

Αύγουστος 2023



**ΣΧΟΛΗ : ΟΙΚΟΝΟΜΙΚΩΝ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ**

ΤΜΗΜΑ : ΠΛΗΡΟΦΟΡΙΚΗΣ

**Ο ΡΟΛΟΣ ΤΩΝ ΜΕΓΑΛΩΝ ΔΕΔΟΜΕΝΩΝ (BIG
DATA) ΚΑΙ ΤΗΣ ΠΡΟΓΝΩΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ
(PREDICTIVE ANALYTICS) ΣΤΗ ΒΙΟΜΗΧΑΝΙΑ
ΑΛΟΥΜΙΝΙΟΥ**

ΚΥΡΛΗΣ ΙΑΣΩΝ

ΑΥΓΟΥΣΤΟΣ/2023



**ΣΧΟΛΗ : ΟΙΚΟΝΟΜΙΚΩΝ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ**

ΤΜΗΜΑ: ΠΛΗΡΟΦΟΡΙΚΗΣ

**Ο ΡΟΛΟΣ ΤΩΝ ΜΕΓΑΛΩΝ ΔΕΔΟΜΕΝΩΝ (BIG
DATA) ΚΑΙ ΤΗΣ ΠΡΟΓΝΩΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ
(PREDICTIVE ANALYTICS) ΣΤΗ ΒΙΟΜΗΧΑΝΙΑ
ΑΛΟΥΜΙΝΙΟΥ**

**Διατριβή η οποία υποβλήθηκε προς απόκτηση εξ
αποστάσεως μεταπτυχιακού τίτλου σπουδών στα
Πληροφοριακά Συστήματα και τη Ψηφιακή Καινοτομία στο
Πανεπιστήμιο Νεάπολις**

ΚΥΡΛΗΣ ΙΑΣΩΝ

ΑΥΓΟΥΣΤΟΣ/2023

Πνευματικά δικαιώματα

Copyright © Όνομα επίθετο φοιτητή, έτος κατάθεσης διατριβής

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της διατριβής από το Πανεπιστημίου Νεάπολις δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Πανεπιστημίου.

Ή ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Ο Κυρλής Ιάσων γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα ότι η παρούσα εργασία με τίτλο «Ο ρόλος των μεγάλων δεδομένων (big data) και της προγνωστικής ανάλυσης (Predictive analytics) στη βιομηχανία αλουμινίου », αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές που έχω χρησιμοποιήσει, έχουν δηλωθεί κατάλληλα στις βιβλιογραφικές παραπομπές και αναφορές. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Ο Δηλών

Κυρλής Ιάσων



Περιεχόμενα

Περίληψη	1
Abstract.....	3
Κεφάλαιο 1 – Εισαγωγή	4
1.1 Χρήση των Μεγάλων Δεδομένων σε μεγάλες επιχειρήσεις.....	4
1.2 Τα Μεγάλα Δεδομένα στον κλάδο των βιομηχανιών αλουμινίου.....	6
1.3 Σκοπός και επιμέρους στόχοι.....	7
1.4 Βασικά Ερευνητικά Ερωτήματα.....	8
1.5 Αναμενόμενα Αποτελέσματα	9
Κεφάλαιο 2 - Βιβλιογραφική Ανασκόπηση.....	10
2.1.Ορισμός.....	10
2.2 Ιστορική αναδρομή	18
2.3 Αναλυτική (Analytics).....	20
2.4. Βιβλιογραφική Ανασκόπηση στις βιομηχανίες	22
2.4.1. Βιομηχανία παραγωγής.....	22
2.4.2. Βιομηχανία αλουμινίου	23
Κεφάλαιο 3 - Μεθοδολογία Έρευνας	26
3.1. Εισαγωγή	26
3.1.1 Συλλογή δεδομένων.....	26
3.1.2. Μεταβλητές.....	26
3.1.2. Προεπεξεργασία δεδομένων	28
3.2. Στατιστικές και ML προσεγγίσεις	29
3.3. Δενδρολογικές Μέθοδοι: Decision Trees and Forests.....	29
3.4. Διαχωρισμός και Μετασχηματισμός Δεδομένων (Splitting and Scalling).....	30
3.5. Η Αντιστάθμιση Μεροληψίας-Διακύμανσης (Bias-Variance Tradeoff).....	31
3.6. Μέθοδοι Ενίσχυσης και Ομαδοποίησης (Boosting and Ensemble Methods)	31
Κεφάλαιο 4 – Ανάλυση και Αποτελέσματα	35

4.1. Εισαγωγή	35
4.2. Ανάλυση Δεδομένων	35
Κεφάλαιο 5 - Συζήτηση αποτελεσμάτων και συμπεράσματα	46
5.1 Ερμηνεία αποτελεσμάτων.....	46
5.2. Συμπεράσματα	47
Βιβλιογραφία	50

