



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
“ΣΥΣΤΗΜΑΤΑ ΑΥΤΟΜΑΤΙΣΜΟΥ”

Παρακολούθηση Κατεργασιών Τόρνευσης με Τεχνικές Μηχανικής Μάθησης

Διπλωματική Εργασία

Χαράλαμπος Καψής

Αριθμός Μητρώου : 02123101

Επιβλέπων Καθηγητής : Γεώργιος-Χριστόφορος Βοσνιάκος

Ιούνιος 2025

Περιεχόμενα

Περίληψη.....	4
Abstract.....	5
Ευχαριστίες.....	6
1 Εισαγωγή.....	7
1.1 Δομή της εργασίας.....	8
1.2 Αντικείμενο και σκοπός της εργασίας.....	9
1.3 Διάγραμμα Ροής.....	10
2 Ανασκόπηση Βιβλιογραφίας.....	12
2.1 Βασικές Αρχές Κατεργασίας Τόρνευσης.....	12
2.2 Τεχνητή Νοημοσύνη στη Βιομηχανία.....	13
2.3 Ποιότητα Προϊόντων και Ταξινόμηση τους.....	15
2.4 Χαρακτηριστικά ποιότητας.....	15
2.5 Επίδραση παραμέτρων κατεργασίας.....	16
2.6 Διεθνή πρότυπα και κριτήρια ποιότητας στην τόρνευση.....	16
3 Εξοπλισμός και δεδομένα.....	18
3.1 Περιγραφή πειραματικού εξοπλισμού τόρνευσης.....	18
3.2 Συλλογή και Αρχική Επεξεργασία Σημάτων.....	20
3.2.1 Χαρακτηριστικά Χρόνου.....	21
3.2.2 Χαρακτηριστικά Συχνότητας.....	21
3.3 Αποθορυβοποίηση των σημάτων.....	22
3.4 Εξαγωγή Χαρακτηριστικών (Features).....	22
3.5 Labels.....	26
4 Μοντέλα Μηχανικής Μάθησης.....	29
4.1 Random Forest Classifier - RFC.....	29
4.2 Multi-Layer Perceptron -MLP.....	30
4.3 Υλοποιήσεις.....	32
4.4 Προεπεξεργασία και Προετοιμασία Δεδομένων.....	32
4.5 1 ^η Υλοποίηση (Imbalanced Data):.....	34
4.6 2 ^η Υλοποίηση (Manually Balanced Data):.....	36

4.7	3 ^η Υλοποίηση (Νέο Feature Extraction):.....	36
5	Αποτελέσματα και Αξιολόγηση.....	38
5.1	Εκπαίδευση & αξιολόγηση στο Αρχικό Datasheet (Imbalanced Data).....	38
5.1.1	Επιλεγμένα Χαρακτηριστικά μέσω Feature Selection.....	38
5.1.2	Random Forest Classifier :.....	39
5.1.3	Multi-Layer Perceptron :.....	39
5.2	Εκπαίδευση & αξιολόγηση με manual Αντιμετώπιση του Class Imbalance.....	42
5.2.1	Random Forest Classifier :.....	42
5.2.2	Multi-Layer Perceptron :.....	42
5.3	Εκπαίδευση & αξιολόγηση με εκ νέου feature extraction.....	45
5.3.1	Random Forest Classifier :.....	46
5.3.2	Multi-Layer Perceptron :.....	48
5.4	Σύγκριση μοντέλων και ερμηνεία ευρημάτων.....	49
6	Συμπεράσματα.....	52
6.1	Περιορισμοί της Μελέτης.....	52
6.2	Μελλοντική Εργασία.....	53
7	Βιβλιογραφία.....	55

Περίληψη

Η 4η Βιομηχανική Επανάσταση (Industry 4.0) έχει αναδείξει την ανάγκη ενσωμάτωσης έξυπνων τεχνολογιών στην παραγωγική διαδικασία, με στόχο την αυτοματοποίηση, τη βελτιστοποίηση της απόδοσης και την ενίσχυση της ποιότητας των παραγόμενων προϊόντων. Σε αυτό το πλαίσιο, η παρούσα εργασία εστιάζει στην ανάπτυξη και αξιολόγηση ενός ευφυούς συστήματος παρακολούθησης κατεργασιών τóρνευσης, βασισμένου σε τεχνικές μηχανικής μάθησης και ανάλυσης σημάτων, με σκοπό την αυτόματη ταξινόμηση της ποιότητας των κατεργασμένων τεμαχίων.

Η μεθοδολογία που ακολουθήθηκε περιλαμβάνει την πειραματική κατεργασία σε CNC τóρνο κατά την οποία συλλέχθηκαν σήματα από τον ελεγκτή της εργαλειομηχανής (ταχύτητα και ροπή στους άξονες X, Z και στην άτρακτο) καθώς και από επιταχυνσιόμετρα τοποθετημένα στους ίδιους άξονες. Η συλλογή των σημάτων πραγματοποιήθηκε μέσω συστήματος DAQ με ρυθμό δειγματοληψίας 1 kHz, διασφαλίζοντας ακρίβεια και συγχρονισμό. Τα σήματα υποβλήθηκαν σε προκαταρκτική επεξεργασία, αποθορυβοποίηση με χρήση κυματιδικής ανάλυσης (Wavelet Denoising – Coiflet2), και εξαγωγή χαρακτηριστικών (feature extraction) τόσο από το πεδίο του χρόνου (RMS, Energy, Kurtosis, Skewness, ZCR, Number of Peaks, Envelope RMS, Envelope Energy, Entropy) όσο και από το πεδίο της συχνότητας (Dominant Frequency, Spectral Centroid, Total Spectrum Power). Τα εξαγόμενα χαρακτηριστικά αποτέλεσαν τα δεδομένα εισόδου για την εκπαίδευση αλγορίθμων μηχανικής μάθησης. Η ποιότητα των τεμαχίων κατηγοριοποιήθηκε σε τρεις κλάσεις (Low, Medium, High), με βάση μετρολογικά χαρακτηριστικά όπως η τραχύτητα επιφάνειας (Ra, Rz), η διαστατική ακρίβεια και το βάθος κοπής. Από αυτά, ως κύρια μεταβλητή-στόχος (label) επιλέχθηκε το Roughness_Ra_Class. Για την ταξινόμηση εφαρμόστηκαν δύο βασικά μοντέλα: ο Random Forest Classifier και ένα Multi-Layer Perceptron (MLP). Πραγματοποιήθηκαν πολλαπλές υλοποιήσεις: (α) με αρχικά μη εξισορροπημένα δεδομένα, (β) με χειροκίνητη εξισορρόπηση (undersampling/oversampling) και (γ) με εμπλουτισμένο σύνολο χαρακτηριστικών. Εφαρμόστηκαν τεχνικές επιλογής χαρακτηριστικών (feature selection) για τη βελτιστοποίηση της απόδοσης. Εφαρμόστηκαν αρκετές διαφορετικές υλοποιήσεις (διαφορετικά μοντέλα, χρήση τεχνικής SMOTE για την αντιμετώπιση της ανισορροπίας μεταξύ των κλάσεων, tuning μοντέλων κλπ.) ούτως ώστε να επιτευχθεί ένα καλύτερο αποτέλεσμα.

Τα αποτελέσματα έδειξαν ότι η προσεκτική επιλογή και εξαγωγή των χαρακτηριστικών, σε συνδυασμό με κατάλληλη προεπεξεργασία και εξισορρόπηση του συνόλου δεδομένων, βελτιώνει σημαντικά την ικανότητα των μοντέλων να προβλέπουν με ακρίβεια την ποιότητα των τεμαχίων. Συμπερασματικά, η εργασία αποδεικνύει ότι η αξιοποίηση σημάτων δόνησης, ταχύτητας και ροπής από τον τóρνο, μέσω τεχνικών μηχανικής μάθησης, επιτρέπει την αποτελεσματική εκτίμηση της ποιότητας κατεργασμένων τεμαχίων. Η προσέγγιση αυτή μπορεί να ενσωματωθεί σε σύγχρονες βιομηχανικές γραμμές παραγωγής, προσφέροντας βελτιωμένο έλεγχο ποιότητας, μείωση αποβλήτων, και ενίσχυση της αυτονομίας του εξοπλισμού – πλήρως εναρμονισμένη με τις αρχές του Industry 4.0.

Abstract

The Fourth Industrial Revolution (Industry 4.0) has highlighted the need to integrate intelligent technologies into the production process, aiming at automation, performance optimization, and enhancement of product quality. Within this context, the present work focuses on the development and evaluation of an intelligent monitoring system for turning operations, based on machine learning techniques and signal analysis, with the goal of automatically classifying the quality of the machined parts.

The adopted methodology includes experimental machining on a CNC lathe, during which signals were collected from the machine tool controller (speed and torque along the X, Z axes and the spindle), as well as from accelerometers mounted on the same axes. Signal acquisition was carried out through a DAQ system with a sampling rate of 1 kHz, ensuring accuracy and synchronization. The signals underwent preliminary processing, including denoising using wavelet analysis (Wavelet Denoising – Coiflet2), and feature extraction from both the time domain (RMS, Energy, Kurtosis, Skewness, ZCR, Number of Peaks, Envelope RMS, Envelope Energy, Entropy) and the frequency domain (Dominant Frequency, Spectral Centroid, Total Spectrum Power). The extracted features were used as input data for training machine learning algorithms.

The quality of the machined parts was categorized into three classes (Low, Medium, High), based on metrological parameters such as surface roughness (Ra, Rz), dimensional accuracy, and cutting depth. Among these, the main target variable (label) selected was the *Roughness_Ra_Class*. Two core classification models were implemented: the Random Forest Classifier and a Multi-Layer Perceptron (MLP). Multiple implementations were carried out: (a) using the initial imbalanced dataset, (b) applying manual balancing (undersampling/oversampling), and (c) using an enriched feature set. Feature selection techniques were applied to optimize performance. Several alternative configurations were also tested (different models, application of the SMOTE technique to address class imbalance, model tuning, etc.) in order to achieve better results.

The results demonstrated that careful feature selection and extraction, combined with appropriate preprocessing and dataset balancing, significantly improve the models' ability to accurately predict the quality of the machined parts. In conclusion, this study confirms that utilizing vibration, speed, and torque signals from the lathe through machine learning techniques enables effective estimation of the quality of machined components. This approach can be integrated into modern industrial production lines, offering enhanced quality control, waste reduction, and increased equipment autonomy — fully aligned with the principles of Industry 4.0.

Ευχαριστίες

Η παρούσα διπλωματική εργασία με τίτλο **“Παρακολούθηση Κατεργασιών Τόρνευσης με Τεχνικές Μηχανικής Μάθησης”** αποτελεί την ολοκλήρωση των σπουδών μου στο ΔΠΜΣ “Συστήματα Αυτοματισμού” της σχολής Μηχανολόγων Μηχανικών του Ε.Μ.Π. Η τελική διαμόρφωση της εργασίας είναι αποτέλεσμα εκτενούς αναζήτησης και έρευνας, πειραμάτων και δοκιμών καθώς και προϊόν υποστήριξης και καθοδήγησης συγκεκριμένων ανθρώπων τους οποίους θα ήθελα σε αυτό το σημείο να ευχαριστήσω θερμά.

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή της διπλωματικής μου εργασίας, Κο. Γεώργιο-Χριστόφορο Βοσνιάκο, για την εμπιστοσύνη που μου έδειξε κατά την επιλογή του θέματος.

Επίσης, ευχαριστώ ιδιαίτερα τον Δρα. Εμμανουήλ Σταθάτο, για τις πολύτιμες συμβουλές, την συνεχή στήριξη και τις ουσιαστικές παρατηρήσεις για την ολοκλήρωση της εργασίας. Ακόμα, τον ευχαριστώ για τη γνώση που μου μετέφερε στο σχετικό μάθημα του μεταπτυχιακού προγράμματος.

Δεν μπορώ να παραλείψω να ευχαριστήσω τους φίλους και συμφοιτητές μου για την υποστήριξη τους και τη συνεχή ενθάρρυνση που μου επέδειξαν καθώς συνέβαλλαν έμμεσα στην προσπάθεια μου για την συνεχή βελτίωση.

Τέλος θα ήθελα να ευχαριστήσω την οικογένεια μου, για την αμέριστη υποστήριξη που μου επέδειξε συνολικά, αποτελώντας ισχυρό κίνητρο για την επιτυχή ολοκλήρωσή του μεταπτυχιακού.

Καψής Χαράλαμπος

Ιούλιος 2025

1 Εισαγωγή

Στο πλαίσιο της **4ης Βιομηχανικής Επανάστασης (Industry 4.0)**, η οποία πραγματεύεται την αυτοματοποίηση των διαδικασιών, ούτως ώστε να αυξηθεί παθητικά η παραγωγικότητα και εν συνεχεία το κέρδος, παρατηρείται η ανάγκη για την ανάπτυξη τεχνολογιών που θα ενισχύσουν αυτήν την ιδέα. Ένα τεράστιο εργαλείο προς αξιοποίηση είναι η τεχνολογία της **Τεχνητής Νοημοσύνης (Artificial Intelligence)**. Η τεχνητή νοημοσύνη, εάν αξιοποιηθεί σωστά, μπορεί να προσφέρει πολύ καλές και εύστοχες λύσεις ως προς την παραγωγικότητα, κάτι που την κάνει αυτομάτως πολύ σημαντική και ελκυστική στους βιομήχανους και στον κλάδο της τεχνολογίας γενικότερα.

