	Χ2	Χ3	Χ4	Χ5	Χ6	Υ1
3	Πραγματικά Δεδομένα	Renfe Mercancías	Ισπανία	15500.0	7200.0	945.0	232.7	210.9	6.9	4.4
	Δεδομένα Στόχος			15500.0	5395.0	945.0	116.4	145.6	3.5	
	Ποσοστιαία μείωση			0.00%	25.07%	0.00%	50.00%	30.98%	50.00%	
9	Πραγματικά Δεδομένα	PKP Cargo	Πολωνία	18800.0	49757.0	14062.0	349.2	897.8	18.5	22.3
	Δεδομένα Στόχος			18800.0	40512.3	10850.7	287.4	449.1	9.2	
	Ποσοστιαία μείωση			0.00%	18.58%	22.84%	17.72%	49.98%	50.00%	
10	Πραγματικά Δεδομένα	VR Transpoint	Φινλανδία	9200.0	8360.0	2800.0	82.7	349.5	5.5	8.8
	Δεδομένα Στόχος			9200.0	4746.6	1479.9	48.8	174.8	2.8	
	Ποσοστιαία μείωση			0.00%	43.22%	47.14%	40.96%	50.00%	50.00%	
12	Πραγματικά Δεδομένα	Mercitalia Rail	Ιταλία	16800.0	12380.0	3000.0	76.1	499.8	10.2	8.9
	Δεδομένα Στόχος			16800.0	6249.7	1995.1	74.8	249.9	5.1	
	Ποσοστιαία μείωση			0.00%	49.52%	33.50%	1.71%	50.00%	50.00%	

Πίνακας 4.11 Προτεινόμενα Επίπεδα Εισροών για Βελτίωση Οικονομικής Αποδοτικότητας ANN

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Revenue (MEUR)
α/α		ΟΡΓΑΝΙΣΜΟΣ		X1	X2	X3	X4	X5	X6	Y2
3	Πραγματικά Δεδομένα	Renfe Mercancías	Ισπανία	15500.0	7200.0	945.0	232.7	210.9	6.9	198.8
	Δεδομένα Στόχος			15500.0	6083.3	767.3	214.2	194.4	6.9	
	Ποσοστιαία μείωση			0.00%	15.51%	18.80%	7.96%	7.81%	0.00%	
8	Πραγματικά Δεδομένα	Green Cargo	Σουηδία	10900.0	5570.0	1800.0	36.5	369.2	2.3	365.4
	Δεδομένα Στόχος			10900.0	4794.1	1431.4	36.5	283.9	2.3	
	Ποσοστιαία μείωση			0.00%	13.93%	20.48%	0.00%	23.11%	0.00%	
9	Πραγματικά Δεδομένα	PKP Cargo	Πολωνία	18800.0	49757.0	14062.0	349.2	897.8	18.5	940.3
	Δεδομένα Στόχος			18800.0	37450.1	10275.9	283.8	644.5	18.5	
	Ποσοστιαία μείωση			0.00%	24.73%	26.92%	18.74%	28.22%	0.00%	
10	Πραγματικά Δεδομένα	VR Transpoint	Φινλανδία	9200.0	8360.0	2800.0	82.7	349.5	5.5	339.9
	Δεδομένα Στόχος			9200.0	4346.5	1400.0	51.7	266.9	5.5	
	Ποσοστιαία μείωση			0.00%	48.01%	50.00%	37.50%	23.64%	0.00%	
12	Πραγματικά Δεδομένα	Mercitalia Rail	Ιταλία	16800.0	12380.0	3000.0	76.1	499.8	10.2	480.0
	Δεδομένα Στόχος			16800.0	7756.1	2457.1	76.1	381.4	10.2	
	Ποσοστιαία μείωση			0.00%	37.35%	18.10%	0.00%	23.69%	0.00%	
13	Πραγματικά Δεδομένα	DB Cargo Netherlands	Ολλανδία	7000.0	7000.0	620.0	51.1	149.0	1.5	137.0
	Δεδομένα Στόχος			7000.0	3500.0	620.0	51.1	117.6	1.5	
	Ποσοστιαία μείωση			0.00%	50.00%	0.00%	0.00%	21.09%	0.00%	

Στη διαδικασία που περιγράφηκε παραπάνω, εφαρμόστηκε το μοντέλο Τεχνητών Νευρωνικών Δικτύων (ANN) και στη συνέχεια τα αποτελέσματά του ελέγχθηκαν ξανά με τη μέθοδο της Περιβάλλουσας Ανάλυσης Δεδομένων. Από τη διαδικασία αυτή φάνηκε ότι είναι εφικτό να προβλεφθούν πιο ρεαλιστικές και στοχευμένες λύσεις για τη βελτίωση της αποδοτικότητας. Με αυτόν τον τρόπο ξεπεράστηκαν αρκετοί από τους περιορισμούς της παραδοσιακής DEA ανάλυσης, ενώ το ANN λειτούργησε ως ένα επιπλέον εργαλείο για την καλύτερη προσαρμογή των εισροών στις ανάγκες κάθε οργανισμού.

Παρακάτω παρουσιάζονται αναλυτικά οι συγκεντρωτικοί πίνακες που εμφανίζουν τις ποσοστιαίες μεταβολές που προκύπτουν από την DEA και συγκρίνονται με αυτές που προκύπτουν από το ANN.

Πίνακας 4.12 Ποσοστιαία διαφορά μεταβολής μεταβλητών DEA-ANN για Τεχνική Αποδοτικότητα

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Μέσος Όρος Βελτίωσης
α/α		ΟΡΓΑΝΙΣΜΟΣ		Χ1	Χ2	Χ3	Χ4	Χ5	Χ6	
3	Ποσοστιαία μείωση DEA	Renfe Mercancías	Ισπανία	68.41%	42.14%	42.14%	87.12%	42.14%	85.01%	-35.15%
	Ποσοστιαία μείωση ANN			0.00%	25.07%	0.00%	50.00%	30.98%	50.00%	
	Ποσοστιαία διαφορά			-68.41%	-17.07%	-42.14%	-37.12%	-11.16%	-35.01%	
9	Ποσοστιαία μείωση DEA	PKP Cargo	Πολωνία	36.57%	51.86%	47.12%	40.14%	36.57%	55.27%	-18.07%
	Ποσοστιαία μείωση ANN			0.00%	18.58%	22.84%	17.72%	49.98%	50.00%	
	Ποσοστιαία διαφορά			-36.57%	-33.28%	-24.28%	-22.43%	13.41%	-5.27%	
10	Ποσοστιαία μείωση DEA	VR Transpoint	Φινλανδία	19.79%	19.79%	22.90%	36.37%	19.79%	54.16%	9.75%
	Ποσοστιαία μείωση ANN			0.00%	43.22%	47.14%	40.96%	50.00%	50.00%	
	Ποσοστιαία διαφορά			-19.79%	23.43%	24.24%	4.59%	30.21%	-4.16%	
12	Ποσοστιαία μείωση DEA	Mercitalia Rail	Ιταλία	42.93%	48.75%	54.76%	42.46%	42.46%	80.19%	-21.14%
	Ποσοστιαία μείωση ANN			0.00%	49.52%	33.50%	1.71%	50.00%	50.00%	
	Ποσοστιαία διαφορά			-42.93%	0.77%	-21.26%	-40.75%	7.54%	-30.19%	

Πίνακας 4.13 Ποσοστιαία διαφορά μεταβολής μεταβλητών DEA-ANN για Οικονομική Αποδοτικότητα