Κατάλογος Γραφικών Παραστάσεων/Εικόνων/Διαγραμμάτων

Εικόνα 1. Τα 5V's των "Big Data"	11
Εικόνα 2. Όγκος δεδομένων (Volume).....	12
Εικόνα 3. Ετήσιο Μέγεθος της Παγκόσμιας Σφαίρας Δεδομένων.....	13
Εικόνα 4. Ταχύτητα δεδομένων (Velocity)	15
Εικόνα 5. Αξία (Value).....	17
Διάγραμμα 4.1. Ιστογράμματα και εκτιμήσεις (KDEs) των μεταβλητών-στόχων πρόβλεψης στη βάση δεδομένων του Αλουμινίου	36
Διάγραμμα 4.2.Συσχετίσεις μεταξύ των μεταβλητών του συνόλου δεδομένων.	37
Διάγραμμα 4.3. Scatter plot γραμμικής παλινδρόμησης OLS.....	38
Διάγραμμα 4.4. Doughnut plot για το μοντέλο γραμμικής παλινδρόμησης.....	39
Διάγραμμα 4.5. Σημαντικότητα ανεξάρτητων μεταβλητών που δημιουργούνται από τον αλγόριθμο δέντρου αποφάσεων RandomForest	40
Διάγραμμα 4.6. Doughnut plot για το μοντέλο RandomForest.....	41
Διάγραμμα 4.7. Σημαντικότητα ανεξάρτητων μεταβλητών που δημιουργούνται από την μέθοδο XGB	42
Διάγραμμα 4.8. Σημαντικότητα ανεξάρτητων μεταβλητών που δημιουργούνται από την μέθοδο ExtraTrees.	43
Διάγραμμα 4.9. Σημαντικότητα ανεξάρτητων μεταβλητών που δημιουργούνται από την μέθοδο LightGBM.	44
Διάγραμμα 4.10.Scatterplots και των 6 μοντέλων πρόβλεψης για το σύνολο δεδομένων του αλουμινίου.	45

Κατάλογος Πινάκων

Πίνακας 4.1. Περιγραφικά μέτρα των 5 μεταβλητών.....	35
--	----

Ονοματεπώνυμο Φοιτητή/Φοιτήτριας:

Τίτλος Μεταπτυχιακής Διατριβής:

Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση εξ αποστάσεως μεταπτυχιακού τίτλου στο Πανεπιστήμιο Νεάπολις και εγκρίθηκε στις [ημερομηνία έγκρισης] από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή:

Πρώτος επιβλέπων (Πανεπιστήμιο Νεάπολις Πάφος).....[ονοματεπώνυμο, βαθμίδα, υπογραφή]

Μέλος Εξεταστικής Επιτροπής:[ονοματεπώνυμο, βαθμίδα, υπογραφή]

Μέλος Εξεταστικής Επιτροπής:[ονοματεπώνυμο, βαθμίδα, υπογραφή]

Ευχαριστίες

Ευχαριστώ θερμά τον επιβλέποντα καθηγητή μου κ. Ζαγόρη Κωνσταντίνο για την κατανόηση του κατά την εκπόνηση της παρούσας μελέτης και την στήριξή του. Επίσης, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους ανθρώπους που ήταν δίπλα μου για την ηθική υποστήριξη που μου προσέφεραν. Ιδιαίτερως, θα ήθελα να ευχαριστήσω τον καθηγητή Χαράλαμπο Αγκυρόπουλο για τις ακαδημαϊκές προτάσεις του, τις υποδείξεις και συμβουλές, καθώς και τις πολύτιμες γνώσεις που μου προσέφερε στην στατιστική ανάλυση των δεδομένων, τις οποίες μοιράστηκε απλόχερα μαζί μου.

Αφιέρωση

Στην μνήμη του πατέρα μου

Περίληψη

Στο σημερινό ταχέως εξελισσόμενο επιχειρηματικό τοπίο, οι οργανισμοί σε όλους τους κλάδους αξιοποιούν τη δύναμη των δεδομένων για να προωθήσουν τη λήψη τεκμηριωμένων αποφάσεων και να αποκτήσουν ανταγωνιστικό πλεονέκτημα. Η βιομηχανία αλουμινίου, με τις περίπλοκες λειτουργίες και τη δυναμική αλυσίδα εφοδιασμού της, δεν αποτελεί εξαίρεση. Υιοθετώντας τα μεγάλα δεδομένα και την προγνωστική ανάλυση, ο κλάδος αυτός έχει τη δυνατότητα να φέρει επανάσταση στις διαδικασίες του, να αυξήσει τη λειτουργική αποδοτικότητα και να ξεκλειδώσει νέες ευκαιρίες.