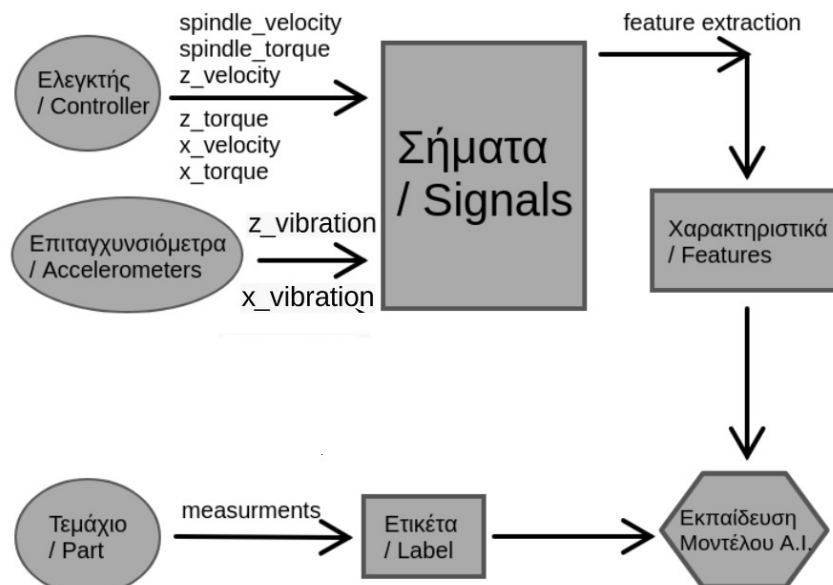
Η ανάπτυξη μιας τέτοιας τεχνολογίας προσφέρει λύσεις σε πάρα πολλά προβλήματα ανεξαρτήτως κλάδου (Βιομηχανία, Ιατρική, Οικονομικά κλπ.), συνεπώς προσφέρει λύση και στο πρόβλημα που αναλύεται στην παρακάτω διπλωματική. Δίνοντας τα κατάλληλα δεδομένα στο μοντέλο **Μηχανικής Μάθησης (Machine Learning)** ή στο **Νευρωνικό Δίκτυο (Neural Network)** που θα αναπτυχθεί και μέσω της κατάλληλης επεξεργασίας των δεδομένων αυτόν, θα παραχθεί ένα αποτέλεσμα το οποίο σίγουρα θα είναι αποδεκτό και πολύ πιο αποδοτικό από άλλες συμβατικές διαδικασίες (π. χ. οπτικός έλεγχος προϊόντος)

Το θέμα στο οποίο καλούμαστε να βρούμε λύση αφορά τον τομέα της **Κατασκευής (Manufacturing)**. Χρησιμοποιώντας μια **Εργαλειομηχανή CNC**, στην προκειμένη περίπτωση ενός **Τόρνου**, γίνεται κατεργασία κοπής προϊόντων (**Τόρνευση / Turning**). Όπως είναι λογικό, η ποιότητα των προϊόντων ποικίλει με βάση τις συνθήκες του περιβάλλοντος, τον χειρισμό της μηχανής κλπ. Αυτό έχει αντίκτυπο στην ποιότητα του παραγόμενου τεμαχίου, κάτι που σημαίνει πως πρέπει να ελέγχεται.

Με χρήση της κατάλληλης διάταξης, συλλέχθηκαν ποιοτικά δεδομένα που αφορούσαν το τεμάχιο (**roughness_Ra, roughness_Rz, diameter_deviation, actual_diameter**) αλλά και την μηχανή (**feed, speed, depth, RPM**) μέσω των οποίων δημιουργήθηκαν κλάσεις ποιότητας του τεμαχίου (**Low, Medium, High**). Συλλέχθηκαν επίσης μέσω του **ελεγκτή (controller)** της μηχανής (**spindle_velocity, spindle_torque, z_velocity, z_torque, x_velocity, x_torque**) αλλά και μέσω επιταχυνσιομέτρων (**z_vibration, x_vibration**) σήματα τα οποία αφορούσαν στοιχεία της μηχανής, μέσω αυτών εξήχθησαν χαρακτηριστικά, τα οποία δώθηκαν στο μοντέλο για να γίνει η εκπαίδευση, ούτως ώστε να γίνει η κατηγοριοποίηση των κλάσεων ποιότητας του παραγόμενου τεμαχίου.

Τα αποτελέσματα που προέκυψαν από την παρούσα μελέτη παρουσίασαν σημαντική διακύμανση, γεγονός που αποδίδεται στις διαφορετικές τεχνικές που εφαρμόστηκαν καθ' όλη τη διάρκεια του πειραματισμού. Συγκεκριμένα, παρατηρήθηκαν αποκλίσεις στην ακρίβεια μεταξύ των αλγορίθμων **Random Forest** και **Πολυεπίπεδων Νευρωνικών Δικτύων**, τόσο πριν όσο και

μετά την εφαρμογή τεχνικών εξισορρόπησης όπως η επαναδειγματοληψία (**upsampling**) και το **SMOTE**. Επιπλέον, η επιλογή των χαρακτηριστικών εισόδου μέσω μεθόδων όπως οι δείκτες σημασίας του **Random Forest** επηρέασε καθοριστικά την απόδοση των μοντέλων. Η τελική βελτιστοποιημένη έκδοση του Νευρωνικού Δικτύου, η οποία περιλάμβανε κανονικοποίηση, χρήση **Batch Normalization** και κατάλληλες υπερπαραμέτρους, πέτυχε ακρίβεια 86.03%, επιβεβαιώνοντας ότι η προσεκτική παραμετροποίηση και προεπεξεργασία οδηγούν σε ουσιαστική βελτίωση της απόδοσης του συστήματος ταξινόμησης. Τα αποτελέσματα αυτά θα σχολιαστούν και θα αναλυθούν στην συνέχεια αυτής της εργασίας, όπως και τα συμπεράσματα που προέκυψαν μέσω αυτών.



Σχήμα 1: Λογικό Διάγραμμα Κατασκευής Μοντέλου

1.1 Δομή της εργασίας

Η εργασία είναι οργανωμένη σε πέντε κύρια κεφάλαια. Στο πρώτο κεφάλαιο παρουσιάζεται η γενική εισαγωγή, το αντικείμενο και οι στόχοι της μελέτης, καθώς και η διάρθρωση της εργασίας. Στο δεύτερο κεφάλαιο ακολουθεί ανασκόπηση της σχετικής βιβλιογραφίας με έμφαση στις σύγχρονες τεχνολογίες παρακολούθησης κατεργασιών και τις εφαρμογές της μηχανικής μάθησης στη βιομηχανία. Στο τρίτο κεφάλαιο περιγράφεται αναλυτικά ο πειραματικός εξοπλισμός, η μεθοδολογία συλλογής δεδομένων και η διαδικασία εξαγωγής χαρακτηριστικών από τα σήματα. Στο τέταρτο κεφάλαιο αναπτύσσεται η στρατηγική προεπεξεργασίας, η εξισορρόπηση δεδομένων και η εκπαίδευση των μοντέλων, παρουσιάζονται τα αποτελέσματα που παράχθηκαν. Τέλος, στο πέμπτο κεφάλαιο συνοψίζονται τα κύρια συμπεράσματα της έρευνας και διατυπώνονται προτάσεις για περαιτέρω επέκταση της παρούσας μελέτης.

1.2 Αντικείμενο και σκοπός της εργασίας

Αντικείμενο της παρούσας εργασίας είναι η ανάπτυξη και αξιολόγηση ενός ευφυούς συστήματος παρακολούθησης της ποιότητας του παραγόμενου τεμαχίου και η καταταξή του σε τρεις προκαθορισμένες κατηγορίες με βάση την ποιότητά τους : **Low, Medium, High**. Αυτή η κατηγοριοποίηση γίνεται με βάση τις ποιοτικές τιμές που προαναφέρθηκαν (**roughness_Ra, roughness_Rz, diameter_deviation, actual_diameter, feed, speed, depth, RPM**) όπου :

- **Roughness_Ra**: Η μέση τραχύτητα επιφάνειας, ο μέσος όρος των αποκλίσεων ύψους από την κεντρική γραμμή. Δείκτης γενικής ποιότητας κατεργασμένης επιφάνειας.
- **Roughness_Rz**: η μέγιστη τιμή της τραχύτητας της επιφάνειας – το άθροισμα των πέντε μεγαλύτερων υψών και βαθύτερων κοιλοτήτων. Αντιπροσωπεύει τις ακραίες τιμές τραχύτητας.
- **diameter_deviation**: Η απόκλιση της διαμέτρου από την ονομαστική. Μετρά πόσο διαφέρει η πραγματική διάμετρος από την επιθυμητή.
- **actual_diameter**: Η πραγματική μετρούμενη διάμετρος του τεμαχίου μετά την κατεργασία.
- **feed**: Ταχύτητα πρόωσης, απεικονίζει πόσο προχωρά το εργαλείο ανά μονάδα χρόνου ή περιστροφής. Επηρεάζει τη μορφολογία και τραχύτητα επιφάνειας.
- **speed**: Ταχύτητα κοπής. Είναι η γραμμική ταχύτητα που μετράται μεταξύ εργαλείου και τεμαχίου στο σημείο επαφής.
- **depth**: Βάθος κοπής, δηλαδή το πάχος του υλικού που αφαιρείται κατά μία διέλευση του εργαλείου.
- **RPM**: Περιστροφές ανά λεπτό του άξονα (spindle), σχετίζεται άμεσα με την ταχύτητα κοπής.

αυτό επιτυγχάνεται μέσω οριοθέτησης τιμών σε αυτά τα measurements μέσω συναρτήσεων (**functions**). Απο αυτές τις ποιοτικές τιμές, χρησιμοποιείται ως **label** στο πείραμα μας η κλάση του **Roughness_Ra**. Αυτές οι πληροφορίες απο μόνες τους δεν προσφέρουν κάτι, εκεί έρχεται το πακέτο των signals (**z_vibration, x_vibration, spindle_velocity, spindle_torque, z_velocity, z_torque, x_velocity, x_torque**), όπου:

- **z_vibration**: Οι δονήσεις κατά τον κατακόρυφο άξονα (Z), μετρούνται από το επιταχυνσιόμετρο στον άξονα Z και σχετίζονται με την ευστάθεια της κοπής και την πιθανή εμφάνιση κραδασμών.
- **x_vibration**: Οι δονήσεις κατά τον οριζόντιο άξονα (X), υποδεικνύουν πλευρική αστάθεια ή ανωμαλίες κατά τη διάρκεια της κατεργασίας, μετρούνται και αυτά απο το επιταχυνσιόμετρο που έχει εγκατασταθεί στον άξονα X.

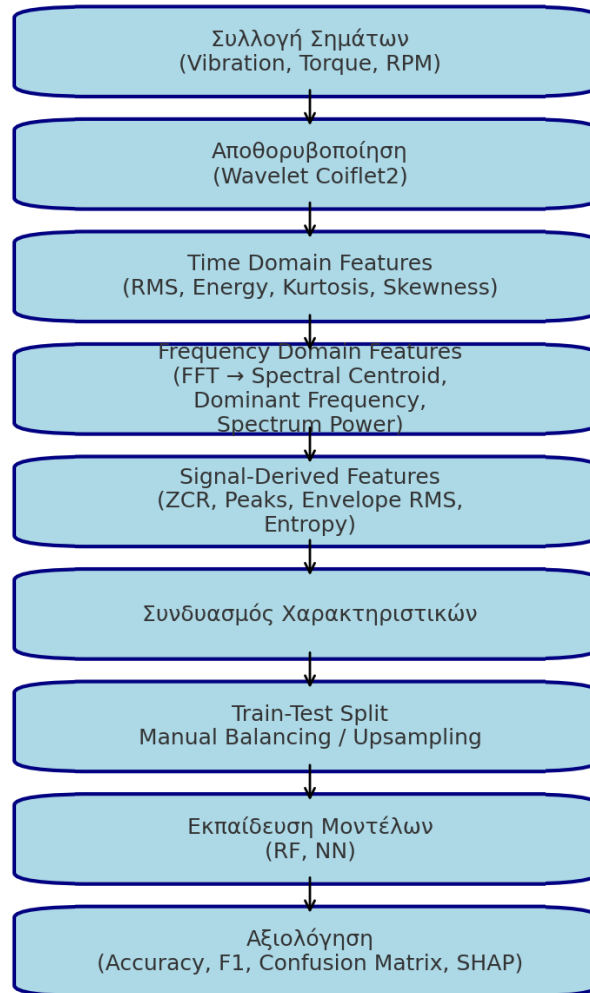
- **spindle_velocity**: Η ταχύτητα περιστροφής του κύριου άξονα (spindle), καθορίζει τη σχετική ταχύτητα κοπής μεταξύ εργαλείου και τεμαχίου. Η πληροφορία αυτή προέρχεται από τον ελεγκτή της μηχανής.
- **spindle_torque**: Η ροπή του άξονα, δείκτης του φορτίου που ασκείται στον κινητήρα κατά την κοπή. Αυξάνεται με τη δυσκολία της κατεργασίας. Και αυτό δίνεται από τον controller.
- **z_velocity**: Η ταχύτητα μετακίνησης του εργαλείου ή του τραπεζιού στον άξονα Z, σχετίζεται με το βάθος κοπής και την κατακόρυφη πρόωση. Προέρχεται από τον controller.
- **z_torque**: Η ροπή κατά μήκος του άξονα Z, εκφράζει την αντίσταση του υλικού στην κάθετη διείσδυση του εργαλείου. Ο ελεγκτής μας δίνει και αυτό το σήμα.
- **x_velocity**: Η ταχύτητα πρόωσης στον άξονα X, συνδέεται με την κύρια κατεύθυνση κατεργασίας (κατά μήκος). Η πληροφορία για αυτό το σήμα δίνεται από τον ελεγκτή.
- **x_torque**: Η ροπή στον άξονα X, υποδεικνύει την αντίσταση κατά την πρόωση και μπορεί να αποκαλύψει προβλήματα στο εργαλείο ή στο υλικό. Οι μετρήσεις δίνονται από τον controller.

το οποίο μέσω του κατάλληλου **feature extraction** εξάγει χαρακτηριστικά, μέσω των οποίων γίνεται η συσχέτιση με το **label** και προκύπτει η πρόβλεψη. Τα προαναφερθέντα σήματα μας δίνονται από τα **επιταχυνσιόμετρα (z_vibration, x_vibration)** και τον **ελεγκτή (spindle_velocity, spindle_torque, z_velocity, z_torque, x_velocity, x_torque)** της μηχανής. Έτσι μπορεί να προβλέπει αυτόματα την κατηγορία ποιότητας ενός νέου τεμαχίου βάσει των σημάτων αισθητήρων.

Συνοψίζοντας, στόχος της διπλωματικής εργασίας είναι ο σχεδιασμός και η υλοποίηση ενός ευφυούς συστήματος ικανού να **ταξινομή την ποιότητα** των κατεργασμένων τεμαχίων σε **χαμηλή, μεσαία ή υψηλή** τραχύτητα βάσει σημάτων αισθητήρων από την τόννευση. Αυτή η τεχνολογία, ταιριάζει απόλυτα με τους στόχους της 4ης Βιομηχανικής Επανάστασης, καθώς αναμένεται να βελτιώσει την αξιοπιστία και την αποδοτικότητα των κατεργασιών τόννευσης, επιτρέποντας την έγκαιρη ανίχνευση τυχόν προβλημάτων ποιότητας και συμβάλλοντας στη μείωση των αποβλήτων και των απρογραμμάτιστων διακοπών παραγωγής, γεγονός που αντιπροσωπεύει απόλυτα την αυτοματοποίηση των διαδικασιών στη βιομηχανία.

1.3 Διάγραμμα Ροής

Η διαδικασία κατασκευής του μοντέλου Μηχανικής Μάθησης που πραγματεύεται αυτή η εργασία απαιτεί μια δομή με μια προκαθορισμένη ροή.