DMU	Δεδομένα	ΣΙΔΗΡΟΔΡΟΜΟΙ		μήκος δικτύου (Km)	Αριθμός τροχαίου υλικού (μονάδες)	Αριθμός προσωπικού (μονάδες)	Κεφάλαιο επένδυσης (MEUR)	Λειτουργικά έξοδα (MEUR)	Ενεργειακή κατανάλωση (g/tkm)	Μέσος Όρος Βελτίωσης
α/α		ΟΡΓΑΝΙΣΜΟΣ		Χ1	Χ2	Χ3	Χ4	Χ5	Χ6	
3	Ποσοστιαία μείωση DEA	Renfe Mercancías	Ισπανία	68.54%	67.73%	11.51%	91.23%	11.51%	74.75%	-45.87%
	Ποσοστιαία μείωση ANN			0.00%	15.51%	18.80%	7.96%	7.81%	0.00%	
	Ποσοστιαία διαφορά			-68.54%	-52.22%	7.29%	-83.28%	-3.70%	-74.75%	
8	Ποσοστιαία μείωση DEA	Green Cargo	Σουηδία	49.59%	19.39%	25.74%	4.21%	4.21%	4.21%	-8.30%
	Ποσοστιαία μείωση ANN			0.00%	13.93%	20.48%	0.00%	23.11%	0.00%	
	Ποσοστιαία διαφορά			-49.59%	-5.46%	-5.26%	-4.21%	18.90%	-4.21%	
9	Ποσοστιαία μείωση DEA	PKP Cargo	Πολωνία	3.73%	30.48%	18.67%	11.78%	3.73%	34.15%	-0.65%
	Ποσοστιαία μείωση ANN			0.00%	24.73%	26.92%	18.74%	28.22%	0.00%	
	Ποσοστιαία διαφορά			-3.73%	-5.75%	8.26%	6.95%	24.49%	-34.15%	
10	Ποσοστιαία μείωση DEA	VR Transpoint	Φινλανδία	10.13%	24.96%	10.13%	25.43%	10.13%	67.65%	1.79%
	Ποσοστιαία μείωση ANN			0.00%	48.01%	50.00%	37.50%	23.64%	0.00%	
	Ποσοστιαία διαφορά			-10.13%	23.05%	39.87%	12.08%	13.51%	-67.65%	
12	Ποσοστιαία μείωση DEA	Mercitalia Rail	Ιταλία	12.19%	56.22%	12.19%	22.34%	12.19%	73.14%	-18.19%
	Ποσοστιαία μείωση ANN			0.00%	37.35%	18.10%	0.00%	23.69%	0.00%	
	Ποσοστιαία διαφορά			-12.19%	-18.87%	5.90%	-22.34%	11.49%	-73.14%	
13	Ποσοστιαία μείωση DEA	DB Cargo Netherlands	Ολλανδία	54.85%	76.94%	13.18%	73.67%	13.18%	14.60%	-29.22%
	Ποσοστιαία μείωση ANN			0.00%	50.00%	0.00%	0.00%	21.09%	0.00%	
	Ποσοστιαία διαφορά			-54.85%	-26.94%	-13.18%	-73.67%	7.90%	-14.60%	

Παρατηρώντας τις ποσοστιαίες μεταβολές στις προτεινόμενες εισροές από τη μέθοδο DEA και το νευρωνικό δίκτυο ANN, προκύπτουν κρίσιμες διαφορές ως προς τον ρεαλισμό και το κατά πόσο επιχειρησιακά εφικτές είναι οι λύσεις. Η DEA, ως μεθοδολογία βασισμένη σε γραμμικό προγραμματισμό, προτείνει συχνά δραστικές και γραμμικές μειώσεις στις εισροές, προκειμένου να μεταφέρει κάθε μη αποδοτική μονάδα στο αποδοτικό μέτωπο. Αν και αυτές οι λύσεις είναι υπολογιστικά ορθές, συχνά δεν ανταποκρίνονται στις πραγματικές συνθήκες λειτουργίας των σιδηροδρομικών οργανισμών. Σε ορισμένες περιπτώσεις, οι μειώσεις που προτείνονται υπερβαίνουν μέχρι και το 60–80%

Αντιθέτως, οι προβλέψεις του ANN χαρακτηρίζονται από μεγαλύτερη ισορροπία και ρεαλισμό. Οι ποσοστιαίες μεταβολές που προτείνει το νευρωνικό δίκτυο είναι εμφανώς μικρότερες στις περισσότερες μεταβλητές, με προσαρμογές που κινούνται εντός ενός λογικού και διαχειρίσιμου εύρους. Πιο συγκεκριμένα το ANN πρότείνει βελτιστοποιημένα αποτελέσματα:

- Για την τεχνική αποτελεσματικότητα σε 17 από τις 24 μεταβλητές
- Για την οικονομική αποτελεσματικότητα σε 24 από τις 36 μεταβλητές

Ενδεικτικά, η πιο μεγάλη βελτίωση για την οικονομική αποτελεσματικότητα παρατηρείται στην μεταβλητή X4 (κεφάλαιο επένδυσης) στον οργανισμό Renfe Mercancías όπου η DEA προτείνει μείωση της μεταβλητής κατά 91.23% ενώ το ANN 7.96% με ποσοστό βελτίωσης 83.28%. Αντίστοιχα, για την τεχνική αποτελεσματικότητα η πιο μεγάλη βελτίωση παρατηρείται στον ίδιο οργανισμό στην μεταβλητή X1 (μήκος δικτύου) όπου η DEA προτείνει μείωση της μεταβλητής 68.41% ενώ το AN 0.00% με ποσοστό βελτίωσης 68.41%.

Επιπλέον, ο μέσος όρος βελτίωσης επιβεβαιώνει με ποσοτικό τρόπο τη διαπίστωση ότι οι λύσεις που προτείνει το ANN είναι ηπιότερες και περισσότερο προσαρμοσμένες στις πραγματικές δυνατότητες των οργανισμών. Ενδεικτικά, παρατηρείται ότι για όλους τους οργανισμούς, εκτός μιας μεμονωμένης περίπτωσης (VR Transpoint), ο μέσος όρος των ποσοστιαίων μειώσεων εισροών που προτάθηκε από το ANN είναι χαμηλότερος από εκείνον της DEA. Αυτό επιβεβαιώνει τη λειτουργία του νευρωνικού δικτύου ως εργαλείου βελτιστοποίησης με πρακτικό προσανατολισμό, καθώς λαμβάνει υπόψη τον συνδυασμό πολλών παραμέτρων αντί να εφαρμόζει αυστηρά γραμμικές μειώσεις βάσει ενός θεωρητικού μοντέλου.

Αυτό αποδεικνύει ότι το μοντέλο ANN, εκπαιδευμένο με δεδομένα DEA και εμπλουτισμένο με παραδείγματα αποδοτικών και μη αποδοτικών μονάδων, έμαθε να προτείνει λύσεις που προσεγγίζουν το αποδοτικό μέτωπο χωρίς να παραβλέπουν τους πρακτικούς περιορισμούς. Επιπλέον, οι προτάσεις του ANN δεν βασίζονται μόνο στη θεωρητική μαθηματική βελτιστοποίηση, αλλά και στην εσωτερική κατανόηση της σχέσης μεταξύ εισροών και αποδοτικότητας, την οποία ανέπτυξε μέσα από την εκπαίδευση στο εμπλουτισμένο σύνολο δεδομένων (dataset).

Η σύγκριση των ποσοστιαίων μειώσεων αποκαλύπτει τη σημαντική προστιθέμενη αξία της χρήσης τεχνητής νοημοσύνης στην ανάλυση αποδοτικότητας. Το ANN παρείχε πιο φιλικές ως προς την υλοποίηση προτάσεις, διατηρώντας σε μεγαλύτερο βαθμό τις κρίσιμες εισροές (όπως ανθρώπινο δυναμικό ή επενδύσεις) και επιτυγχάνοντας ταυτόχρονα αξιολογη βελτίωση της αποδοτικότητας. Η ικανότητα του μοντέλου να προτείνει ρεαλιστικά σενάρια καθιστά τη μέθοδο DEA–ANN ένα εξελιγμένο εργαλείο

σχεδιασμού, πιο προσαρμοσμένο στις πραγματικές ανάγκες των σιδηροδρομικών φορέων συγκριτικά με την παραδοσιακή προσέγγιση DEA.

Τα αποτελέσματα που προέκυψαν από την ανάλυση τόσο με τη μέθοδο DEA όσο και με την υβριδική προσέγγιση DEA–ANN προσφέρουν μια ολοκληρωμένη εικόνα της αποδοτικότητας των εξεταζόμενων οργανισμών και της δυνατότητας στοχευμένων βελτιώσεων. Στο επόμενο κεφάλαιο, επιχειρείται η σύνθεση των βασικών ευρημάτων της μελέτης, με στόχο την ανάδειξη των κύριων συμπερασμάτων και τη διατύπωση προτάσεων για μελλοντική ερευνητική δραστηριότητα.