Ένα από τα βασικά οφέλη των μεγάλων δεδομένων και της προγνωστικής ανάλυσης στη βιομηχανία είναι η βελτιστοποίηση των διαδικασιών. Αναλύοντας τεράστιες ποσότητες δεδομένων, οι εταιρείες μπορούν να εντοπίσουν μοτίβα και τάσεις που αποκαλύπτουν ανεπάρκειες στις δραστηριότητές τους. Αυτή η διορατικότητα τους επιτρέπει να εξορθολογίσουν τις ροές εργασίας παραγωγής, να μειώσουν τη σπατάλη και να βελτιώσουν τη συνολική παραγωγικότητα. Για παράδειγμα, αναλύοντας ιστορικά δεδομένα και πληροφορίες αισθητήρων σε πραγματικό χρόνο, οι οργανισμοί μπορούν να βελτιστοποιήσουν τις παραμέτρους των κλιβάνων, τις συνθέσεις κραμάτων και τις τεχνικές χύτευσης, με αποτέλεσμα να παράγουν προϊόντα υψηλότερης ποιότητας με μειωμένη κατανάλωση ενέργειας.

Η βιομηχανία αλουμινίου βασίζεται σε μια παγκόσμια και διασυνδεδεμένη αλυσίδα εφοδιασμού και τα μεγάλα δεδομένα σε συνδυασμό με την προγνωστική ανάλυση μπορούν να ενισχύσουν τη διαχείρισή της. Η ορατότητα σε πραγματικό χρόνο για τα επίπεδα αποθεμάτων, τα πρότυπα ζήτησης και τις αλυσίδες μεταφοράς και αποθήκευσης επιτρέπει στους οργανισμούς να βελτιστοποιούν τα επίπεδα αποθεμάτων, να προβλέπουν τις διακυμάνσεις της ζήτησης και να λαμβάνουν αποφάσεις βάσει δεδομένων σχετικά με την προμήθεια και τη διανομή. Κάτι τέτοιο έχει ως αποτέλεσμα τη βελτίωση της επιχειρησιακής απόδοσης, τη μείωση του κόστους και τη βελτίωση της ικανοποίησης των πελατών.

Η βιωσιμότητα αποτελεί αυξανόμενη ανησυχία στη βιομηχανία αλουμινίου. Τα μεγάλα δεδομένα και η προγνωστική ανάλυση μπορούν να συμβάλουν καθοριστικά στις πρωτοβουλίες και προοπτικές βιωσιμότητας κάθε είδους επιχείρησης. Με την ανάλυση δεδομένων αναφορικά για παράδειγμα με την κατανάλωση ενέργειας και τις παραμέτρους παραγωγής, οι επιχειρήσεις μπορούν να εντοπίσουν περιοχές αναποτελεσματικότητας και να εφαρμόσουν στοχευμένες βελτιώσεις, οδηγώντας σε εξοικονόμηση της ενέργειας και

μείωση των αποβλήτων. Η προγνωστική ανάλυση μπορεί επίσης να προβλέψει τη ζήτηση ενέργειας και να επιτρέψει τη δυναμική εξισορρόπηση φορτίου, βελτιστοποιώντας την κατανάλωση σε όλη τη διαδικασία παραγωγής.

Λέξεις κλειδιά: Μεγάλα Δεδομένα, προγνωστική ανάλυση, βιομηχανία αλουμινίου, παραγωγή, διαχείριση αποθεμάτων, βιωσιμότητα.

Abstract

In today's rapidly evolving business landscape, organisations across all industries are harnessing the power of data to drive informed decision making and gain competitive advantage. The aluminum industry, with its complex operations and dynamic supply chain, is no exception. By adopting big data and predictive analytics, this industry has the potential to revolutionise its processes, increase operational efficiency and unlock new opportunities.

One of the key benefits of big data and predictive analytics in the industry is process optimization. By analyzing vast amounts of data, companies can identify patterns and trends that reveal inefficiencies in their operations. This insight allows them to streamline production workflows, reduce waste and improve overall productivity.

The aluminium industry relies on a global and interconnected supply chain and big data combined with predictive analytics can enhance its management. Real-time visibility into inventory levels, demand patterns and transportation logistics allows organizations to optimize inventory levels, anticipate demand fluctuations and make data-driven decisions on supply and distribution. This results in improved operational efficiency, reduced costs and improved customer satisfaction.

Sustainability is a growing concern in the aluminum industry, and big data and predictive analytics can contribute to sustainability initiatives. By analyzing data, indicatively on energy consumption and production parameters, organizations can identify areas of inefficiency and implement targeted improvements, leading to energy optimization and waste reduction. Predictive analytics can also predict demand, energy demand and enable dynamic load balancing, optimising energy use throughout the production process.

Keywords: Big Data, predictive analytics, aluminum industry, production, demand, logistics, sustainability.