Σχήμα 2: Διάγραμμα Ροής Επεξεργασίας Σημάτων

2 Ανασκόπηση Βιβλιογραφίας

2.1 Βασικές Αρχές Κατεργασίας Τόρνευσης

Η κατεργασία τόρνευσης είναι μια βασική κατεργασία κοπής μετάλλων, όπου ένα περιστρεφόμενο τεμάχιο κατεργάζεται με κοπτικό εργαλείο για την παραγωγή κυλινδρικών ή περιστροφικών εξαρτημάτων. Οι εργαλειομηχανές CNC τόρνου συγκρατούν και περιστρέφουν το τεμάχιο, ενώ ένα ακίνητο κοπτικό εργαλείο αφαιρεί υλικό. Τα κύρια **παραμετρικά χαρακτηριστικά** που ορίζουν τις συνθήκες κοπής είναι η **ταχύτητα κοπής** (συνάρτηση της γωνιακής ταχύτητας του τεμαχίου και της διαμέτρου του, συνήθως σε **m/min** ή στροφές ανά λεπτό), η **πρόωση** (ρυθμός τροφοδοσίας εργαλείου ανά περιστροφή, π. χ. **mm/rev**) και το **βάθος κοπής** (πάχος στρώσης υλικού που αφαιρείται σε κάθε πέρασμα, σε **mm**). Η επιλογή αυτών των παραμέτρων επηρεάζει άμεσα τις δυνάμεις κοπής, τη φθορά εργαλείου και το αποτέλεσμα στην επιφάνεια του τεμαχίου. Για παράδειγμα, αύξηση της πρόωσης τείνει να αυξάνει τις δυνάμεις κοπής και τους κραδασμούς, ενώ η αύξηση της ταχύτητας κοπής μπορεί να μειώσει τις δυνάμεις κοπής λόγω φαινομένων θερμομαλακώσεως. Έχει παρατηρηθεί ότι υπερβολικά μεγάλες δυνάμεις κοπής μπορούν να οδηγήσουν σε τραχύτερη επιφάνεια και βλάβες στο κατεργασμένο κομμάτι. Επίσης, οι **δονήσεις** (ταλαντώσεις εργαλείου/τεμαχίου γνωστές και ως φαινόμενο **chatter**) αποτελούν σύνηθες πρόβλημα στην τόρνευση, το οποίο επιδεινώνει τη φθορά του εργαλείου και επηρεάζει αρνητικά την επιφανειακή ποιότητα. Συνεπώς, η βέλτιστη ρύθμιση των παραμέτρων (**ταχύτητα, πρόωση, βάθος**) είναι κρίσιμη τόσο για την επίτευξη καλής ποιότητας, όσο και για την παράταση της ζωής του εργαλείου.

Σε περιβάλλον παραγωγής, η **παρακολούθηση της διεργασίας τόρνευσης** με κατάλληλους αισθητήρες είναι θεμελιώδης για τη διασφάλιση σταθερής ποιότητας και την αποφυγή ζημιών. Πλήθος φυσικών μεγεθών μπορούν να μετρηθούν σε πραγματικό χρόνο: ενδεικτικά, η **ισχύς ή το ρεύμα του κινητήρα**, οι **δυνάμεις κοπής**, η **ροπή**, η **ακουστική εκπομπή**, οι **κραδασμοί/δονήσεις**, **εικόνες/οπτικά σήματα**, η **θερμοκρασία**, οι **μετατοπίσεις** και οι **μηχανικές παραμορφώσεις**. Οι αντίστοιχοι **αισθητήρες** περιλαμβάνουν πιεζοηλεκτρικούς δυναμομετρητές για μέτρηση δυνάμεων, αισθητήρες ροπής ή μετρητές ρεύματος άξονα για την **εκτίμηση κοπτικής δύναμης**, επιταχυνσιόμετρα για δονήσεις, μικρόφωνα ή αισθητήρες ακουστικών εκπομπών για **ήχο**, θερμοζεύγη ή υπέρυθρες για **θερμοκρασία**, κ. ά. Σε σύγχρονους CNC τόρνους, τέτοιοι αισθητήρες μπορούν να ενσωματωθούν ή να προστεθούν εξωτερικά, παρέχοντας δεδομένα επι τόπου από τη διεργασία. Η συλλογή και ανάλυση αυτών των σημάτων σε πραγματικό χρόνο επιτρέπει την **επιτήρηση της κατάστασης** του εργαλείου και της κατεργασίας, με στόχο τη βελτίωση της παραγωγικότητας, της αυτοματοποίησης και της ποιότητας. Για παράδειγμα, σε πρόσφατη μελέτη τόρνευσης χάλυβα **11SMn30** χρησιμοποιήθηκε δυναμομετρητής και αισθητήρας **MPU6050** για την καταγραφή δυνάμεων και δονήσεων κατά την κατεργασία, επιτρέποντας τη στατιστική συσχέτιση των κοπτικών παραμέτρων με τα σήματα και την ποιότητα επιφάνειας (π.χ. συσχέτιση $RMS-Ra \approx 0,92$) [8] [9].

Σχετικά με τις πληροφορίες που πήραμε από τα σήματα, λόγω του μεγάλου όγκου των ακατέργαστων σημάτων, είναι απαραίτητη η μετατροπή τους σε κατάλληλα *χαρακτηριστικά (features)* – δηλαδή σε μια περιορισμένη σειρά από περιγραφικές παραμέτρους που συνοψίζουν τις πληροφορίες των σημάτων. Σύμφωνα με τη σχετική βιβλιογραφία [1], για τη μείωση της διάστασης των σημάτων κατεργασίας εφαρμόζονται παραδοσιακά τεχνικές εξαγωγής χαρακτηριστικών στον χρόνο και στη συχνότητα (όπως RMS, FFT κλπ.), με στόχο τη διατήρηση κρίσιμης πληροφορίας. Ωστόσο, η μεθοδολογία του παρόντος έργου αποφεύγει τη χειροκίνητη εξαγωγή και χρησιμοποιεί συνελκτικό νευρωνικό δίκτυο (1D CNN) για αυτόματη ανάδειξη των σημαντικών προτύπων απευθείας από τα ακατέργαστα σήματα.

Οι κατάλληλοι συνδυασμοί χαρακτηριστικών από το πεδίο χρόνου και συχνότητας παρέχουν μια ολοκληρωμένη περιγραφή της κατάστασης κοπής. Κάθε χαρακτηριστικό είναι ευαίσθητο σε διαφορετικό κομμάτι του σήματος, επομένως ο συνδυασμός πολλών χαρακτηριστικών επιτρέπει πολύπλευρη ανάλυση και αυξάνει την ικανότητα διάκρισης μεταξύ συνθηκών. Στη βιβλιογραφία έχουν αναφερθεί παρόμοια σετ χαρακτηριστικών για την παρακολούθηση εργαλείων και την ανίχνευση φθοράς [1] [3] [4]: π.χ. το RMS και η ενέργεια σχετίζονται με το επίπεδο δονήσεων, ενώ υψηλές τιμές κυρτότητας, ασυμμετρίας ή συντελεστή αιχμής έχουν συνδεθεί με εμφάνιση κραδασμών λόγω μικρο-θραύσεων ή ταλαντώσεων κοπή

2.2 Τεχνητή Νοημοσύνη στη Βιομηχανία

Η **Τεχνητή Νοημοσύνη (TN)**, και ειδικότερα τα πεδία της **Μηχανικής Μάθησης (Machine Learning)** και της **Βαθιάς Μάθησης (Deep Learning)**, αποτελούν βασικούς άξονες της 4ης **Βιομηχανικής Επανάστασης (Industry 4.0)**, προάγοντας την έξυπνη αυτοματοποίηση της παραγωγής, τη διασύνδεση εξοπλισμού, και τη δυναμική προσαρμογή της παραγωγικής διαδικασίας σε πραγματικό χρόνο [10] [14]. Η εφαρμογή τεχνικών **TN** στη βιομηχανία επιτρέπει την αποτελεσματικότερη παρακολούθηση διεργασιών, τη βελτιστοποίηση χρήσης πόρων, καθώς και την προγνωστική συντήρηση εξοπλισμού, οδηγώντας σε μείωση του λειτουργικού κόστους, αύξηση της διαθεσιμότητας και ελαχιστοποίηση του απρογραμμάτιστου χρόνου εκτός λειτουργίας [12] [13] [14].

Πλήθος ερευνητικών εργασιών έχουν τεκμηριώσει σημαντικές προόδους στον τομέα του **condition monitoring** και της **έγκαιρης διάγνωσης προβλημάτων κατεργασιών**, μέσω της συλλογής και ανάλυσης αισθητηριακών σημάτων (όπως δονήσεις, κοπτικές δυνάμεις, θερμοκρασία ή ρεύμα κινητήρα) με τη χρήση εποπτευόμενων ή μη εποπτευόμενων αλγορίθμων [3] [10] [13] [15]. Οι τεχνικές αυτές βρίσκουν εφαρμογή όχι μόνο στη συντήρηση (π.χ. **predictive maintenance**), αλλά και στην **πρόβλεψη της ποιότητας** του παραγόμενου τεμαχίου, στη **διασφάλιση σταθερών συνθηκών κατεργασίας**, καθώς και στην **ανίχνευση αποκλίσεων** που σχετίζονται με φθορά εργαλείων, απώλεια ακριβείας ή κακή ρύθμιση των μηχανών [6] [11] [14].

Σε σύγχρονες έξυπνες γραμμές παραγωγής, οι αισθητήρες είναι ενσωματωμένοι στο εργαλείο, στο τεμάχιο ή στις μονάδες μετάδοσης κίνησης, και αποστέλλουν **δεδομένα σε πραγματικό χρόνο** προς επεξεργασία. Αυτά τα δεδομένα αναλύονται μέσω τεχνικών όπως τα **Support Vector Machines (SVM)**, τα **Random Forests**, ή τα **Artificial Neural Networks (ANNs)** για την πρόβλεψη κρίσιμων δεικτών όπως η **τραχύτητα επιφάνειας**, η **φθορά κοπτικού εργαλείου**, η **διάμετρος κατεργασμένου τεμαχίου**, και η **παραβίαση ανοχών** [6] [11] [13] [15].

Η επιλογή των συγκεκριμένων μοντέλων βασίστηκε στη βιβλιογραφία [1] [5] [6] παρακολούθησης εργαλείων κοπής: μελέτες ανασκόπησης έχουν δείξει ότι μέθοδοι όπως το **Random Forest Classifiers** και τα **Artificial Neural Networks** συγκαταλέγονται στις πιο διαδεδομένες προσεγγίσεις για την παρακολούθηση φθοράς εργαλείων στις βιομηχανικές κατεργασίες. Επίσης βάση βιβλιογραφίας [7], τα πολυεπίπεδα νευρωνικά δίκτυα είναι από τις πλέον δημοφιλείς μεθόδους στην παρακολούθηση εργαλείων τόρνου – μάλιστα έχει καταγραφεί ότι στις περισσότερες μελέτες **turning** χρησιμοποιούνται **ANN** μοντέλα για την πρόγνωση της φθοράς, λόγω της υψηλής τους ικανότητας μάθησης μη-γραμμικών σχέσεων.

Επιπλέον, η αρχιτεκτονική των **Κυβερνο-Φυσικών Συστημάτων (Cyber-Physical Systems – CPS)** διευκολύνει την ολοκλήρωση φυσικών μηχανών με υπολογιστικά υποσυστήματα TN και δικτύωση μέσω **IoT**. Η σύνδεση αυτή οδηγεί σε **αυτορρυθμιζόμενα, αποκεντρωμένα συστήματα ελέγχου** που λαμβάνουν αποφάσεις βάσει πραγματικών δεδομένων και ιστορικών προτύπων λειτουργίας [10] [14] [16].

Πέρα από την πρόβλεψη της φθοράς εργαλείων ή της ποιότητας επιφάνειας, τα δεδομένα που συλλέγονται σε πραγματικό χρόνο αξιοποιούνται όλο και συχνότερα σε πλαίσια αυτομάθησης και αυτό-βελτιστοποίησης της διαδικασίας κοπής. Τεχνικές **Reinforcement Learning (RL)** εφαρμόζονται πειραματικά για τη δυναμική προσαρμογή παραμέτρων όπως ταχύτητα κοπής ή βάθος, βάσει των αισθητηριακών ενδείξεων, επιτυγχάνοντας έτσι μία κυκλική ροή δεδομένων – από τη μέτρηση στην απόφαση και πάλι πίσω [13] [14]

Αξιοσημείωτο είναι επίσης το γεγονός ότι η χρήση **TN** επιτρέπει την ενοποίηση ανομοιογενών πηγών πληροφορίας, όπως σήματα από επιταχυνσιόμετρα, ροπές άξονα και αισθητήρες ρεύματος, επιτρέποντας τη δημιουργία μοντέλων υψηλής γενικευσιμότητας. Μελέτες όπως αυτή των M. Kaur και A. Aggarwal [8] [9], αξιοποιούν αισθητήρες τύπου **MPU6050** για τη συσχέτιση δονήσεων και τραχύτητας, αναδεικνύοντας τη σημασία των μη-παραδοσιακών αισθητηρίων στις διαδικασίες κοπής.

Ενδεικτικά, ο πίνακας που ακολουθεί συνοψίζει τις βασικές εφαρμογές,σε επίπεδο **AI**,στον βιομηχανικό κλάδο :

Πεδίο Εφαρμογής	Περιγραφή	Μοντέλο Τεχνητής Νοημοσύνης
Προγνωστική Συντήρηση (Predictive Maintenance)	Πρόβλεψη βλαβών/αστοχιών εξοπλισμού πριν την εμφάνισή	Random Forest, LSTM, SVM

Πεδίο Εφαρμογής	Περιγραφή	Μοντέλο Τεχνητής Νοημοσύνης
	τους	
Παρακολούθηση Φθοράς Εργαλείων	Εκτίμηση της κατάστασης κοπτικών εργαλείων από σήματα δόνησης, δύναμης, ροπής	Neural Networks, k-NN, Decision Trees
Πρόβλεψη Ποιότητας Προϊόντος	Συσχέτιση παραμέτρων κατεργασίας και αισθητηριακών σημάτων με δείκτες ποιότητας τεμαχίων	RFC,SVM, XGBoost, Deep Learning
Ανίχνευση Ανωμαλιών Διεργασίας	Αναγνώριση αποκλίσεων από ομαλή λειτουργία σε πραγματικό χρόνο	Autoencoders, Isolation Forest, PCA+Clustering
Βελτιστοποίηση Παραμέτρων	Επιλογή των βέλτιστων συνθηκών κατεργασίας με στόχο την απόδοση ή ποιότητα	Bayesian Optimization, Genetic Algorithms

2.3 Ποιότητα Προϊόντων και Ταξινόμηση τους

Ως ποιότητα της κατεργασίας στην τόννευση νοείται ο έλεγχος της γεωμετρίας και της κατάστασης επιφανείας των παραγόμενων τεμαχίων. Η ποιότητα προϊόντος περιλαμβάνει τις διαστατικές αποκλίσεις, τα σφάλματα μορφής, καθώς και το φινιρίσμα επιφάνειας – χαρακτηριστικά που σχετίζονται άμεσα με την απόδοση και τη λειτουργικότητα του τεμαχίου στο τελικό σύστημα [11]. Ευρέως διαδεδομένοι δείκτες φινιρίσματος είναι οι παράμετροι τραχύτητας **Ra** (μέση απόλυτη απόκλιση) και **Rz** (μέσο ύψος κορυφογραμμής). Χαμηλές τιμές **Ra–Rz** συνδέονται με βελτιωμένη μηχανική αντοχή, μειωμένη τριβή και αυξημένη αισθητική ποιότητα [6] [11]. Αντίστοιχα, η διαστατική ακρίβεια (π.χ. κυκλικότητα, διάμετρος) συνδέεται με την πιστή αναπαραγωγή του σχεδιασμού και τη συμβατότητα συναρμογής.