5. Συμπεράσματα

5.1 Γενικά

Η παρούσα διπλωματική εργασία επιχειρεί να συμβάλει στην εκτίμηση και ενίσχυση της αποδοτικότητας των ευρωπαϊκών σιδηροδρομικών οργανισμών εμπορευματικών μεταφορών, αξιοποιώντας έναν συνδυασμό παραδοσιακών και σύγχρονων μεθοδολογιών επιχειρησιακής έρευνας και τεχνητής νοημοσύνης. Η ερευνητική προσέγγιση που υιοθετήθηκε ενσωματώνει τα πλεονεκτήματα του μοντέλου DEA, το οποίο αποτελεί ένα ευρέως αποδεκτό εργαλείο για τη συγκριτική αξιολόγηση της αποδοτικότητας ομοειδών μονάδων, με τις δυνατότητες των Τεχνητών Νευρωνικών Δικτύων (ANN) να παρέχουν προσαρμοστικές προτάσεις βελτιστοποίησης βασισμένες σε δεδομένα.

Η μεθοδολογία συνδυάζει τη διαγνωστική ικανότητα της DEA, που επιτρέπει τον εντοπισμό αποδοτικών και μη αποδοτικών οργανισμών και την ανάλυση των υπολειμμάτων (slacks), με την προβλεπτική ισχύ των ANN, τα οποία εκπαιδεύονται πάνω σε DEA-αξιολογημένα σύνολα δεδομένων για να παράγουν ρεαλιστικά σενάρια βελτίωσης. Με τον τρόπο αυτό, η παρούσα διπλωματική εργασία επιδιώκει να υπερβεί τα όρια της DEA, προσφέροντας ένα δυναμικό και πρακτικά αξιοποιήσιμο πλαίσιο υποστήριξης αποφάσεων για τη βελτίωση της αποδοτικότητας των οργανισμών.

Τα αποτελέσματα που προέκυψαν επιβεβαιώνουν την προστιθέμενη αξία του συνδυαστικού αυτού μοντέλου. Η DEA εντόπισε πλήρως αποδοτικές και μη αποδοτικές μονάδες τόσο σε τεχνικούς όσο και σε οικονομικούς όρους, ενώ τα ANN παρήγαγαν προτάσεις βελτιωμένων συνδυασμών εισροών, οι οποίες, μετά από επανεξέταση μέσω της DEA, οδήγησαν σε σημαντικές βελτιώσεις των δεικτών αποδοτικότητας για όλες τις μη αποδοτικές μονάδες.

Συνολικά, η διερεύνηση που πραγματοποιήθηκε αναδεικνύει ότι η υβριδική προσέγγιση DEA-ANN μπορεί να αποτελέσει ένα πολύτιμο εργαλείο για τους φορείς σιδηροδρομικών εμπορευματικών μεταφορών, προσφέροντας όχι μόνο διαγνωστική αξιολόγηση αλλά και εφαρμόσιμες συστάσεις για τη βελτιστοποίηση της χρήσης πόρων και την ενίσχυση της αποδοτικότητας.

5.2 Βασικά συμπεράσματα

Η παρούσα διπλωματική εργασία ανέπτυξε και εφάρμοσε ένα υβριδικό μοντέλο αξιολόγησης και βελτιστοποίησης της αποδοτικότητας σιδηροδρομικών οργανισμών εμπορευματικών μεταφορών, συνδυάζοντας τη μέθοδο Περιβάλλουσας Ανάλυσης Δεδομένων (DEA) με τα Τεχνητά Νευρωνικά Δίκτυα (ANN). Το μοντέλο αυτό επιδιώκει όχι μόνο να αξιολογήσει την αποδοτικότητα των εξεταζόμενων οργανισμών αλλά και να προτείνει ρεαλιστικές και εφαρμόσιμες λύσεις βελτίωσής της.

Η εφαρμογή της DEA αποκάλυψε ουσιώδεις διαφοροποιήσεις στο επίπεδο τεχνικής και οικονομικής αποδοτικότητας μεταξύ των 13 ευρωπαϊκών σιδηροδρομικών εμπορευματικών οργανισμών που εξετάστηκαν. Συγκεκριμένα:

- Σε τεχνικό επίπεδο (με εκροή τα ton-kilometres): 9 οργανισμοί χαρακτηρίστηκαν ως πλήρως αποδοτικοί, ενώ οι υπόλοιποι παρουσίασαν σημαντικά περιθώρια βελτίωσης. Ο μέσος όρος αποδοτικότητας κυμάνθηκε στο 0.892.
- Σε οικονομικό επίπεδο (με εκροή τα έσοδα): 7 οργανισμοί θεωρήθηκαν πλήρως αποδοτικοί, με τον μέσο όρο αποδοτικότητας να ανέρχεται στο 0.958.

Η ανάλυση των slacks κατέδειξε ότι σε αρκετές περιπτώσεις, οι οργανισμοί διατηρούσαν πλεονάζουσες εισροές, ειδικά σε μεταβλητές όπως ο αριθμός προσωπικού, η ενεργειακή κατανάλωση και οι κεφαλαιουχικές δαπάνες, προσδιορίζοντας συγκεκριμένα πεδία με σαφές δυναμικό για περαιτέρω εξορθολογισμό.

Ωστόσο, η DEA ως μέθοδος παρουσίασε ορισμένους περιορισμούς:

- Τα προτεινόμενα σενάρια βελτίωσης συχνά ήταν μη ρεαλιστικά ή επιχειρησιακά δυσχερή στην υλοποίηση, καθώς προέβλεπαν γραμμικές και ενίοτε δραστικές μειώσεις εισροών.
- Η μέθοδος, εκ φύσεως στατική και μη προβλεπτική, δεν επιτρέπει την πρόβλεψη αποδοτικών εισροών για νέες ή διαφορετικές συνθήκες λειτουργίας.

Η ενσωμάτωση των ANN στο μοντέλο DEA-ANN ήρθε να αντιμετωπίσει αυτά τα μειονεκτήματα. Με την εκπαίδευση του νευρωνικού δικτύου πάνω σε δεδομένα DEA (εισροές – εκροές – αποτελεσματικότητα θ), κατέστη δυνατή η πρόβλεψη εναλλακτικών, προσαρμοσμένων και ρεαλιστικών σεναρίων εισροών που μεγιστοποιούν την αποδοτικότητα κάθε οργανισμού.

Τα κύρια ευρήματα από την εφαρμογή του μοντέλου ANN ήταν τα εξής:

- Όλοι οι μη αποδοτικοί οργανισμοί βελτίωσαν τις τιμές αποδοτικότητάς τους όταν οι προβλεπόμενες από τα ANN εισροές επανεκτιμήθηκαν μέσω DEA.
- Τα ANN προσέφεραν ρεαλιστικότερες προσαρμογές εισροών σε σχέση με τις αρχικές προβολές της DEA, λαμβάνοντας υπόψη πρακτικούς περιορισμούς (π.χ. μη δυνατότητα δραστήριας μείωσης του ανθρώπινου δυναμικού ή του μήκους σιδηροδρομικού δικτύου).
- Η χρήση ANN επέτρεψε την πρόβλεψη αποδοτικών σεναρίων ακόμη και για εισροές ή καταστάσεις που δεν υπήρχαν στο αρχικό δείγμα, προσδίδοντας στο μοντέλο προγνωστική ικανότητα, από την οποία η παραδοσιακή DEA υπολείπεται.
- Ο συνδυασμός DEA-ANN δημιούργησε έναν δυναμικό κύκλο βελτιστοποίησης, όπου οι προβλεπόμενες λύσεις επανελέγχονται και επανατροφοδοτούνται στο σύστημα, επιτρέποντας τη σταδιακή προσέγγιση του αποδοτικού μετώπου.

Συγκρίνοντας τα προτεινόμενα επίπεδα εισροών που προέκυψαν από την ανάλυση DEA με εκείνα που παρήχθησαν από το μοντέλο ANN, παρατηρούνται σημαντικές διαφοροποιήσεις. Η DEA, εστιάζοντας αποκλειστικά στη βέλτιστη τεχνική αποδοτικότητα, πρότεινε σε ορισμένες περιπτώσεις εξαιρετικά υψηλά ποσοστά μείωσης εισροών, που ενδέχεται να είναι μη ρεαλιστικά σε επιχειρησιακό επίπεδο. Αντίθετα, οι προβλέψεις των ANN αποδείχθηκαν πιο ισορροπημένες, με πιο ήπιες

προσαρμογές στις εισροές, ειδικά για μεταβλητές όπως το προσωπικό (X3), τα λειτουργικά έξοδα (X5) και η ενεργειακή κατανάλωση (X6).