2.4 Χαρακτηριστικά ποιότητας

- **Τραχύτητα επιφάνειας (R_a , R_z):** Εκφράζει την ανωμαλία της επιφάνειας σε μικροκλίμακα. Χαμηλές τιμές R_a – R_z σημαίνουν λεία επιφάνεια, άρα μικρότερη φθορά και κόπωση των μερών. Η τραχύτητα μετρείται επί τόπου (με στυλό-μέτρη) ή έμμεσα μέσω αισθητήρων (**δόνηση, δυνάμεις**). [6] [18]
- **Διαστατική ακρίβεια:** Αφορά τις αποκλίσεις των διαστάσεων από το σχεδιασμό (**π.χ. διάμετρος ή κυκλικότητα**). Επηρεάζεται από την ακαμψία, τη φθορά εργαλείου και τις δυνάμεις κοπής. Σε λεπτά τεμάχια, η ακρίβεια μπορεί να προβλεφθεί από μετρήσεις ριζικών συνιστωσών της δύναμης κοπής και κραδασμών. Αξιολογείται μέσω της απόκλισης από τη σχεδιασμένη διάμετρο ή γεωμετρία. Επηρεάζεται από παράγοντες όπως η ακαμψία του συστήματος, η φθορά εργαλείου, οι θερμικές διαστολές και οι κοπτικές δυνάμεις. [11] [17]

2.5 Επίδραση παραμέτρων κατεργασίας

Οι συνθήκες κοπής (**πρόωση, ταχύτητα, βάθος κοπής**) επηρεάζουν έντονα τα χαρακτηριστικά ποιότητας. Μελέτες δείχνουν ότι η **πρόωση (feed)** έχει την ισχυρότερη επίδραση στην τραχύτητα, ακολουθούμενη από την **ταχύτητα κοπής** και το **βάθος κοπής**. Συγκεκριμένα, με αυξημένη τροφοδότηση η επιφάνεια γίνεται συνήθως τραχύτερη (**μεγαλύτερα όρια κύρτωσης**), ενώ υψηλότερες ταχύτητες κοπής βελτιώνουν το φινίρισμα (**μείωση R_a**). Το βάθος κοπής έχει γενικά μικρότερη επίδραση, αν και πολύ μεγάλα βάθη τείνουν να αυξάνουν την τραχύτητα [11] [17]. Γενικότερα, μικρές μεταβολές στις συνθήκες αυτές οδηγούν σε σημαντικές διαφορές στην τραχύτητα.

2.6 Διεθνή πρότυπα και κριτήρια ποιότητας στην τórνευση

Η αξιολόγηση της ποιότητας στην τórνευση βασίζεται σε διεθνώς καθορισμένα πρότυπα, τα οποία εξασφαλίζουν τη συγκρισιμότητα και την επαναληψιμότητα των μετρήσεων σε βιομηχανικό περιβάλλον. Ειδικότερα, για την εκτίμηση της τραχύτητας επιφάνειας εφαρμόζονται τα πρότυπα **ISO 4287** και **ISO 1302**, τα οποία καθορίζουν τις παραμέτρους περιγραφής της μικρογεωμετρίας, όπως το **R_a , R_z , R_q (τετραγωνική μέση τιμή)** κ.ά., καθώς και τις προδιαγραφές των μεθόδων μέτρησης (π.χ. μετρολογικά όργανα, μήκος αξιολόγησης, φίλτρα). Η σωστή επιλογή των παραμέτρων εξαρτάται από τη λειτουργική απαίτηση της επιφάνειας και το είδος της κατεργασίας (**π.χ. τραχιά, ημιτελής ή τελική τórνευση**) [19].

Η σήμανση των απαιτήσεων τραχύτητας στα τεχνικά σχέδια καθορίζεται από το **ISO 1302**, το οποίο συνδέεται με τις υπόλοιπες γεωμετρικές ανοχές μέσω του συστήματος **GD&T (Geometric Dimensioning and Tolerancing)**. Η εφαρμογή **GD&T** επιτρέπει τον ακριβή προσδιορισμό ορίων αποκλίσεων σε χαρακτηριστικά όπως ευθύτητα, κυλινδρικότητα, παράλληλοτητα και ομοαξονικότητα. Κατά αυτόν τον τρόπο, ο ποιοτικός έλεγχος δεν περιορίζεται σε απομονωμένες

μετρήσεις διαμέτρου ή τραχύτητας, αλλά εξετάζει τη γεωμετρική συμβατότητα του τεμαχίου με το σύνολο του συστήματος συναρμολόγησης.

Σε εφαρμογές υψηλής ακρίβειας, όπως στην αεροναυπηγική ή στην ιατρική μηχανολογία, η συμμόρφωση με τα παραπάνω πρότυπα είναι κρίσιμη, καθώς επηρεάζει την αλληλεπίδραση της επιφάνειας με άλλα εξαρτήματα (**τριβή, φθορά, λιπαντικότητα**), αλλά και την ασφάλεια και αντοχή του τελικού προϊόντος. Συνεπώς, η χρήση πιστοποιημένων προτύπων και διαδικασιών μέτρησης αποτελεί αναπόσπαστο τμήμα κάθε σύγχρονου **CNC** ποιοτικού ελέγχου [19] [20].

3 Εξοπλισμός και δεδομένα

3.1 Περιγραφή πειραματικού εξοπλισμού τόρνευσης

Για την εκπόνηση του πειράματος η κατεργασία πραγματοποιήθηκε σε τόρνο **CNC** τύπου **Okuma™ LB10II**, ο οποίος είναι εφοδιασμένος με σύστημα ελέγχου **OSP700L**. Ο τόρνος διαθέτει μέγιστη διάμετρο τόρνευσης **170 mm**, μέγιστο μήκος παραγόμενου τεμαχίου **300 mm**, μέγιστη ταχύτητα ατράκτου **10.000 RPM** και πύργο εργαλείων με **12 θέσεις**. Η φόρτωση του **G-code** έγινε μέσω θύρας σειριακής επικοινωνίας με δυνατότητα **DNC**. Για την κατεργασία χρησιμοποιήθηκαν κοπτικά ένθετα **Sandvik Coromant™ CNMG 12 04 08-MR 4305**, τοποθετημένα σε βάση εργαλείου **Sandvik Coromant™ PCLNR 2020K 12**.

Δεδομένου ότι το μηχάνημα δεν διαθέτει ενσωματωμένη δυνατότητα εξαγωγής σήματος σε πραγματικό χρόνο (π. χ. ταχύτητα ατράκτου, γραμμικές ταχύτητες αξόνων, ροπές κοπής), έγινε άμεση σύνδεση με την πλακέτα ελέγχου του συστήματος CNC. Μέσω αυτής αντλήθηκαν αναλογικά σήματα για τις ταχύτητες και τις ροπές των κινητήρων. Οι ταχύτητες αντιστοιχούν στην περιστροφική κίνηση της ατράκτου και στις γραμμικές κινήσεις των αξόνων X (ακτινικά) και Z (διαμήκη), ενώ οι ροπές συνδέονται με τις συνιστώσες των δυνάμεων κοπής. Η σύνδεση επιτεύχθηκε με ειδικούς ακροδέκτες τύπου **hook**, οι οποίοι οδήγησαν σε θύρα 10 ακροδεκτών, στερεωμένη σε ειδική εγκοπή στο κάλυμμα του τόρνου, επιτρέποντας άμεση πρόσβαση στα ακατέργαστα σήματα.

Για την καταγραφή των δονήσεων, χρησιμοποιήθηκαν δύο μονοαξονικά επιταχυνσιόμετρα **KISTLER**, τύπου **8640A50 (εύρος ±50 g)** και **8704B500 (εύρος ±500 g)**. Για λόγους ασφαλείας και οικονομικής αποδοτικότητας, τα επιταχυνσιόμετρα τοποθετήθηκαν πίσω από τον πύργο εργαλείων με μαγνητική βάση, ευθυγραμμισμένα κατά τους άξονες X και Z. Η θέση τοποθέτησης δεν είναι αυστηρά κρίσιμη, αλλά απαιτείται σταθερότητα καθ' όλη τη διάρκεια των δοκιμών ώστε να διατηρηθούν σταθερά χαρακτηριστικά σήματος. Η δόνηση διαδίδεται μέσω της κινηματικής αλυσίδας του τόρνου, παράγοντας σήμα με συνιστώσες από όλους τους άξονες.

Τα σήματα των επιταχυνσιομέτρων συνδέθηκαν σε ενισχυτές **KISTLER Type 5118B2**, μέσω καλωδίων χαμηλού θορύβου. Για το επιταχυνσιόμετρο χαμηλότερης ευαισθησίας επιλέχθηκε ενίσχυση $10\times$ ώστε οι μετρήσεις και από τα δύο να βρίσκονται στην ίδια κλίμακα. Όλα τα σήματα (από κινητήρες και αισθητήρες) τροφοδοτήθηκαν σε σύστημα απόκτησης δεδομένων **ADVANTECH USB-4711A**, εξασφαλίζοντας συγχρονισμό και απευθείας μεταφορά σε υπολογιστή μέσω USB. Η καταγραφή έγινε με ρυθμό δειγματοληψίας **1 kHz**, ενώ χρησιμοποιήθηκε λογισμικό **DAQnavi** για τον έλεγχο της απόκτησης.

Αναλυτικά ο εξοπλισμός που χρησιμοποιήθηκε :

- **CNC Τόρνος:** Χρησιμοποιήθηκε τόρνος CNC τύπου **Okuma™ LB10II**, ο οποίος αποτελεί κέντρο τόρνευσης ακριβείας με σύστημα ελέγχου **OSP700L**. Ο τόρνος διαθέτει δύο κύριους άξονες κίνησης, τον **X (ακτινικό)** και τον **Z (διαμήκη)**, μέσω των οποίων

καθοδηγείται το κοπτικό εργαλείο κατά μήκος του περιστρεφόμενου τεμαχίου. Η άτρακτος διαθέτει μέγιστη ταχύτητα περιστροφής **10.000 rpm** και ο πύργος εργαλείων έχει 12 θέσεις, επιτρέποντας την αυτοματοποιημένη αλλαγή εργαλείων και την εκτέλεση πολύπλοκων κοπών με υψηλή ακρίβεια.

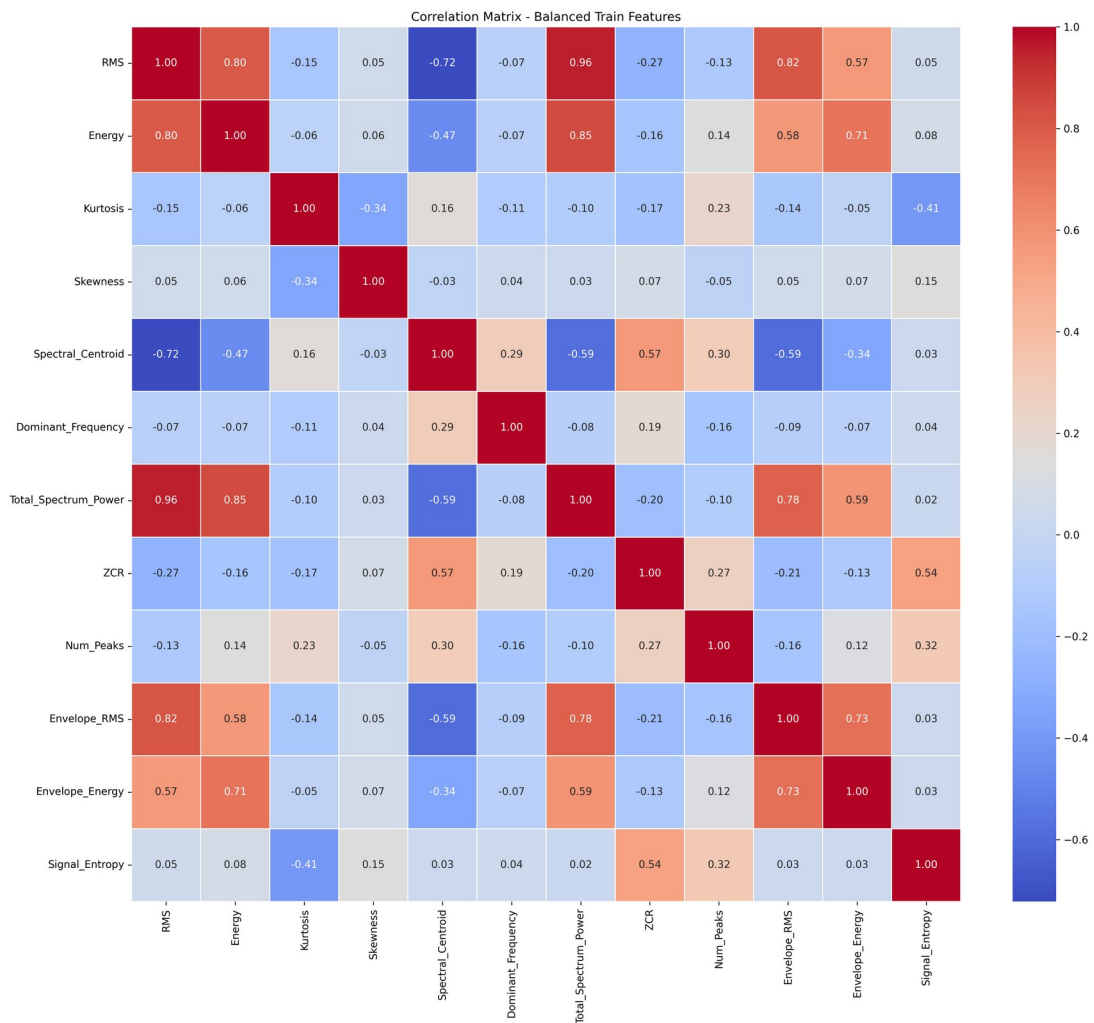
- **Κοπτικά εργαλεία και τεμάχια:** Για την κατεργασία χρησιμοποιήθηκαν κοπτικά ένθετα **Sandvik Coromant™ CNMG 12 04 08-MR 4305**, τοποθετημένα σε βάση εργαλείου **Sandvik PCLNR 2020K 12**. Το υλικό των τεμαχίων ήταν χάλυβας **CK45**, σε μορφή ράβδων διαμέτρου **32 mm** και μήκους **75 mm**, οι οποίες κατεργάστηκαν σε μήκος **45 mm**. Το συγκεκριμένο υλικό επιλέχθηκε λόγω της ευρείας χρήσης του σε εφαρμογές γενικής μηχανουργικής κατεργασίας.
- **Αισθητήρες:** Στην περιοχή κοντά στον πύργο εργαλείων τοποθετήθηκαν δύο μονοαξονικά επιταχυνσιόμετρα **Kistler 8640A50 (εύρος ±50 g)** και **Kistler 8704B500 (±500 g)**. Τα επιταχυνσιόμετρα στερεώθηκαν με μαγνητική βάση, προσανατολισμένα κατά τους άξονες X και Z, ώστε να καταγράφουν τις δονήσεις που δημιουργούνται από την κοπή. Η θέση (πίσω από το **turret**) τοποθέτησης επιλέχθηκε έτσι ώστε να ελαχιστοποιούνται παρεμβολές από το περιβάλλον και να διατηρείται σταθερό το σήμα καθ' όλη τη διάρκεια των δοκιμών, αλλά επίσης και για την προστασία των επιταχυνσιομέτρων από άμεση επαφή με το κοπτικό εργαλείο και τα ρινίσματα.
- **Παρακολούθηση στροφών και ροπής κινητήρων:** Ο τórνος δεν διαθέτει αυτόνομο σύστημα μέτρησης στροφών και ροπής προς εξωτερική έξοδο, για τον λόγο αυτό αντλήθηκαν αναλογικά σήματα απευθείας από την πλακέτα ελέγχου του CNC. Καταγράφηκαν η ταχύτητα περιστροφής της ατράκτου και οι γραμμικές ταχύτητες και ροπές των κινητήρων στους άξονες X και Z. Τα σήματα αυτά παρέχουν πληροφορία για τις συνιστώσες των δυνάμεων κοπής και το ενεργειακό φορτίο της κατεργασίας.
- **Σύστημα συλλογής δεδομένων:** Όλα τα αναλογικά σήματα από τους αισθητήρες, συνδέθηκαν σε πολυκαναλικό σύστημα απόκτησης δεδομένων **ADVANTECH USB-4711A**, το οποίο διασφαλίζει συγχρονισμένη καταγραφή σε ρυθμό δειγματοληψίας **1 kHz**. Για την ενίσχυση των σημάτων των επιταχυνσιομέτρων χρησιμοποιήθηκαν ενισχυτές **Kistler Type 5118B2** με ρύθμιση κέρδους $10\times$ όπου απαιτούνταν. Η απόκτηση και αποθήκευση των δεδομένων έγινε μέσω ειδικού λογισμικού **DAQNavi**, εγκατεστημένου σε υπολογιστή.