Επιπλέον, παρατηρήθηκε ότι το ANN τείνει να συγκρατεί τις μειώσεις σε κρίσιμες εισροές, όπως το τροχαίο υλικό (X2), προτείνοντας σενάρια πιο συμβατά με την πραγματική λειτουργία των οργανισμών. Συνολικά, η προσέγγιση ANN παρήγαγε περισσότερο εφαρμόσιμες και επιχειρησιακά υλοποιήσιμες λύσεις, αναδεικνύοντας τη σημαντική προστιθέμενη αξία του συνδυαστικού μοντέλου DEA-ANN.

Η μεθοδολογία αυτή προσφέρει στους διαχειριστές των σιδηροδρομικών εμπορευματικών οργανισμών ένα πρακτικό εργαλείο λήψης αποφάσεων και σχεδιασμού βελτιώσεων, συμβάλλοντας στον στρατηγικό προγραμματισμό και στην αποδοτικότερη κατανομή των πόρων.

5.3 Προτάσεις για περαιτέρω έρευνα

Η παρούσα διπλωματική ανέδειξε την αξία της υβριδικής προσέγγισης DEA-ANN για την εκτίμηση και βελτιστοποίηση της αποδοτικότητας στον τομέα των σιδηροδρομικών εμπορευματικών μεταφορών. Ωστόσο, τα ευρήματα και οι περιορισμοί που προέκυψαν από την ανάλυση ανοίγουν τον δρόμο για περαιτέρω ερευνητικές προσπάθειες που μπορούν να ενισχύσουν ακόμα περισσότερο τη μεθοδολογία και την πρακτική της εφαρμογή.

Πρωτίστως, θα ήταν ιδιαίτερα χρήσιμο να διερευνηθεί η εφαρμογή της προσέγγισης αυτής σε μεγαλύτερα και πιο αντιπροσωπευτικά σύνολα δεδομένων, που να καλύπτουν ευρύτερο φάσμα σιδηροδρομικών οργανισμών και διαφορετικά επιχειρησιακά μοντέλα. Ένα πιο πλούσιο dataset θα επέτρεπε στα ANN να γενικεύσουν αποτελεσματικότερα, ενισχύοντας την ακρίβεια και την αξιοπιστία των προτεινόμενων σεναρίων εισροών.

Παράλληλα, μελλοντική έρευνα θα μπορούσε να ενσωματώσει στο μοντέλο ανεπιθύμητες εκροές (undesirable outputs), όπως οι εκπομπές CO₂, ώστε η εκτίμηση αποδοτικότητας να λαμβάνει υπόψη και τις περιβαλλοντικές διαστάσεις της λειτουργίας των οργανισμών, συμβάλλοντας στη διαμόρφωση στρατηγικών με υψηλότερη βιωσιμότητα.

Επιπλέον, η ενσωμάτωση πιο εξελιγμένων τεχνικών τεχνητής νοημοσύνης, όπως τα Γενετικά Αλγορίθμια (Genetic Algorithms) και η Ενισχυτική Μάθηση (Reinforcement Learning), μπορεί να προσδώσει ακόμα μεγαλύτερη προβλεπτική ισχύ και να βελτιώσει την αναζήτηση βέλτιστων λύσεων στο πλαίσιο μη γραμμικών και πολυπαραγοντικών προβλημάτων.

Τέλος, ένα σημαντικό πεδίο για περαιτέρω διερεύνηση αποτελεί η εφαρμογή της προσέγγισης DEA-ANN σε δυναμικά περιβάλλοντα (Dynamic DEA), επιτρέποντας τη μελέτη της αποδοτικότητας σε βάθος χρόνου και την αξιολόγηση της επίδρασης διαχρονικών στρατηγικών βελτίωσης. Η συνέχιση της έρευνας προς αυτές τις κατευθύνσεις θα μπορούσε να οδηγήσει στη δημιουργία ενός πλήρως προσαρμόσιμου, έξυπνου συστήματος υποστήριξης αποφάσεων για σιδηροδρομικούς οργανισμούς, ενισχύοντας περαιτέρω τη συμβολή της παρούσας μελέτης στην επίλυση σύγχρονων προκλήσεων στον τομέα των εμπορευματικών μεταφορών.

6. Βιβλιογραφία

Abate, M., Lijesen, M., Pels, E., & Roeleveld, A., (2013). The impact of reliability on the productivity of railroad companies. *Transp. Res.: Part E*, pp. 5141–5149.

Albino, V., & Garavelli, A. C. (1998). A neural network application to subcontractor rating in construction firms. *International Journal of Project Management*, Vol. 16, pp. 9–14.

Amirteimoori, A., Allahviranloo, T., Zadmirzaei, M., & Hasanzadeh, F. (2023). On the environmental performance analysis: A combined fuzzy data envelopment analysis and artificial intelligence algorithms, *Expert Systems with Applications*, Vol. 224, Article 119953. DOI: 10.1016/j.eswa.2023.119953

Amirteimoori, A., Zadmirzaei, M., & Hassanzadeh, F. (2021). Developing a new integrated artificial Immune system and fuzzy non-discretionary DEA approach. *Soft Computing*, Vol. 25, pp. 8109–8127.

Andrikopoulos, A.A., & Loizides, J. (1998). Cost structure and productivity growth in European railway systems. *Appl. Econ.* Vol. 30, pp. 1625–1639.

Archer, N., & Wang, S. (1993). Application of the back propagation neural network algorithm with monotonicity constraints for two-group classification problems. *Decis. Sci.* Vol. 24, pp. 60–75.

Athanassopoulos, A.D., & Curram, S.P. (1996). A comparison of data envelopment analysis and artificial neural networks as tools for assessing the efficiency of decision making units, *Journal of the Operational Research Society*, Vol. 47, pp. 1000–1016.

Azadeh, A., Ghaderi, S.F., & Izadbakhsh, H. (2008). Integration of DEA and AHP with computer simulation for railway system improvement and optimization, *Applied Mathematics and Computation*, Vol. 195, pp. 775–785. DOI: 10.1016/j.amc.2007.05.023

Azadeh, A., Saberi, M., & Anvari, M. (2010). An integrated artificial neural network algorithm for performance assessment and optimization of decision making units. *Expert Syst. Appl.* Vol. 36, pp. 5688–5697.

Azadeh, A., Saberi, M., Moghaddam, R.T., & Javanmardi, L. (2011). An integrated data envelopment analysis-artificial neural network-rough set algorithm for assessment of personnel efficiency. *Expert Syst. Appl.* Vol. 38, pp. 1364–1373.

Banaszewska, A., Cruijssen, F., Dullaert, W., & Gerdessen, J.C. (2012). A framework for measuring efficiency levels– the case of express depots. *Int. J. Prod. Econ.* Vol. 139, pp.484–495.

Barros, C.P., & Wanke, P. (2014). Insurance companies in Mozambique: a two-stage DEA and neural networks on efficiency and capacity slacks. *Appl. Econ.* Vol. 46, pp. 3591–3600.

Battese, G.E., & Coelli, T.J. (1988). Prediction of firm level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, Vol. 38, pp. 387-399.

Battese, G.E., & Coelli, T.J. (1991). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. Department of Econometrics, University of New England, Australia.

Bjurek, H., Hjalmarsson, L., & Forsund, F.R. (1990). Parametric and nonparametric estimation of efficiency in service production: a comparison. *Frontier analysis, parametric and non parametric approaches*, edited by A.Y. Lewin and C.A. Knox Love in *Journal of Econometrics Annals 1990-4*, 46-1/2, pp. 213-228.

Blagojević, A., Vesković, S., Kasalica, S., Gojić, A., & Allamani, A. (2020). The Application of the Fuzzy AHP and DEA for Measuring the Efficiency of Freight Transport Railway Undertakings. *Operational Research in Engineering Sciences: Theory and Applications*, Vol. 3, Is.2, pp. 1–23 DOI: <https://doi.org/10.31181/oresta2003001b>

Breusch, TX., Mizon, G.E., & Schmidt, P. (1989). Efficient estimation using panel data. *Econometrica* 57, pp. 695-700.

Carlucci, D., Renna, P., & Schiuma, G. (2013). Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. *Health Care Manag. Sci.* Vol. 16, pp. 37–44.

Cavalieria, S., Maccarrone, P., & Pinto, R. (2004). Parametric vs. neural network models for the estimation of production costs: a case study in the automotive industry. *Int. J. Prod. Econ.* Vol. 91, pp. 165–177.

Çakır, S. (2024). Best output prediction in OECD railways using DEA in conjunction with machine learning algorithms, *Annals of Operations Research*, Vol. 335, pp. 59–77. DOI: 10.1007/s10479-023-05668-w

Çelebi, D., & Bayraktar, D. (2008). An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. *Expert Syst. Appl.* Vol. 35, pp. 1698–1710.

Chandra, P., Cooper, W., Shanling, L., Rahman, A. (1998). Using DEA to evaluate 29 Canadian textile companies—considering returns to scale. *Int. J. Prod. Econ.* Vol. 54, pp. 129–141.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units, *European Journal of Operational Research*, Vol. 2, Is. 429, 429-444

Coelli, T., Estache, A., Perelman, S., & Trujillo, L. (2003). *A Primer on Efficiency Measurement for Utilities and Transport Regulators*. The World Bank Institute

Coelli, T., Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: with application to European railways. *Eur. J. Oper. Res.* Vol. 117, pp. 326–339.

Coelli, T., Perelman, S. (2000). Technical efficiency of European railways: a distance function approach. *Appl. Econ.* Vol. 32, pp. 1967–1976.

Cornwell, C.P., Schmidt, P., & Sickles, R.C. (1990). Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels. *Frontier analysis, parametric and non parametric approaches*, edited by A.Y. Lewin and C.A. Knox Love11 in *Journal of Econometrics Annals* 1990-4, 46-1/2, pp. 185-200.

Costa, Á., & Markellos R. N. (1997). Evaluating public transport efficiency with neural network models, *Transportation Research-C: Vol. 5, No. 5*, pp. 301–312. PII:S0968-090X(97)00017-X

Cowie, J. & Riddington, G. (1996). Measuring the efficiency of European railways, *Applied Economics*, 28:8, 1027-1035, DOI: 10.1080/000368496328155

Doomernik, J. E. (2014). Performance and efficiency of High-speed Rail systems, *Transportation Research Procedia*, 8, 136 – 144 DOI: 10.1016/j.trpro.2015.06.049

Fausett, L., (1994). *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Prentice Hall, Englewood Cliffs, NJ.

Ferreira da Silva, F. G., Magalhães de Oliveira, R. L. M., & Marinov, M. (2020). An Analysis of the Effects on Rail Operational Efficiency Due to a Merger between

Brazilian Rail Companies: The Case of RUMO-ALL. Sustainability, 12(4827), DOI:10.3390/su12124827

Fu, L. M. (1994). Neural Networks in Computer Intelligence. McGraw Hill, New York

Gathon, H.-J., & Perelman, S. (1992). Measuring Technical Efficiency in European Railways: A Panel Data Approach. Journal of Productivity Analysis, 3, 135-151.

Holvad, T. (2020). Efficiency analyses for the railway sector: An overview of key issues. Research in Transportation Economics, 82, 100877, <https://doi.org/10.1016/j.retrec.2020.100877>

Hu, M., Zhang, G., Jiang, C., Patuwo, E. (1999). A cross validation analysis of neural network out-of-sample performance in exchange rate forecasting. Decis. Sci. Vol. 30, pp. 197–216.

ITF (2019). Efficiency in Railway Operations and Infrastructure Management. ITF Roundtable Reports, No. 177, OECD Publishing, Paris.

Jondrow, J., Lovell, C.A.K., Materov, IS., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics Vol. 19, pp. 233-238.

Karakeke, V. (2020). A Freight Transport Mode Selection Tool Based on a Total Cost Approach Analysis: Application on the Orient-East Mediterranean Corridor, Aristotle University of Thessaloniki, MSc Thesis

Kumbhakar, S.C. (1988). Production frontiers, panel data and time-varying technical inefficiency. Frontier analysis, parametric and non parametric approaches, edited by A.Y. Lewin and C.A. Knox Lovell in Journal Of Econometrics Annals 1990-4, 46-1/2, pp. 201-212.

Kwon, H.-B. (2017). Exploring the predictive potential of artificial neural networks in conjunction with DEA in railroad performance modeling, International Journal of Production Economics, Vol. 183, pp. 159–170. DOI: 10.1016/j.ijpe.2016.10.022

Liang, L., & Wu, D. (2005). An application of pattern recognition on scoring Chinese corporations financial conditions based on backpropagation neural network. Computers and Operations Research, Vol. 32, pp. 1115–1129.

Maltseva, V., Na, J., Kim, G., & Ha, H.-K. (2020). Efficiency analysis of Russian rail freight transportation companies with super slack-based measurement data

envelopment analysis, *Journal of International Logistics and Trade*, 18(2), 77-89.
<https://doi.org/10.24006/jilt.2020.18.2.077>

Mazanec, J. (2018). Measuring Performance of the Freight Rail Transport Companies in the Slovak Republic, *East African Scholars Multidisciplinary Bulletin*, Vol. 1, Is. 3, pp. 45-50 DOI: 10.36349/easmb.2018.v01i03.002

McKinnon, A. (2015). Performance Measurement in Freight Transport: Its Contribution to the Design, Implementation and Monitoring of Public Policy, Prepared for the Roundtable on Logistics Development Strategies and their Performance Measurements

Moeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 18, pp. 435-44.

Nash, C. (1985). European railway comparisons what can we learn? in *International Railway Economics*, eds K. Button and D. Pit eld, Gower, Aldershot.

Oum, T. H. & Yu, C. (1994). Economic efficiency of railway systems and implications for public policy, *Journal of Transport Economic and Policy*, Vol. 28, No. 2, pp. 121-38.

Pazour, J. A., Meller, R. D., & Pohl, L. M. (2010). A Model to Design a National High-Speed Rail Network for Freight Distribution. *Transportation Research Part A*, 44, 119–135, DOI:10.1016/j.tra.2009.11.006

Pendharkar, P.C. & Rodger, J.A. (2003). Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption, *Decision Support Systems*, Vol. 36, pp. 117–136. DOI: 10.1016/S0167-9236(02)00138-0

Poitier, K., & Cho, S. (2011). Estimation of true efficient frontier of organisational performance using data envelopment analysis and support vector machine learning. *International Journal of Information and Decision Sciences.*, Vol. 3, pp. 148–172.

Ramón, N., Ruiz, J. L., & Sirvent, I. (2018). Two-step benchmarking: Setting more realistically achievable targets. *Expert Systems with Applications*, Vol. 92, pp. 124–131.

Ripley, B. D. (1993). Statistical aspects of neural networks. In *Networks and Chaos--Statistical and Probabilistic Aspects*, ed. O. E. Barndorff-Nielsen, J. L. Jensen, and W. S. Kendall, pp. 40-123.

Sameni, M. K., Preston, J., & Sameni, M. K. (2016). Evaluating efficiency of passenger railway stations: A DEA approach, *Research in Transportation Business & Management*, 20, 33-38, <http://dx.doi.org/10.1016/j.rtbm.2016.06.001>

Ståhlberg, A. (2023). Evaluation of efficiency and measures of improvement for Rail Freight Corridors in Europe, (Master's thesis, KTH Royal Institute of Technology)

Sueyoshi, T., Machida, H., Sugiyama, M., Arai, T., & Yamada, Y. (1997). Privatization of Japan national railways: DEA time series approaches. *Journal of the Operations Research Society of Japan*, Vol. 40, pp.202–205.

Sumo, P. D., Ji, X., & Cai, L. (2023). Performance prediction of a textile reverse logistics system using DEA and ANFIS hybrid models. *Journal of Intelligent & Fuzzy Systems*, (Preprint), pp. 1–11.

Thompson, L. S., & Bente, H. (2014). What is Rail Efficiency and How Can it Be Changed? *International Transport Forum Discussion Paper*, No. 2014-23.

Tomikawa, T., & Goto, M. (2022). Efficiency assessment of Japanese National Railways before and after privatization and divestiture using data envelopment analysis. *Transport Policy*, Vol, 118, pp. 44–55.