Ο προαναφερθείς εξοπλισμός λειτουργεί ως βάση για την παραγωγή δεδομένων προς ανάλυση και εξαγωγή χαρακτηριστικών, τα οποία με τη σειρά τους τροφοδοτούν τα μοντέλα **ML & NN** που θα στηθούν. Τα δεδομένα που συλλέγονται από τις δοκιμές τóρνευσης αποτελούν τη βάση για την ανάπτυξη των εν λόγω μοντέλων ως προς την απόδοση και την ποιότητα της κατεργασίας.



Σχήμα 3: Κατεργασία τόρνευσης

3.2 Συλλογή και Αρχική Επεξεργασία Σημάτων



Σχήμα 4: Ενδεικτικός Πίνακας Συσχέτισης Εξαγόμενων Χαρακτηριστικών (Correlation Matrix)

Η παρακολούθηση της ποιότητας τεμαχίου σε **CNC** τόρνευση απαιτεί τη συλλογή ακατέργαστων σημάτων από τον μηχανικό εξοπλισμό, με σκοπό την **εξαγωγή χαρακτηριστικών** που συνδέονται με την τραχύτητα επιφάνειας, τη διαστατική ακρίβεια και τελικά με την λειτουργική καταλληλότητα του τεμαχίου. Η απόκτηση αυτών των σημάτων βασίζεται στη χρήση **αισθητήρων δονήσεων και αναλογικών εξόδων από την πλακέτα ελέγχου του CNC**, επιτρέποντας τη **μη παρεμβατική παρακολούθηση** των συνθηκών κοπής.

Τα σήματα των αισθητήρων είναι **χρονοσειρές**, οι οποίες περιλαμβάνουν τόσο ωφέλιμες όσο και θορυβώδεις πληροφορίες. Η απευθείας εισαγωγή τέτοιων σημάτων σε αλγόριθμους μηχανικής μάθησης είναι ανεπιθύμητη λόγω **υπερδιάστασης (curse of dimensionality)**, **υψηλού υπολογιστικού κόστους** και πιθανής **υπερπροσαρμογής (overfitting)**.

Για τον λόγο αυτό, εφαρμόστηκε **εξαγωγή χαρακτηριστικών (feature extraction)** με σκοπό την **συμπύκνωση της πληροφορίας** σε αριθμητικά επεξεργάσιμες μεταβλητές, διατηρώντας παράλληλα τα διακριτικά γνωρίσματα κάθε σήματος [13].

3.2.1 Χαρακτηριστικά Χρόνου

Από το αρχικό σήμα επιταχύνσεων και ροπών υπολογίστηκαν βασικά στατιστικά:

- **Μέση τιμή (Mean)** και **τυπική απόκλιση (STD)** για τη μέτρηση κεντρικής τάσης και διασποράς.
- **RMS (Root Mean Square)**, χρήσιμος δείκτης για τη συνολική "ένταση" του σήματος.
- **Κυρτότητα (Kurtosis)** και **συμμετρία (Skewness)**, για την ανίχνευση peak ή ανωμαλιών [6].

3.2.2 Χαρακτηριστικά Συχνότητας

Στο πεδίο της συχνότητας εφαρμόστηκε **Fast Fourier Transform (FFT)** ώστε να εντοπιστούν:

- **Κυρίαρχες συχνότητες (Dominant Frequencies)**
- **Φασματική ενέργεια (Spectral Energy)**
- **Σχετική συμβολή κάθε ζώνης συχνοτήτων (Band Power Ratios)**

Επιπλέον, μετρήθηκαν δείκτες όπως:

- **Εντροπία σήματος (Signal Entropy)**
- **Ενέργεια περιβλήματος (Envelope Energy)**

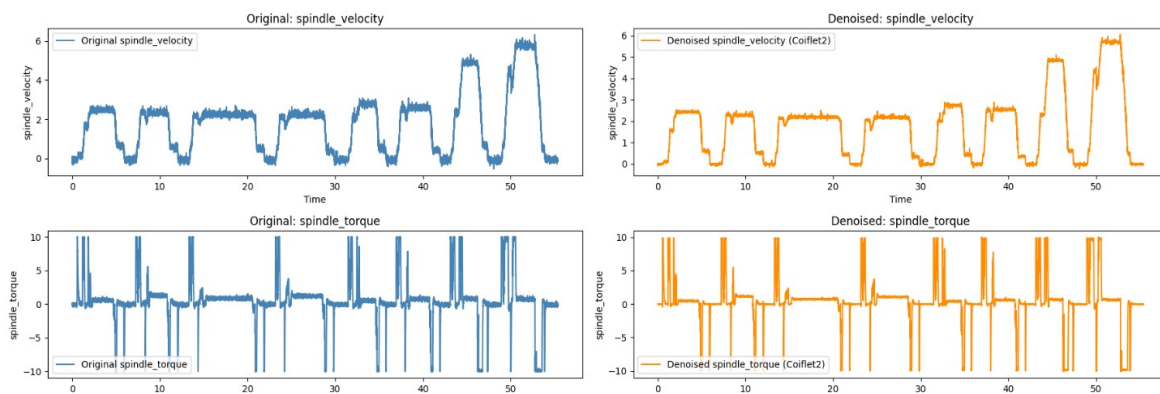
Τα **features** που εξήχθησαν θα αναφερθούν αναλυτικά στην συνέχεια.

Η επιλογή των χαρακτηριστικών βασίστηκε τόσο σε **βιβλιογραφικές κατευθύνσεις όσο και εμπειρικά**, με σκοπό τη διατήρηση των πιο πληροφοριακών και ανεξάρτητων μεταβλητών για κάθε σήμα.

Τα εξαγόμενα χαρακτηριστικά αποθηκεύτηκαν ανά δοκιμή και χρησιμοποιούνται για την ταξινόμηση ή παλινδρόμηση της ποιότητας κατεργασίας, όπως περιγράφεται στο επόμενο κεφάλαιο.

3.3 Αποθορυβοποίηση των σημάτων

Για την αξιόπιστη εξαγωγή χαρακτηριστικών από τα σήματα δονήσεων και ροπής που συλλέχθηκαν κατά τη διάρκεια της κοπής, εφαρμόστηκε μέθοδος αποθορυβοποίησης βασισμένη στην κυματιδική ανάλυση (**Wavelet Denoising**). Συγκεκριμένα, χρησιμοποιήθηκε το κυματιδίο **Coiflet2**, το οποίο παρουσιάζει καλή χρονική και συχνοτική εντοπισμότητα, προσφέροντας ισορροπία μεταξύ ομαλής αναπαράστασης και ακριβούς ανίχνευσης αιχμών. Η διαδικασία περιλάμβανε πολυεπίπεδη ανάλυση μέσω **Discrete Wavelet Transform (DWT)**, εφαρμογή τεχνικής **thresholding** στους λεπτομερείς συντελεστές για την απομάκρυνση ανεπιθύμητου θορύβου υψηλής συχνότητας και ανακατασκευή του σήματος μέσω **Inverse DWT**. Η αποθορυβοποίηση αυτή οδήγησε σε πιο καθαρά σήματα, γεγονός που βελτίωσε την ακρίβεια και τη σταθερότητα των υπολογισμένων χαρακτηριστικών (όπως **RMS**, **Energy**, **Envelope RMS**, **Entropy**), τα οποία αποτέλεσαν εισόδους για τα μοντέλα μηχανικής μάθησης και τα νευρωνικά δίκτυα της παρούσας μελέτης.



Σχήμα 5: Αποτελέσματα Αποθορυβοποίησης

Όπως φαίνεται ξεκάθαρα στο παραπάνω γράφημα, η χρονοσειρά (raw signal) περιλαμβάνει αρκετό θόρυβο, τον οποίο μέσω της μεθόδου αποθορυβοποίησης που επιλέχθηκε, επιτεύχθηκε ο περιορισμός του. Αυτή η λεπτομέρεια είναι πολύ σημαντική για την εξαγωγή χαρακτηριστικών, καθώς η πληροφορία του θορύβου θα επηρεάσει τα αποτελέσματα χωρίς ουσία.

3.4 Εξαγωγή Χαρακτηριστικών (Features)

Για την εκπαίδευση του μοντέλου αλλά και την κατηγοριοποίηση της ποιότητας του παραγόμενου τεμαχίου, υλοποιήθηκε ένα πλήρες σύνολο χαρακτηριστικών (**features**) που εξαγονται απευθείας από τα σήματα δονήσεων, ταχύτητας και ροπής της ατράκτου. Η εξαγωγή καλύπτει τόσο το χρονικό όσο και το συχνοτικό πεδίο, με στόχο να αποτυπωθούν συνολικά η

ένταση, η μεταβλητότητα, η περιοδικότητα και η αταξία του σήματος. Συγκεκριμένα, στο χρονικό πεδίο υπολογίστηκαν: η ενεργός τιμή (**RMS**), η ολική ενέργεια σήματος (**Energy**), ο συντελεστής κύρτωσης (**Kurtosis**), ο συντελεστής ασυμμετρίας (**Skewness**), ο ρυθμός διελεύσεων από το μηδέν (**Zero Crossing Rate, ZCR**), ο αριθμός κορυφών (**Number of Peaks**), το RMS του περιβλήματος σήματος (**Envelope RMS**), η ενέργεια του περιβλήματος (**Envelope Energy**) και η εντροπία σήματος (**Signal Entropy**) ως μέτρο αταξίας και πολυπλοκότητας. Στο πεδίο της συχνότητας, μέσω του μετασχηματισμού **FFT**, εξήχθησαν η επικρατέστερη συχνότητα (**Dominant Frequency**), το φασματικό κέντρο μάζας (**Spectral Centroid**) και η συνολική ισχύς φάσματος (**Total Spectrum Power**), που καταγράφουν πού συγκεντρώνεται ενεργειακά το φάσμα και ποια περιοδικά φαινόμενα υπερισχύουν.

Όλα τα παραπάνω χαρακτηριστικά υπολογίστηκαν για κάθε σήμα και για όλα τα διαθέσιμα κανάλια μέτρησης, συνθέτοντας ένα πλούσιο διάγραμμα χαρακτηριστικών για κάθε δείγμα. Πριν την εξαγωγή των χαρακτηριστικών εφαρμόστηκε σταθερή **τμηματοποίηση (segmentation)** των σημάτων σε χρονικά παράθυρα διάρκειας **0.5 δευτερολέπτων**, βάση των στηλών **start & end time** από τα **measurments**, ώστε να καταστεί δυνατή η αποτύπωση των χρονικά τοπικών μεταβολών της δυναμικής συμπεριφοράς του συστήματος. Η επιλογή των συγκεκριμένων χαρακτηριστικών βασίστηκε σε βιβλιογραφικά **[1] [3] [4] [6]** τεκμήρια και προηγούμενες μελέτες που αναδεικνύουν τη σημασία τους στην παρακολούθηση ταλαντωτικών και φασματικών φαινομένων κατά την κατεργασία. Τα εξαγόμενα αυτά χαρακτηριστικά αποτέλεσαν το πρωταρχικό σύνολο εισόδου για την εκπαίδευση και τη βελτιστοποίηση των μοντέλων μηχανικής μάθησης που εφαρμόστηκαν στο πλαίσιο της παρούσας εργασίας.

Αναλυτικά τα χαρακτηριστικά που εξήχθησαν :

Πεδίο του Χρόνου (Time Domain) :

- **RMS (Root Mean Square)**

Η RMS είναι η ενεργός τιμή του σήματος και εκφράζει το μέσο τετραγωνικό πλάτος των δονήσεων. Είναι ένας από τους βασικότερους δείκτες για το συνολικό δυναμικό φορτίο κατά την κοπή. Υψηλές τιμές RMS δείχνουν πιο έντονες δονήσεις, κάτι που μπορεί να σχετίζεται με επιθετικές συνθήκες κατεργασίας ή πιθανή κακή ποιότητα επιφάνειας.