Wanke, P., & Azad, M. A. K. (2018). Efficiency in Asian railways: A comparison between data envelopment analysis approaches. *Transportation Planning and Technology*, Vol. 41, pp. 573–599.

Wetzel, H. (2008). European railway deregulation: the influence of regulatory and environmental conditions on efficiency, *Working Paper Series in Economics*, No. 86, Leuphana Universität Lüneburg, Institut für Volkswirtschaftslehre, Lüneburg.

Wu, C., Zhang, N., & Xu, L. (2021). Travelers on the railway: An economic growth model of the effects of railway transportation infrastructure on consumption and sustainable economic growth. *Sustainability*, Vol. 13, 6863.

Wu, D. (2009). Supplier selection: A hybrid model using DEA, decision tree and neural network, *Expert Systems with Applications*, Vol. 36, pp. 9105–9112. DOI: 10.1016/j.eswa.2008.12.039

Wu,D. (2006). Detecting information technology impact on firm performance using DEA and decision tree. *International Journal of Information Technology and Management*, Vol. 5, pp. 162–174.

Yu, M.-M. & Lin, E. T. J. (2007). Efficiency and effectiveness in railway performance using a multi-activity network DEA model, *Omega*, 36(6), 1005–1017, DOI: 10.1016/j.omega.2007.06.003

Zhang, H., Wang, X., Chen, L., Luo, Y., & Peng, S. (2022). Evaluation of the Operational Efficiency and Energy Efficiency of Rail Transit in China's Megacities Using a DEA Model. *Energies*, 15, 7758, <https://doi.org/10.3390/en15207758>

Zong, Y., Wu, J., Yu, K., & Yang, X. (2023). Efficiency benchmarking and its determinants in high-speed railways: Reference for China. *Research in Transportation Economics*, 102, 101355. <https://doi.org/10.1016/j.retrec.2023.101355>

7. Παραρτήματα

Το παρόν κεφάλαιο περιλαμβάνει τον πλήρη πηγαίο κώδικα που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας και συνοδεύει τη μεθοδολογική εφαρμογή του συνδυαστικού μοντέλου DEA–ANN. Η ενότητα αυτή αποσκοπεί στην ενίσχυση της διαφάνειας, της αναπαραγωγιμότητας και της τεχνικής πληρότητας της ερευνητικής διαδικασίας, επιτρέποντας στον αναγνώστη να κατανοήσει τη ροή της υλοποίησης, τη λογική των επιμέρους αλγορίθμων και τις επεξεργασίες που πραγματοποιήθηκαν στα δεδομένα.

Ο κώδικας έχει υλοποιηθεί στη γλώσσα Python με χρήση των ακόλουθων βιβλιοθηκών:

- **pandas** και **numpy** για επεξεργασία δεδομένων,
- **itertools** και **tqdm** για συνδυασμούς και εμφάνιση προόδου,
- **scipy.optimize.linprog** για επίλυση του DEA μοντέλου,
- **scikit-learn** για διαχωρισμό του dataset σε σύνολα εκπαίδευσης και ελέγχου,
- **tensorflow.keras** για κατασκευή, εκπαίδευση και φόρτωση του τεχνητού νευρωνικού δικτύου (ANN).

Οργανώνεται σε οκτώ (8) κύρια scripts. Κάθε script συνοδεύεται από τίτλο και παρατίθεται αυτούσιος ο πηγαίος κώδικας. Τα scripts έχουν διαχωριστεί σε δύο ομάδες, σύμφωνα με την επιλεγμένη εκροή στην ανάλυση αποδοτικότητας:

- Οικονομική αποδοτικότητα, με εκροή τα ετήσια έσοδα (Revenue)
- Τεχνική αποδοτικότητα, με εκροή το μεταφορικό έργο σε τονοχιλιόμετρα (Ton-kilometres)

Για κάθε εκδοχή (Revenue / Ton-kilometres), υλοποιούνται τα εξής βασικά βήματα:

1. Δημιουργία συνθετικού dataset μέσω συστηματικής μεταβολής εισροών
2. Υπολογισμός DEA αποδοτικότητας (θ) για κάθε εγγραφή του dataset
3. Εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου (ANN) με βάση τα αποτελέσματα DEA
4. Αξιολόγηση πολλαπλών προβλέψεων του ANN και επιλογή της βέλτιστης ως προς την αποδοτικότητα

Τα scripts είναι αριθμημένα και παρουσιάζονται με ενιαία μορφοποίηση, ώστε να διευκολύνεται η κατανόηση και η αναφορά τους σε άλλα σημεία της εργασίας.

- Script 1 – generate_systematic_sampled_half.py

```
import pandas as pd
import numpy as np
from itertools import product

# === STEP 1: Define Original Firm Data ===
firms = pd.DataFrame({
```

```

    "DMU": ["DB Cargo", "Xefafet", "Renfe", "CD Cargo", "OBB", "ZSSK",
"Hupac", "Green Cargo",
        "PKP Cargo", "VR", "RCH", "Mercitalia", "DB NL"],
    "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18000, 7606,
7606, 16000, 7000],
    "X2": [81536, 7677, 7200, 20411, 24973, 4289, 8500, 5170, 49757, 8360,
6000, 12380, 7000],
    "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460,
3000, 620],
    "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7,
42.3, 76.1, 51.1],
    "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82,
349.5, 475.7, 499.8, 149],
    "X6": [20.1, 1.99, 6.9, 7.22, 2.9, 6.5, 10.34, 2.28, 18.49, 5.5, 13.96,
10.2, 1.5],
    "Y2": [5582, 714, 198.81, 530, 1910, 303.6, 667, 365.4, 940.3, 339.9,
473, 480, 137]
})

# === STEP 2: Define Scaling Strategy ===
scales = np.arange(0.5, 1.01, 0.1) # 50% to 100% in 10% steps (6 values)
scale_combinations = list(product(scales, repeat=5)) # for X2-X6 → 6^5 =
7776 rows per firm

# === STEP 3: Generate Combinations for Each Firm ===
all_rows = []

for _, row in firms.iterrows():
    base_inputs = np.array([row["X2"], row["X3"], row["X4"], row["X5"],
row["X6"]])
    for scale_set in scale_combinations:
        scaled_inputs = base_inputs * np.array(scale_set)
        all_rows.append({
            "DMU": row["DMU"],
            "X1": row["X1"],
            "Y2": row["Y2"],
            "X2": scaled_inputs[0],
            "X3": scaled_inputs[1],
            "X4": scaled_inputs[2],
            "X5": scaled_inputs[3],
            "X6": scaled_inputs[4],
        })
df_full = pd.DataFrame(all_rows)

# === STEP 4: Sample Every 20th Row to Reduce Size ===

```



```
df_sampled_half = df_full.groupby("DMU", group_keys=False).apply(lambda x:
x.iloc[::20]).reset_index(drop=True)
```

```
# === STEP 5: Export Final Dataset ===
```

```
df_sampled_half.to_excel("systematic_dataset_all_firms_sampled_half.xlsx",
index=False)
```

```
print("☑ Dataset created and saved.")
```

▪ Script 2 – generate_systematic_sampled_half_BTK.py

```
import pandas as pd
```

```
import numpy as np
```

```
from itertools import product
```

```
# === STEP 1: Define Original Firm Data (BTK version) ===
```

```
firms_btk = pd.DataFrame({
    "DMU": [
        "DB Cargo", "Xefafet", "Renfe", "CD Cargo", "OBB", "ZSSK", "Hupac",
        "Green Cargo", "PKP Cargo", "VR", "RCH", "Mercitalia", "DB NL"
    ],
    "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18000, 7606,
7606, 16000, 7000],
    "X2": [81536, 7677, 7200, 20411, 24973, 2239, 8500, 5570, 49757, 8360,
6000, 12380, 7000],
    "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460,
3000, 620],
    "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7,
42.3, 76.1, 51.1],
    "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82,
349.5, 475.7, 499.8, 149],
    "X6": [20.1, 1.99, 6.9, 7.22, 2.9, 6.5, 10.34, 2.28, 18.49, 5.5, 13.96,
10.2, 1.5],
    "Y1": [74.5, 16.7, 4.42, 19.05, 20.7, 5.18, 13.3, 10.2, 22.3, 8.8, 5.85,
8.85, 6.26]
})
```

```
# === STEP 2: Define Scaling Strategy for X2-X6 ===
```

```
scales = np.arange(0.5, 1.01, 0.1)
```

```
scale_combinations = list(product(scales, repeat=5)) # For X2-X6
```

```
# === STEP 3: Generate Systematic Dataset ===
```

```
all_rows = []
```

```

for _, row in firms_btk.iterrows():
    base_inputs = np.array([row["X2"], row["X3"], row["X4"], row["X5"],
row["X6"]])
    for scale_set in scale_combinations:
        scaled_inputs = base_inputs * np.array(scale_set)
        all_rows.append({
            "DMU": row["DMU"],
            "X1": row["X1"],
            "Y1": row["Y1"],
            "X2": scaled_inputs[0],
            "X3": scaled_inputs[1],
            "X4": scaled_inputs[2],
            "X5": scaled_inputs[3],
            "X6": scaled_inputs[4],
        })

df_full_btk = pd.DataFrame(all_rows)

# === STEP 4: Sample Every 20th Row ===
df_sampled_btk = df_full_btk.groupby("DMU", group_keys=False).apply(lambda
x: x.iloc[::20]).reset_index(drop=True)

# === STEP 5: Export ===
df_sampled_btk.to_excel("systematic_dataset_all_firms_sampled_half_BTK.xlsx"
, index=False)

print("☑ File saved as
'systematic_dataset_all_firms_sampled_half_BTK.xlsx'")