- **Energy (Συνολική Ενέργεια Σήματος)**

Η συνολική ενέργεια δείχνει πόση ισχύς «μεταφέρεται» από τις δονήσεις στο εργαλείο και στο τεμάχιο. Υπολογίζεται ως το άθροισμα των τετραγώνων όλων των τιμών του σήματος. Υψηλή ενέργεια σημαίνει μεγαλύτερη καταπόνηση και μπορεί να υποδηλώνει σκληρότερα υλικά ή ακατάλληλες παραμέτρους κοπής.

- **Kurtosis (Συντελεστής Κύρτωσης)**

Η κύρτωση μετρά την «αιχμηρότητα» της κατανομής του σήματος. Ανιχνεύει σπάνιες αλλά έντονες κορυφές (κρούσεις) στο σήμα. Στη τόννευση, υψηλή κύρτωση υποδηλώνει ξαφνικά χτυπήματα ή αστάθειες, που επηρεάζουν την ποιότητα.

- **Skewness (Συντελεστής Ασυμμετρίας)**

Η ασυμμετρία δείχνει πόσο συμμετρική είναι η κατανομή των τιμών γύρω από τον μέσο όρο. Μη μηδενική ασυμμετρία φανερώνει τάση προς ακραίες τιμές από τη μία πλευρά. Είναι χρήσιμη για την ανίχνευση μη ισορροπημένων φορτίων ή τάσεων κατά την κατεργασία.

- **Zero Crossing Rate (ZCR)**

Το ZCR μετρά πόσο συχνά το σήμα αλλάζει πρόσημο. Είναι ένας απλός τρόπος να εκτιμήσουμε έμμεσα την παρουσία υψηλών συχνοτήτων. Υψηλό ZCR σημαίνει γρήγορες ταλαντώσεις, χαμηλό ZCR δείχνει πιο «ομαλή» κίνηση του εργαλείου.

- **Number of Peaks (Αριθμός Κορυφών)**

Αυτό το χαρακτηριστικό μετρά πόσες τοπικές κορυφές υπερβαίνουν ένα κατώφλι πλάτους. Αύξηση των κορυφών δείχνει έντονη ταλαντωτική ή κρουστική συμπεριφορά, βοηθώντας στον εντοπισμό ανωμαλιών ή φαινομένων chatter.

- **Envelope RMS (RMS Περιβλήματος)**

Το RMS του περιβλήματος του σήματος υπολογίζεται αφού αφαιρεθούν οι υψηλές συχνότητες και κρατηθεί το «γενικό περίγραμμα». Χρησιμεύει για την ανίχνευση πιο αργών δυναμικών φαινομένων, τα οποία μπορεί να κρύβονται στο raw σήμα.

- **Envelope Energy (Ενέργεια Περιβλήματος)**

Η ενέργεια του περιβλήματος αναδεικνύει την ένταση της χαμηλοσυχνής δυναμικής συνιστώσας. Συμπληρώνει το raw Energy και βοηθά να κατανοηθεί η συνολική δυναμική φόρτιση, ειδικά όταν συνυπάρχουν χαμηλές και υψηλές συχνότητες.

- **Signal Entropy (Εντροπία Σήματος)**

Η εντροπία εκφράζει το πόσο τυχαίο ή αταξινόμητο είναι το σήμα. Υψηλή εντροπία σημαίνει πολύπλοκη, απρόβλεπτη ταλάντωση. Η μέτρηση αυτή επιτρέπει την ανίχνευση χαοτικών ή «θορυβωδών» φαινομένων που επηρεάζουν την ποιότητα.

Πεδίο της Συχνότητας (Spectrum Domain) :

- **Dominant Frequency (Επικρατούσα Συχνότητα)**

Πρόκειται για τη συχνότητα με τη μεγαλύτερη ισχύ στο φάσμα FFT. Δείχνει το κύριο φαινόμενο ταλάντωσης του συστήματος. Πολύτιμο για την ανίχνευση περιοδικών φαινομένων όπως chatter ή resonance.

- **Spectral Centroid (Φασματικό Κέντρο Μάζας)**

Το κέντρο βάρους του φάσματος. Περιγράφει που συγκεντρώνεται ενεργειακά η ισχύς των συχνοτήτων. Χρησιμοποιείται για να εκτιμηθεί αν το σήμα είναι πιο «βαρύ» σε χαμηλές ή υψηλές συχνότητες.

- **Total Spectrum Power (Συνολική Ισχύς Φάσματος)**

Η ολική ενέργεια του σήματος στο συχνοτικό πεδίο. Επιβεβαιώνει το συνολικό δυναμικό φορτίο και επιτρέπει έλεγχο συνέπειας με την ενέργεια του χρονικού πεδίου μέσω της σχέσης Parseval.

Τα παραπάνω χαρακτηριστικά του χρόνου εξήχθησαν από τις χρονοσειρές που είχαμε στην διάθεση μας μέσω των αισθητήρων (και του φάσματος τους όταν τους εφαρμόσαμε μετασχηματισμό **Fourier**) συμπυκνώθηκε η πληροφορία των πολλών δειγμάτων του σήματος σε λίγες μόνο αριθμητικές τιμές διατηρώντας όμως την ουσία των σημαντικών **χρονικών μοτίβων** (ένταση, μεταβλητότητα, αιχμές) του σήματος. Οι τιμές αυτές τροφοδοτούν το μοντέλο ως είσοδοι και επιτρέπουν στο σύστημα μηχανικής μάθησης να αντιληφθεί το **συνολικό επίπεδο** και τη **συμπεριφορά** των ταλαντώσεων κατά την κοπή.

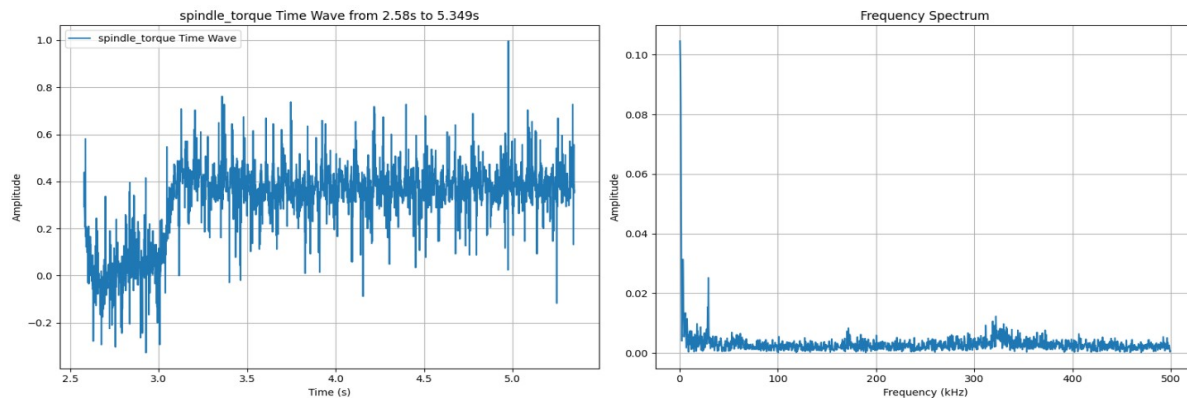
Τα φασματικά χαρακτηριστικά υπολογίστηκαν για τα ίδια χρονικά τμήματα σήματος, εφαρμόζοντας **FFT** σε κάθε χρονικό σήμα, εξάγοντας τιμές όπως παραπάνω. Έτσι, κάθε τμήμα κατεργασίας περιγράφεται τελικά από έναν συνδυασμό χαρακτηριστικών **χρονικού πεδίου** (π. χ. RMS, μέγιστη τιμή, ενέργεια, κ. λπ.) και **του πεδίου της συχνότητας** (π. χ. κορυφαία συχνότητα, centroid, εύρος ζώνης φάσματος).

Στην αρχική υλοποίηση δεν χρησιμοποιήθηκαν **όλα** τα features που αναφέρονται, αλλά κάποια από αυτά προστέθηκαν σε επόμενες υλοποιήσεις για βελτίωση της απόδοσης (**ZCR, Num_Peaks, Envelope_RMS, Envelope_Energy, Signal_Entropy**).

Ο εμπλουτισμός του συνόλου δεδομένων με χαρακτηριστικά και από τις δύο κατηγορίες εξασφαλίζει ότι το μοντέλο μηχανικής μάθησης λαμβάνει πληροφορία τόσο για το **γενικό επίπεδο και τις στατιστικές ιδιότητες των σημάτων**, όσο και για τις **συχνότητες-κλειδιά** που αντανακλούν συγκεκριμένα φαινόμενα της διεργασίας. Κάθε χαρακτηριστικό συμβάλλει διαφορετικά. Δηλαδή, οι χρονικές πληροφορίες όπως το RMS και η ενέργεια υποδεικνύουν πόσο έντονες είναι συνολικά οι δονήσεις, ενώ τιμές όπως η κύρτωση ή η ασυμμετρία μπορούν να εντοπίσουν ασυνήθιστα **spikes** στο σήμα. Από την άλλη πλευρά, τα φασματικά χαρακτηριστικά αποκαλύπτουν **περιοδικότητες ή συντονισμούς** που δεν φαίνονται άμεσα στο χρόνο, επιτρέποντας στο μοντέλο να αναγνωρίσει μοτίβα όπως την εμφάνιση δονήσεων σε χαρακτηριστικές συχνότητες.

Οι υπολογισμένες αυτές παράμετροι προστέθηκαν ως στήλες χαρακτηριστικών στο σύνολο δεδομένων (π. χ. στο αρχείο **balanced_combined_features_2.csv** που επεκτάθηκε με νέες στήλες για κάθε **feature**). Κατά την εκπαίδευση των μοντέλων, κάθε δείγμα διαδικασίας (δηλαδή κάθε τμήμα κοπής) αντιστοιχεί πλέον σε ένα διάνυσμα χαρακτηριστικών που περιλαμβάνει όλες τις παραπάνω τιμές. Με αυτόν τον τρόπο, τα μοντέλα μηχανικής μάθησης (π. χ. ένα **Random Forest Classifier** όπως χρησιμοποιήθηκε και σε αυτήν την εργασία) μπορούν να **εκπαιδευτούν** πάνω σε αυτά τα δεδομένα ώστε να μάθουν τη συσχέτιση των χαρακτηριστικών με την ποιότητα της κατεργασίας. Συνεπώς, η συμπερίληψη πλούσιων χαρακτηριστικών χρονικού και συχνοτικού πεδίου ενισχύει την απόδοση των μοντέλων, διότι τους παρέχει μια καλύτερη “εικόνα” της κατάστασης του συστήματος. Οι δοκιμές που ακολούθησαν έδειξαν ότι τα χαρακτηριστικά βοήθησαν το μοντέλο να διακρίνει πιο αποτελεσματικά περιπτώσεις που ήταν δυσδιάκριτες με

μόνο απλές στατιστικές, βελτιώνοντας τελικά την **ακρίβεια πρόγνωσης** της ποιότητας και επιτρέποντας πιο έγκαιρο εντοπισμό φθοράς ή ανωμαλιών στη διεργασία.



Σχήμα 6: Παράδειγμα Μετασχηματισμού Fourier (FFT) σε Χρονοσειρά

3.5 Labels

Για την ανάπτυξη και αξιολόγηση των μοντέλων, χρησιμοποιήθηκε ως **ετικέτα στόχου (label)** η μεταβλητή **Roughness_Ra_Class**, η οποία αντιπροσωπεύει την **ποιοτική κατηγοριοποίηση της τραχύτητας επιφάνειας** με βάση τη μετρούμενη τιμή **Ra (μέση τραχύτητα)**. Η κλάση αυτή κατατάσσει κάθε δείγμα σε μία από τις τρεις κατηγορίες ποιότητας: **χαμηλή, μέση ή υψηλή τραχύτητα**, αντιστοιχώντας έτσι σε μια εκτίμηση της ποιότητας κατεργασίας του τεμαχίου.

Η ετικέτα δεν ήταν εξ αρχής διαθέσιμη στο αρχικό σύνολο σημάτων. Προέκυψε **μέσω διαδικασίας κατηγοριοποίησης πάνω σε ανεξάρτητα αρχεία μετρήσεων (L01 και L02)**, τα οποία περιείχαν τις φυσικές μετρήσεις που ελήφθησαν εκτός τόρνου, όπως:

- roughness_Ra
- roughness_Rz
- diameter_deviation
- actual_diameter
- feed
- speed
- depth
- RPM

Τα δεδομένα αυτά **ήταν χρονικά επισημασμένα**, με πεδία **start_time** και **end_time**, υποδεικνύοντας το χρονικό διάστημα κατεργασίας κάθε τεμαχίου. Για τη σύνδεση των ποιοτικών μετρήσεων με τα σήματα, αλλά και για την επαύξηση των **λιγοστών** δεδομένων που ήταν διαθέσιμα, **αναπτύχθηκαν συναρτήσεις που “έσπαγαν” τα σήματα σε χρονικά παράθυρα (segments)** διάρκειας **0.5 δευτερολέπτων**, όπως προαναφέρθηκε. Σε κάθε παράθυρο,

αντιστοιχούσε μια ποιοτική ετικέτα βάσει των τιμών **Ra**. Οι συναρτήσεις αυτές υλοποιούσαν **κανόνες κατωφλίου (threshold-based rules)** για την ταξινόμηση της τραχύτητας **Ra** σε τρεις ποιοτικές κλάσεις (π.χ. $Ra \leq 2 \rightarrow \text{Low}$, $2 < Ra \leq 4.25 \rightarrow \text{Medium}$, $Ra > 4.25 \rightarrow \text{High}$), σύμφωνα με πρότυπα ανοχών και τεχνολογικές απαιτήσεις.

Κατ' αυτόν τον τρόπο, κάθε υποσύνολο σημάτων που αντιστοιχούσε σε συγκεκριμένο χρονικό τμήμα, **εμπλουτίστηκε με ετικέτα ποιότητας**, επιτρέποντας τη δημιουργία πλήρως εποπτευόμενου **dataset** για εκπαίδευση μοντέλων μηχανικής μάθησης. Η διαδικασία αυτή αποτέλεσε κρίσιμο βήμα, καθώς συνέδεσε **απλά χρονικά σήματα με πραγματικές μετρούμενες τιμές ποιότητας**, προσδίδοντας σημαντική σημασιολογική αξία στα δεδομένα.

Αν και στην παρούσα μελέτη αξιοποιήθηκε **ένα μόνο ποιοτικό μέγεθος** — η μέση τραχύτητα επιφάνειας (**Roughness_Ra**) — ως **έξοδος** για την εκπαίδευση του μοντέλου, αυτό δεν αποτελεί περιορισμό ως προς τη δυνατότητα επέκτασης του συστήματος. Συγκεκριμένα, το **Ra** επιλέχθηκε λόγω της ισχυρής του συσχέτισης με την αισθητική και λειτουργική ποιότητα του κατεργασμένου τεμαχίου, όπως επιβεβαιώνεται σε πληθώρα μελετών [1] [6] [11] [18]. Η αριθμητική του τιμή κατηγοριοποιήθηκε σε τρεις διακριτές ποιοτικές ετικέτες (**Low, Medium, High**), διαμορφώνοντας έτσι ένα πρόβλημα μονο-εξόδου ταξινόμησης με **τρεις κατηγορίες (multiclass classification)**.

Ωστόσο, η χρήση ενός μόνο **target label** δεν αναιρεί τη δυνατότητα για **πολυδιάστατη** μάθηση. Η διαθέσιμη βάση δεδομένων περιλαμβάνει επιπλέον κρίσιμες μεταβλητές, όπως **Roughness_Rz, Diameter_Deviation, Actual_Diameter**, καθώς και παραμέτρους κατεργασίας όπως **Feed, Speed, Depth of Cut, RPM**, που μπορούν να αξιοποιηθούν είτε ως παράλληλες έξοδοι (σε **multi-output regression**) είτε ως επιπλέον στόχοι ταξινόμησης (**multi-label classification**). Τέτοιες προσεγγίσεις υποστηρίζονται βιβλιογραφικά και αποσκοπούν σε πολυπαραγοντική αξιολόγηση της ποιότητας κατεργασίας [6] [11] [13].

Η επιλογή του **Ra** ως μοναδικού **label** βασίζεται στη διεθνή πρακτική, καθώς θεωρείται πρότυπο μέτρο για την αξιολόγηση της ποιότητας επιφάνειας σύμφωνα και με το **ISO 4287 [19]**, ενώ και πολλές εφαρμογές μηχανικής μάθησης το χρησιμοποιούν ως βασικό εξαρτημένο μεταβλητό [1] [6] [7] [12].

Συγκεκριμένα, η επιλογή αυτή έγινε για τους εξής λόγους :

1. Απλούστερη Υλοποίηση.
2. Απαιτούνται **λιγότερα δείγματα** για να επιτευχθεί καλή ακρίβεια.
3. Ιδανικό εάν ο στόχος είναι να βγει μια συνολική εκτίμηση (π. χ. «καλή ποιότητα», όχι πολλά μεμονωμένα μεγέθη).

Όμως : Παρέχει περιορισμένη πληροφορία (μία μόνο πτυχή της ποιότητας) σε αντίθεση με πολλά target labels.

Απο την άλλη πλευρά , σχετικά με τα πολλά labels :

1. Προβλέπει περισσότερες πληροφορίες ταυτόχρονα.
2. Είναι ένα πολύ πιο γενικό σύστημα.

Όμως :

1. Απαιτεί μια αρκετά πιο περίπλοκη αρχιτεκτονική
2. Δύσκολη βαθμονόμηση & **loss function**
3. Πιο δύσκολη αξιολόγηση και ερμηνεία
4. Περισσότερα δεδομένα ανά label για καλή ακρίβεια (κάτι το οποίο δεν είχαμε)

Συνεπώς, με βάση αυτά τα δεδομένα για τα πλεονεκτήματα και τα μειονεκτήματα της κάθε κατηγορίας, καταλήξαμε στο να κινηθούμε σε λύση με ένα **label**.

4 Μοντέλα Μηχανικής Μάθησης

Εφόσον τα σήματα μετατραπούν σε ένα πίνακα χαρακτηριστικών και προετοιμαστούν κατάλληλα, είναι πλέον έτοιμα να χρησιμοποιηθούν για την εκπαίδευση του μοντέλου. Στην παρούσα εργασία υλοποιήθηκαν δύο κατηγορίες μοντέλων: (α) Ένα **συμβατικό μοντέλο ML - Random Forest Classifier** και (β) ένα **Νευρωνικό Δίκτυο (Πολλαπλών Επιπέδων)**.

4.1 Random Forest Classifier - RFC

- **Random Forest Classifier – RFC:** Τα τυχαία δάση είναι μέθοδοι συλλογικής μάθησης (**ensemble learning**) που αποτελούνται από μεγάλο αριθμό δέντρων αποφάσεων. Κατά την ταξινόμηση, κάθε δέντρο παράγει μια «ψήφο» για την κατηγορία πρόβλεψης και το **Random Forest** αποφασίζει τελικά προβλέποντας την κατηγορία που έλαβε τις περισσότερες ψήφους από τα επιμέρους δέντρα (**πλειοψηφική απόφαση**). Αυτό το σύνολο απλών ταξινομητών (**δέντρων**) προσφέρει υψηλή ακρίβεια και ανθεκτικότητα σε θόρυβο ή υπερεκπαίδευση, επειδή τα λάθη των μεμονωμένων δέντρων τείνουν να ακυρώνονται αμοιβαία. Στην περίπτωση μας, το **Random Forest** αποδείχθηκε κατάλληλο λόγω των πολλών χαρακτηριστικών εισόδου και της πιθανής μη γραμμικότητάς τους – τα δέντρα μπορούν να μοντελοποιήσουν πολύπλοκες αλληλεπιδράσεις μεταξύ χαρακτηριστικών χωρίς ισχυρές παραδοχές. Πράγματι, αναφορές δείχνουν ότι τα **ensemble** δέντρων όπως το **RF** είναι αποτελεσματικά σε προβλήματα παρακολούθησης εργαλείων, συχνά ξεπερνώντας μεμονωμένα δέντρα ή απλούστερους αλγορίθμους.

Κατα την εκπόνηση της εργασίας δοκιμάστηκαν διάφορες παραμετροποιήσεις στο μοντέλο :

- **n_estimators (Αριθμός Δέντρων)**
Ορίζει πόσα δέντρα θα περιλαμβάνονται στο δάσος. Περισσότερα δέντρα αυξάνουν τη σταθερότητα του αποτελέσματος αλλά και τον υπολογιστικό χρόνο. Επιλέχθηκε τιμή 100 για επαρκή ακρίβεια χωρίς υπερβολική καθυστέρηση.
- **max_depth (Μέγιστο Βάθος Δέντρου)**
Επιτρέπει σε κάθε δέντρο να επεκταθεί όσο χρειάζεται για να μοντελοποιήσει τα δεδομένα πλήρως. Δεν επιβλήθηκε όριο, ώστε να διατηρηθεί υψηλή ευκαμψία.
- **random_state (Τυχαίος Σπόρος)**
Ορίστηκε σταθερός αριθμός για πλήρη επαναληψιμότητα των αποτελεσμάτων και δυνατότητα άμεσης σύγκρισης μεταξύ πειραμάτων.
- **criterion (Κριτήριο Διάσπασης)**
Προσδιορίζει τον τρόπο μέτρησης της «καθαρότητας» σε κάθε κόμβο. Χρησιμοποιήθηκε η **Gini impurity** καθώς είναι αποδοτική και αξιόπιστη για ταξινομήσεις πολλαπλών κλάσεων.

- **min_samples_split (Ελάχιστα Δείγματα για Διάσπαση)**
Ορίζει τον ελάχιστο αριθμό δειγμάτων που απαιτούνται για να διασπαστεί ένας κόμβος. Κρατήθηκε στην προεπιλεγμένη τιμή για ισορροπία μεταξύ υπερεκπαίδευσης και υποεκπαίδευσης.
- **min_samples_leaf (Ελάχιστα Δείγματα σε Φύλλο)**
Ορίζει το ελάχιστο πλήθος δειγμάτων που πρέπει να έχει κάθε τελικό φύλλο. Επίσης διατηρήθηκε στη default τιμή ώστε να διασφαλιστεί ευκαμψία.
- **bootstrap (Δειγματοληψία με Επανατοποθέτηση)**
Ενεργό. Επιτρέπει σε κάθε δέντρο να εκπαιδεύεται σε διαφορετικό υποσύνολο δεδομένων, προσφέροντας ποικιλία και ανθεκτικότητα στο θόρυβο.

Σε κάθε μια από τις 3 υλοποιήσεις χρησιμοποιήθηκαν διαφορετικά χαρακτηριστικά υπερπαραμέτρων, θα σχολιαστεί αναλυτικά στην συνέχεια.

(Εκτός από το **RFC**, δοκιμάστηκαν και άλλοι συμβατικοί ταξινομητές για λόγους σύγκρισης, όπως π. χ. **SVM**, απλά Δέντρα Απόφασης, k-πλησιέστεροι γείτονες (**k-NN**) και μεθοδοί παλινδρόμησης/ταξινόμησης **Bayes**. Ωστόσο, τα αποτελέσματά τους δεν παρουσίασαν υπεροχή έναντι του **Random Forest** και έτσι δεν δόθηκε έμφαση σε αυτά.)

4.2 Multi-Layer Perceptron -MLP

Πολυεπίπεδο πυκνό νευρωνικό δίκτυο (Multilayer Perceptron): Πρόκειται για ένα τεχνητό νευρωνικό δίκτυο προώθησης (**feed-forward ANN**) με πολλαπλά κρυφά επίπεδα πλήρως διασυνδεδεμένων νευρώνων. Τα τεχνητά νευρωνικά δίκτυα (**ANN**) είναι εμπνευσμένα από τα βιολογικά νευρωνικά συστήματα, αποτελούμενα από μονάδες (**neurons**) που δέχονται εισόδους, εφαρμόζουν μια μη-γραμμική συναρτησιακή μετασχηματισμό (συνήθως συνάρτηση ενεργοποίησης) και μεταδίδουν το αποτέλεσμα στα επόμενα επίπεδα μέσω βαρύνσεων. Κατά την εκπαίδευση, οι βαρύνσεις αυτές προσαρμόζονται με βάση έναν αλγόριθμο όπως η οπισθοδιάδοση σφάλματος (**Backpropagation**), έτσι ώστε το δίκτυο να μάθει τη συσχέτιση εισόδων-εξόδων. Στην υλοποίησή που πραγματοποιήθηκε, το πυκνό νευρωνικό δίκτυο έλαβε ως εισόδους τα εξαγόμενα χαρακτηριστικά σήματος και εκπαιδεύτηκε να αποδίδει την κατάσταση της κατεργασίας (π. χ. κανονική λειτουργία ή κάποιο σφάλμα/φθορά εργαλείου). Τα δίκτυα αυτά λειτουργούν ως καθολικοί προσεγγιστές συναρτήσεων και μπορούν να μοντελοποιήσουν εξαιρετικά περίπλοκες σχέσεις μεταξύ των χαρακτηριστικών, εφόσον υπάρχουν επαρκή δεδομένα εκπαίδευσης. Στην περίπτωση μας, το δίκτυο διαμορφώθηκε με μερικές κρυφές στρώσεις και κατάλληλο αριθμό νευρώνων σε κάθε στρώση (βασισμένο και σε δοκιμές επικύρωσης) ώστε να επιτυγχάνει καλή γενίκευση χωρίς υπερεκπαίδευση. Μετά την εκπαίδευση, το μοντέλο **MLP** πέτυχε υψηλή ακρίβεια στην ταξινόμηση των συνθηκών κατεργασίας, καταδεικνύοντας την ικανότητα των νευρωνικών δικτύων να συλλαμβάνουν τα σύνθετα πρότυπα στα δεδομένα αισθητήρων. Για το νευρωνικό δίκτυο χρησιμοποιήθηκε αρχιτεκτονική **Multi-Layer Perceptron (MLP)**, υλοποιημένη με το Keras και TensorFlow. Το δίκτυο περιλαμβάνει

τρία πλήρως συνδεδεμένα κρυφά επίπεδα, **dropout** για **regularization** και **softmax** έξοδο για **multi-class** ταξινόμηση. Η εκπαίδευση έγινε με κατηγορική cross-entropy και προσαρμοσμένα **class weights** για το πρόβλημα ανισορροπίας.

Οι υπερπαραμέτροι που έγιναν **tuning** κατά την διάρκεια των δοκιμών ήταν οι εξής :

- **hidden_layer_sizes**: Ορίστηκαν δύο επίπεδα κρυφών νευρώνων με 100 νευρώνες το καθένα. Η επιλογή αυτή προσφέρει επαρκή υπολογιστική ισχύ για την εκμάθηση μη γραμμικών σχέσεων μεταξύ των χαρακτηριστικών εισόδου και των κατηγοριών εξόδου, χωρίς να επιβαρύνει υπερβολικά το χρόνο εκπαίδευσης.
- **activation**: Χρησιμοποιήθηκε η συνάρτηση ενεργοποίησης **ReLU (Rectified Linear Unit)**, η οποία είναι η πλέον δημοφιλής επιλογή για νευρωνικά δίκτυα, καθώς επιτρέπει γρήγορη και αποδοτική εκπαίδευση και μειώνει τον κίνδυνο **vanishing gradient**.
- **solver**: Για την προσαρμογή των βαρών εφαρμόστηκε ο αλγόριθμος βελτιστοποίησης **Adam**. Ο **Adam** συνδυάζει τα πλεονεκτήματα του **Adaptive Gradient Algorithm** και του **RMSProp**, παρέχοντας σταθερή σύγκλιση ακόμη και σε προβλήματα με θορυβώδη δεδομένα.
- **max_iter**: Η μέγιστη διάρκεια εκπαίδευσης (αριθμός εποχών) αυξήθηκε πάνω από το προεπιλεγμένο όριο των 200 εποχών, εξασφαλίζοντας ότι το δίκτυο έχει επαρκή χρόνο για να φτάσει σε ελάχιστο σφάλμα και να συγκλίνει.
- **random_state**: Ορίστηκε τιμή **seed** στον τυχαίο αρχικοποιητή βαρών, ώστε τα αποτελέσματα να είναι αναπαραγωγίμα σε επαναλαμβανόμενες δοκιμές και να διευκολύνεται η αξιολόγηση μοντέλων υπό τα ίδια αρχικά δεδομένα.

Το πώς ορίστηκαν οι υπερπαραμέτροι του Νευρωνικού Δικτύου με βάση την υλοποίηση θα αναλυθεί παρακάτω.

Τα προαναφερθέντα μοντέλα και τεχνικές επιλογής χαρακτηριστικών συνδέονται άμεσα με την ανάγκη της βιομηχανίας για αξιόπιστη **επιτήρηση ποιότητας κατεργασμένου τεμαχίου απο τόρνευση**. Με τον κατάλληλο συνδυασμό εξαγωγής χαρακτηριστικών (που ανιχνεύουν σημάδια φθοράς ή ανωμαλιών στα σήματα) και μεθόδων ταξινόμησης/πρόβλεψης (που οδηγούν αυτά τα σημάδια σε αποφάσεις για την ποιότητα του προϊόντος), μπορούμε να έχουμε ένα σύστημα **παρακολούθησης κατάστασης εργαλείου (Tool Condition Monitoring – TCM)**. Με τον τρόπο αυτό, τα μοντέλα μηχανικής μάθησης που παρουσιάστηκαν μπορούν να ενταχθούν σε ένα έξυπνο βιομηχανικό σύστημα (Industry 4. 0) που θα **βελτιώνει την παραγωγικότητα** μειώνοντας την ανάγκη ελέγχου ποιότητας του τεμαχίου (αφού θα προβλέπει το σύστημα τεχνητής νοημοσύνης).