```

▪ Script 3 – calculate_theta_sampled_half.py

```

import pandas as pd
import numpy as np
from scipy.optimize import linprog
from tqdm import tqdm

# Load the reduced dataset
df = pd.read_excel("systematic_dataset_all_firms_sampled_half.xlsx")

# DEA Efficiency Function (Input-oriented, CRS)
def dea_efficiency(inputs, outputs, eval_index):
    n, m_inputs = inputs.shape

```

```

c = np.zeros(n + 1)
c[0] = 1
A_ub, b_ub = [], []
for j in range(m_inputs):
    row = np.zeros(n + 1)
    row[0] = -inputs[eval_index, j]
    row[1:] = inputs[:, j]
    A_ub.append(row)
    b_ub.append(0)
A_eq, b_eq = [], []
for j in range(outputs.shape[1]):
    row = np.zeros(n + 1)
    row[1:] = -outputs[:, j]
    A_eq.append(row)
    b_eq.append(-outputs[eval_index, j])
bounds = [(0, 1)] + [(0, None)] * n
res = linprog(c, A_ub=A_ub, b_ub=b_ub, A_eq=A_eq, b_eq=b_eq,
bounds=bounds, method="highs")
return res.x[0] if res.success else np.nan

# Prepare inputs and outputs
inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
outputs = df[["Y2"]].values

# Calculate theta with progress bar
theta_scores = []
for i in tqdm(range(len(df)), desc="Calculating DEA Efficiency (Half
Sample)"):
    theta = dea_efficiency(inputs, outputs, i)
    theta_scores.append(theta)

# Save results
df["Theta"] = theta_scores
df.to_excel("systematic_dataset_all_firms_sampled_half_with_theta.xlsx",
index=False)
print("☑ DEA efficiency calculation complete and file saved.")

```

- Script 4 – calculate_theta_sampled_half_BTK.py

```

import pandas as pd
import numpy as np

```

```

from scipy.optimize import linprog
from tqdm import tqdm

# === STEP 1: Load Dataset ===
df = pd.read_excel("systematic_dataset_all_firms_sampled_half_BTK.xlsx")

# === STEP 2: Define DEA Function (Input-Oriented, CRS) ===
def dea_efficiency(inputs, outputs, eval_index):
    n, m_inputs = inputs.shape
    c = np.zeros(n + 1)
    c[0] = 1 # Objective: min  $\theta$ 

    A_ub = []
    b_ub = []

    for j in range(m_inputs):
        row = np.zeros(n + 1)
        row[0] = -inputs[eval_index, j]
        row[1:] = inputs[:, j]
        A_ub.append(row)
        b_ub.append(0)

    A_eq = []
    b_eq = []

    for j in range(outputs.shape[1]):
        row = np.zeros(n + 1)
        row[1:] = -outputs[:, j]
        A_eq.append(row)
        b_eq.append(-outputs[eval_index, j])

    bounds = [(0, 1)] + [(0, None)] * n

    res = linprog(c, A_ub=A_ub, b_ub=b_ub,
                  A_eq=A_eq, b_eq=b_eq, bounds=bounds,
                  method="highs")

    return res.x[0] if res.success else np.nan

# === STEP 3: Prepare Inputs and Outputs ===
inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
outputs = df[["Y1"]].values

```

```
# === STEP 4: Calculate  $\theta$  for Each Row ===
theta_scores = []
for i in tqdm(range(len(df)), desc="Calculating DEA Efficiency"):
    theta = dea_efficiency(inputs, outputs, i)
    theta_scores.append(theta)

# === STEP 5: Add  $\theta$  Column and Save ===
df["Theta"] = theta_scores
df.to_excel("systematic_dataset_all_firms_sampled_half_BTK_with_theta.xlsx",
index=False)

print("☑ DEA scores calculated and saved to
'systematic_dataset_all_firms_sampled_half_BTK_with_theta.xlsx'")
```

▪ Script 5 – train_ann_with_theta.py

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.optimizers import Adam

# Load the DEA-evaluated dataset
df =
pd.read_excel("systematic_dataset_all_firms_sampled_half_with_theta.xlsx")

# Drop any rows with missing theta (NaN)
df = df.dropna(subset=["Theta"])

# Define input (X1, Y2, Theta) and output (X2-X6)
X = df[["X1", "Y2", "Theta"]].values
y = df[["X2", "X3", "X4", "X5", "X6"]].values

# Split the dataset
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# Build the ANN model
model = Sequential()
model.add(Dense(64, input_dim=3, activation='relu'))
```

```

model.add(Dense(64, activation='relu'))
model.add(Dense(5)) # 5 outputs for X2-X6

# Compile the model
model.compile(optimizer=Adam(learning_rate=0.001), loss='mse')

# Train the model
model.fit(X_train, y_train, epochs=500, batch_size=16, validation_split=0.1)

# Save the trained model
model.save("ann_theta_optimizer_model.h5")

# Evaluate the model
loss = model.evaluate(X_test, y_test)
print(f"☑ Model trained. Test MSE Loss: {loss}")

```

- Script 6 – train_ann_with_theta_BTK.py

```

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.optimizers import Adam

# === STEP 1: Load Dataset ===
df =
pd.read_excel("systematic_dataset_all_firms_sampled_half_BTK_with_theta.xlsx")
df = df.dropna(subset=["Theta"])

# === STEP 2: Define Inputs and Outputs ===
X = df[["X1", "Y1", "Theta"]].values
y = df[["X2", "X3", "X4", "X5", "X6"]].values

# === STEP 3: Split Dataset ===
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# === STEP 4: Build ANN Model ===
model = Sequential()

```

```

model.add(Dense(64, input_dim=3, activation='relu'))
model.add(Dense(64, activation='relu'))
model.add(Dense(5)) # Outputs: X2-X6

# === STEP 5: Compile and Train ===
model.compile(optimizer=Adam(learning_rate=0.001), loss='mse')
model.fit(X_train, y_train, epochs=500, batch_size=16, validation_split=0.1)

# === STEP 6: Save Model ===
model.save("ann_theta_optimizer_model_BTK.h5")

# === STEP 7: Evaluate Model ===
loss = model.evaluate(X_test, y_test)
print(f"☑ ANN for BTK trained. Test MSE Loss: {loss}")

```

- Script 7 – eval_ann_with_theta_model_REVENUE_FINAL.py

```

import pandas as pd
import numpy as np
from tensorflow.keras.models import load_model
from scipy.optimize import linprog

# DEA Efficiency Function
def dea_efficiency(inputs, outputs, eval_index):
    n, m_inputs = inputs.shape
    c = np.zeros(n + 1)
    c[0] = 1
    A_ub, b_ub = [], []
    for j in range(m_inputs):
        row = np.zeros(n + 1)
        row[0] = -inputs[eval_index, j]
        row[1:] = inputs[:, j]
        A_ub.append(row)
        b_ub.append(0)
    A_eq, b_eq = [], []
    for j in range(outputs.shape[1]):
        row = np.zeros(n + 1)
        row[1:] = -outputs[:, j]
        A_eq.append(row)
        b_eq.append(-outputs[eval_index, j])

```

```

bounds = [(0, 1)] + [(0, None)] * n
res = linprog(c, A_ub, b_ub, A_eq, b_eq, bounds, method="highs")
return res.x[0] if res.success else np.nan

# STEP 1: Load Correct Firm Data (Revenue version with user provided
corrections)
df = pd.DataFrame({
    "DMU": [
        "DB Cargo", "Xefafet", "Renfe", "CD Cargo", "OBB", "ZSSK", "Hupac",
        "Green Cargo", "PKP Cargo", "VR", "Rail Cargo Hungaria",
        "Mercitalia", "DB NL"
    ],
    "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18800, 9200,
           7606, 16800, 7000],
    "X2": [81536, 7677, 7200, 20411, 24973, 2239, 8500, 5570, 49757, 8360,
           2600, 12380, 7000],
    "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460,
           3000, 620],
    "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7,
           42.3, 76.1, 51.1],
    "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82,
           349.5, 475.7, 499.8, 149],
    "X6": [20.1, 1.99, 6.9, 7.22, 2.9, 6.5, 10.34, 2.28, 18.49, 5.5, 13.96,
           10.2, 1.5],
    "Y2": [5582, 714, 198.81, 530, 1910, 303.6, 667, 365.4, 940.3, 339.9,
           473, 480, 137],
    "Theta_DEAP": [1.000, 1.000, 0.885, 1.000, 1.000, 1.000, 1.000, 0.958,
                   0.963, 0.899, 1.000, 0.878, 0.868]
})

# STEP 2: Load Trained Revenue Model
model = load_model("ann_theta_optimizer_model.h5", compile=False)

X1_vals = df["X1"].values
Y2_vals = df["Y2"].values
theta_deap = df["Theta_DEAP"].values
original_inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
outputs = df[["Y2"]].values

# STEP 3: Predict Optimized Inputs and Evaluate Multi-Try
final_inputs = original_inputs.copy()
theta_ann_best = []

for i in range(len(df)):
    if theta_deap[i] == 1.0:

```



```

        theta_ann_best.append(1.0)
        continue

    best_theta = -1
    best_input = None
    theta_range = np.linspace(theta_deap[i], 1.0, num=10)

    for theta_target in theta_range:
        input_ann = np.array([[X1_vals[i], Y2_vals[i], theta_target]])
        pred = model.predict(input_ann, verbose=0)[0]
        base = original_inputs[i, 1:]
        clipped = np.clip(pred, base * 0.5, base)
        test_input = final_inputs.copy()
        test_input[i] = np.concatenate([[X1_vals[i]], clipped])
        score = dea_efficiency(test_input, outputs, i)
        if not np.isnan(score) and score > best_theta:
            best_theta = score
            best_input = clipped

    if best_input is not None:
        final_inputs[i] = np.concatenate([[X1_vals[i]], best_input])
    theta_ann_best.append(best_theta)

# STEP 4: Save Final Output
df["Theta_ANN_best"] = theta_ann_best
for j, col in enumerate(["X2_best", "X3_best", "X4_best", "X5_best",
                        "X6_best"]):
    df[col] = final_inputs[:, j + 1]
df["Improvement_%"] = ((df["Theta_ANN_best"] - df["Theta_DEAP"]) /
df["Theta_DEAP"]) * 100

df.to_excel("ann_theta_prediction_results_REVENUE_FINAL.xlsx", index=False)
print("☑ ANN optimization + DEA evaluation (Revenue FINAL) completed and
saved.")

```

- Script 8 – eval_ann_with_theta_model_BTK_FINAL.py

```

import pandas as pd
import numpy as np
from tensorflow.keras.models import load_model
from scipy.optimize import linprog

```

```

# DEA Efficiency Function
def dea_efficiency(inputs, outputs, eval_index):
    n, m_inputs = inputs.shape
    c = np.zeros(n + 1)
    c[0] = 1
    A_ub, b_ub = [], []
    for j in range(m_inputs):
        row = np.zeros(n + 1)
        row[0] = -inputs[eval_index, j]
        row[1:] = inputs[:, j]
        A_ub.append(row)
        b_ub.append(0)
    A_eq, b_eq = [], []
    for j in range(outputs.shape[1]):
        row = np.zeros(n + 1)
        row[1:] = -outputs[:, j]
        A_eq.append(row)
        b_eq.append(-outputs[eval_index, j])
    bounds = [(0, 1)] + [(0, None)] * n
    res = linprog(c, A_ub, b_ub, A_eq, b_eq, bounds, method="highs")
    return res.x[0] if res.success else np.nan

# STEP 1: Load Correct Firm Data (BTK version with manual Theta_DEAP)
df = pd.DataFrame({
    "DMU": [
        "DB Cargo", "Xefafet", "Renfe", "CD Cargo", "OBB", "ZSSK", "Hupac",
        "Green Cargo", "PKP Cargo", "VR", "Rail Cargo Hungaria",
        "Mercitalia", "DB NL"
    ],
    "X1": [33400, 28000, 15500, 9500, 5000, 3600, 5300, 10900, 18800, 9200,
           7606, 16800, 7000],
    "X2": [81536, 7677, 7200, 20411, 24973, 2239, 8500, 5570, 49757, 8360,
           2600, 12380, 7000],
    "X3": [31350, 5000, 945, 6628, 5912, 3771, 700, 1800, 14062, 2800, 1460,
           3000, 620],
    "X4": [319.0, 105.0, 232.7, 180.0, 170.5, 69.8, 35.0, 36.5, 349.2, 82.7,
           42.3, 76.1, 51.1],
    "X5": [6079, 630, 210.89, 488.5, 1920, 301.8, 668.44, 369.2, 897.82,
           349.5, 475.7, 499.8, 149],
    "X6": [20.1, 1.99, 6.9, 7.22, 2.9, 6.5, 10.34, 2.28, 18.49, 5.5, 13.96,
           10.2, 1.5],
    "Y1": [74.5, 16.7, 4.42, 19.05, 20.7, 5.18, 13.3, 10.2, 22.3, 8.8, 5.85,
           8.85, 6.26],

```

```

    "Theta_DEAP": [1.000, 1.000, 0.579, 1.000, 1.000, 1.000, 1.000, 1.000,
0.634, 0.802, 1.000, 0.575, 1.000]
})

# STEP 2: Load Trained BTK Model
model = load_model("ann_theta_optimizer_model_BTK.h5", compile=False)

X1_vals = df["X1"].values
Y1_vals = df["Y1"].values
theta_deap = df["Theta_DEAP"].values
original_inputs = df[["X1", "X2", "X3", "X4", "X5", "X6"]].values
outputs = df[["Y1"]].values

# STEP 3: Predict Optimized Inputs and Evaluate Multi-Try
final_inputs = original_inputs.copy()
theta_ann_best = []

for i in range(len(df)):
    if theta_deap[i] == 1.0:
        theta_ann_best.append(1.0)
        continue

    best_theta = -1
    best_input = None
    theta_range = np.linspace(theta_deap[i], 1.0, num=10)

    for theta_target in theta_range:
        input_ann = np.array([[X1_vals[i], Y1_vals[i], theta_target]])
        pred = model.predict(input_ann, verbose=0)[0]
        base = original_inputs[i, 1:]
        clipped = np.clip(pred, base * 0.5, base)
        test_input = final_inputs.copy()
        test_input[i] = np.concatenate([[X1_vals[i]], clipped])
        score = dea_efficiency(test_input, outputs, i)
        if not np.isnan(score) and score > best_theta:
            best_theta = score
            best_input = clipped

    if best_input is not None:
        final_inputs[i] = np.concatenate([[X1_vals[i]], best_input])
    theta_ann_best.append(best_theta)

```

```
# STEP 4: Save Final Output
df["Theta_ANN_best"] = theta_ann_best
for j, col in enumerate(["X2_best", "X3_best", "X4_best", "X5_best",
                        "X6_best"]):
    df[col] = final_inputs[:, j + 1]

df["Improvement_%"] = ((df["Theta_ANN_best"] - df["Theta_DEAP"]) /
                        df["Theta_DEAP"]) * 100

df.to_excel("ann_theta_prediction_results_BTK_FINAL.xlsx", index=False)
print("☑ ANN optimization + DEA evaluation (BTK FINAL) completed and
saved.")
```