4.3 Υλοποιήσεις

Στο πλαίσιο της παρούσας εργασίας αναπτύχθηκαν τρεις διακριτές υλοποιήσεις μοντέλων μηχανικής μάθησης και τεχνητών νευρωνικών δικτύων, με στόχο την ταξινόμηση της ποιότητας

επιφάνειας βάσει χαρακτηριστικών που προέκυψαν από την επεξεργασία σημάτων και μετρήσεων από τον τόρνο. Αν και οι φάσεις της προεπεξεργασίας δεδομένων διατηρήθηκαν σταθερές, οι επιμέρους υλοποιήσεις διαφοροποιούνται ως προς τη στρατηγική εξισορρόπησης των κλάσεων και τις επιλεγμένες αρχιτεκτονικές και παραμετροποιήσεις των μοντέλων, κυρίως του **Random Forest Classifier (RFC)** και του **Multilayer Perceptron (MLP)**.

- **Υλοποίηση 1:** Στην αρχική αυτή φάση, τα δύο βασικά μοντέλα τεχνητής νοημοσύνης (**RFC και MLP**) εκπαιδεύτηκαν πάνω σε ένα μη εξισορροπημένο σύνολο δεδομένων, χρησιμοποιώντας ένα πρώτο σύνολο χαρακτηριστικών που προέκυψαν μέσω χειροκίνητης εξαγωγής χαρακτηριστικών (**RMS, Energy, Kurtosis, Skewness, Spectral Centroid, Dominant Frequency, Total Spectrum Power**). Πραγματοποιήθηκαν αρχικές δοκιμές παραμετροποίησης των μοντέλων με στόχο τη βελτιστοποίηση της απόδοσής τους.
- **Υλοποίηση 2:** Στη δεύτερη προσέγγιση εφαρμόστηκε εξισορρόπηση των κλάσεων στο αρχικό **dataset**, κυρίως μέσω της τεχνικής επαναδειγματοληψίας (**upsampling**), ενώ πραγματοποιήθηκαν και πειραματισμοί με τη μέθοδο **SMOTE** για συγκριτικούς λόγους. Παράλληλα, δοκιμάστηκαν επιπλέον υπερπαραμετροποιήσεις των μοντέλων, οδηγώντας σε αξιοσημείωτη βελτίωση της ακρίβειας και της γενικευσιμότητας.
- **Υλοποίηση 3:** Με σκοπό την περαιτέρω ενίσχυση της αποδοτικότητας, στη φάση αυτή υλοποιήθηκε μια νέα διαδικασία εξαγωγής χαρακτηριστικών, αξιοποιώντας μετρικές απευθείας από τα σήματα (**ZCR, Num Peaks, Envelope RMS, Envelope Energy, Signal Entropy**). Επιπλέον, εφαρμόστηκε εκτεταμένο tuning των μοντέλων, μεταξύ των οποίων και χρήση του **GridSearchCV** για την εύρεση βέλτιστων υπερπαραμέτρων στο **Random Forest Classifier**.

Κάθε υλοποίηση σχεδιάστηκε ως φυσική συνέχεια της προηγούμενης, με στόχο τη σταδιακή βελτίωση της ταξινομητικής ικανότητας των μοντέλων. Για τον λόγο αυτό, οι τεχνικές εξαγωγής χαρακτηριστικών και οι διαδικασίες βελτιστοποίησης δεν εφαρμόστηκαν εξ αρχής στο σύνολό τους, αλλά εισήχθησαν προοδευτικά, ώστε να διασφαλιστεί η κατανόηση της επίδρασης κάθε παράγοντα στην τελική απόδοση.

Τα αποτελέσματα που προέκυψαν από την κάθε υλοποίηση, θα αναλυθούν στην συνέχεια.

4.4 Προεπεξεργασία και Προετοιμασία Δεδομένων

Η προετοιμασία των δεδομένων αποτέλεσε κρίσιμο στάδιο για την επιτυχή εφαρμογή των μοντέλων μηχανικής μάθησης. Αρχικά, συλλέχθηκαν τα δεδομένα σήματος (μετρήσεις όπως δονήσεις, ροπές, ταχύτητες, κ.λπ.) από τις κατεργασίες που πραγματοποιήθηκαν στα πειράματα L01 και L02, μέσω των δύο επιταχυνσιομέτρων και το ελεγκτή της μηχανής όπως αναλύθηκε προηγουμένως. Παράλληλα, καταγράφηκαν και οι πραγματικές μετρήσεις ποιότητας του τεμαχίου (π.χ. τραχύτητα Ra, Rz, πραγματική διάμετρος, απόκλιση διαμέτρου) σε συγκεκριμένα χρονικά διαστήματα (start_time, end_time) σε ξεχωριστά αρχεία.

Για να καταστεί δυνατός ο συσχετισμός των σημάτων με τις ποιοτικές μετρήσεις, εφαρμόστηκε **segmentation** των σημάτων βάσει των χρονικών παραθύρων που δόθηκαν στα αρχεία μετρήσεων. Κάθε σήμα τεμαχίστηκε σε επιμέρους τμήματα διάρκειας 0.5 δευτερολέπτων, όπως προκύπτει από την εφαρμογή συναρτήσεων που βασίζονται σε **start** και **end time**. Έτσι, για κάθε χρονικό **segment**, δημιουργήθηκε ένα σύνολο χαρακτηριστικών (**feature vector**) που αντιστοιχεί σε εκείνο το χρονικό διάστημα της κατεργασίας.

Από κάθε **segment** εξάχθηκαν στατιστικά, δυναμικά και φασματικά χαρακτηριστικά (**mean, RMS, ZCR, peaks, entropy, energy, envelope RMS κ.ά.**), ξεχωριστά για κάθε κανάλι μέτρησης (π.χ. **x_vibration, z_vibration, spindle torque κ.λπ.**). Στη συνέχεια, τα χαρακτηριστικά αυτά αποτέλεσαν την είσοδο για την εκπαίδευση των μοντέλων.

Η κατηγορική ετικέτα στόχου (**label**) **Roughness_Ra_Class** αποδόθηκε σε κάθε δείγμα, με βάση την τιμή της τραχύτητας **Ra** στην οποία αντιστοιχούσε το χρονικό παράθυρο. Οι τιμές **Ra** κατηγοριοποιήθηκαν με βάση κατάλληλα **thresholds** σε τρεις κατηγορίες: **Low, Medium** και **High**, μέσω κατάλληλης συνάρτησης **label assignment**, η οποία εφαρμόστηκε στα αρχεία μετρήσεων.

Πριν την εκπαίδευση των μοντέλων εφαρμόστηκαν τα εξής βήματα προεπεξεργασίας:

- **Καθαρισμός δεδομένων:** Αφαιρέθηκαν στήλες όπως **Signal, Measurement_File, start_time, end_time**, καθώς δεν προσέφεραν πληροφορία για την πρόβλεψη.
- **Κανονικοποίηση:** Εφαρμόστηκε **StandardScaler** για την κανονικοποίηση όλων των αριθμητικών χαρακτηριστικών ώστε να έχουν μηδενική μέση τιμή και μοναδιαία διασπορά.
- **Κωδικοποίηση ετικετών:** Οι κατηγορικές ετικέτες (**Low, Medium, High**) μετατράπηκαν σε αριθμητικές τιμές μέσω **Label Encoding**, και ακολούθως σε **one-hot vectors** για χρήση σε νευρωνικά δίκτυα.
- **Διαχωρισμός σε training και test set:** Χωρίστηκε το **dataset** σε **training** και **test set** σε αναλογία **80:20** με χρήση **stratified sampling** για διατήρηση της κατανομής των ετικετών.
- **Εξισορρόπηση δεδομένων:** Σε μια περίπτωση εφαρμόστηκε **upsampling** των μειοψηφικών κλάσεων μόνο στο **training set** ή χρήση **SMOTE**, προκειμένου να μειωθεί η επίδραση της ανισορροπίας των τάξεων κατά την εκπαίδευση.

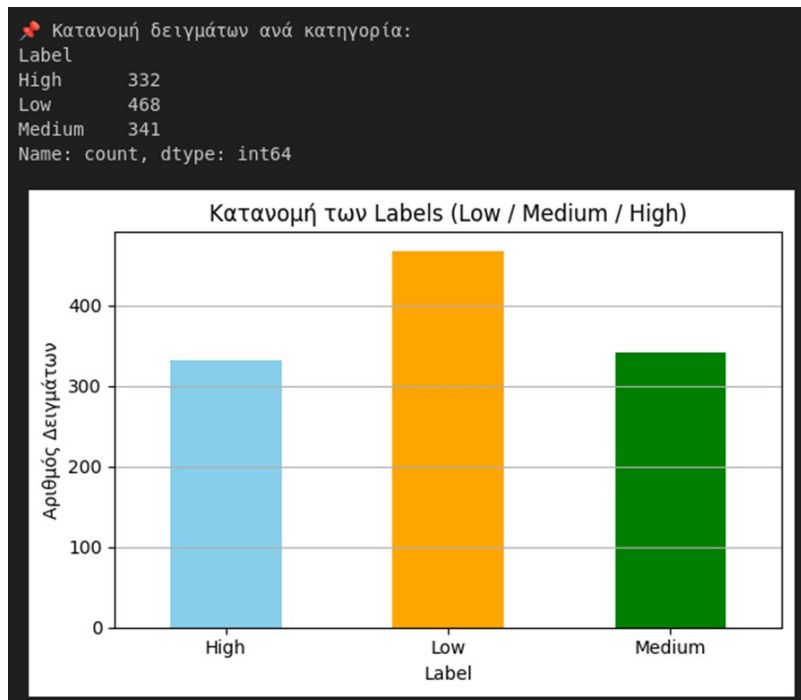
Αξίζει να σημειωθεί ότι για λόγους αξιοπιστίας και ρεαλισμού, η εξισορρόπηση εφαρμόστηκε **μόνο στο training set**, ενώ το **test set** διατηρήθηκε στην αρχική του μορφή, ώστε τα τελικά αποτελέσματα αξιολόγησης να αντικατοπτρίζουν ρεαλιστικά την επίδοση των μοντέλων.

Για κάθε ξεχωριστή υλοποίηση που πραγματοποιήθηκε, δημιουργήθηκαν ξεχωριστά **csv** για να γίνει η εκπαίδευση.

4.5 1^η Υλοποίηση (Imbalanced Data):

Η αρχική υλοποίηση βασίστηκε σε ένα μη εξισορροπημένο **dataset** (Σχήμα 7), το οποίο περιείχε άνιση κατανομή δειγμάτων μεταξύ των τριών ποιοτικών κλάσεων του **label (Low, Medium, High)**. Το σύνολο αυτό προήλθε από τα αρχεία μετρήσεων (**L01 & L02**) (Σχήμα 8), στα οποία καταγράφονται τα χρονικά παράθυρα κατεργασίας (**start-end time**) και οι αντίστοιχες ποιοτικές μετρήσεις (**roughness, διάμετρος, συνθήκες κοπής**). Οι εγγραφές αυτών των αρχείων αντιστοιχούν σε τμήματα σήματος διάρκειας περίπου **0,5 δευτερολέπτων**.

Για κάθε τέτοιο τμήμα πραγματοποιήθηκε εξαγωγή χαρακτηριστικών από τα αντίστοιχα σήματα δονήσεων και άλλων μετρούμενων μεγεθών, δημιουργώντας ένα ενιαίο dataset με χαρακτηριστικά όπως: **RMS, Energy, Kurtosis, Skewness, Spectral Centroid, Dominant Frequency, Total Spectrum Power**. Η ετικέτα στόχου (**Roughness_Ra_Class**) αποδόθηκε σε κάθε εγγραφή με βάση το χρονικό παράθυρο, πραγματοποιώντας συσχέτιση μεταξύ των αρχείων σημάτων και των μετρήσεων κατεργασιών. Το αρχείο **combined_extracted_features_with_labels.csv** (Σχήμα 9) αποτέλεσε την κύρια πηγή δεδομένων για την υλοποίηση αυτή.



Σχήμα 7: Διάγραμμα Απεικόνισης Class Imbalance

```

1 pass_id,measurement_location,start_end_time,roughness_Ra,roughness_Rz,diameter_deviation,actual_diameter,feed,speed,depth,RPM,Roughness_Ra_Class,Roughness_Rz_Class
2 Roughing 1,L1,(2.58, 3.08),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
3 Roughing 1,L1,(3.08, 3.58),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
4 Roughing 1,L1,(3.58, 4.08),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
5 Roughing 1,L1,(4.08, 4.58),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
6 Roughing 1,L1,(4.58, 5.08),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
7 Roughing 1,L1,(5.08, 5.58),1.31,6.8,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,Low,High,Low,Low,Low
8 Roughing 1,L2,(5.35, 5.85),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
9 Roughing 1,L2,(5.85, 6.35),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
10 Roughing 1,L2,(6.35, 6.85),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
11 Roughing 1,L2,(6.85, 7.35),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
12 Roughing 1,L2,(7.35, 7.85),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
13 Roughing 1,L2,(7.85, 8.12),1.28,6.77,-0.004,30.996,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
14 Roughing 1,L3,(8.121, 8.621),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
15 Roughing 1,L3,(8.621, 9.121),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
16 Roughing 1,L3,(9.121, 9.621),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
17 Roughing 1,L3,(9.621, 10.121),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
18 Roughing 1,L3,(10.121, 10.621),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
19 Roughing 1,L3,(10.621, 10.89),1.27,6.67,-0.005,30.995,0.2,160.0,0.5,1642.89,Low,Low,High,Low,Low,Low,Low
20 Roughing 2,L1,(14.992, 15.492),2.72,11.7,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
21 Roughing 2,L1,(15.492, 15.992),2.72,11.7,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
22 Roughing 2,L1,(15.992, 16.492),2.72,11.7,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
23 Roughing 2,L2,(16.436, 16.936),2.96,11.93,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
24 Roughing 2,L2,(16.936, 17.436),2.96,11.93,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
25 Roughing 2,L2,(17.436, 17.879),2.96,11.93,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
26 Roughing 2,L3,(18.38, 18.38),2.72,11.3,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
27 Roughing 2,L3,(18.38, 18.88),2.72,11.3,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
28 Roughing 2,L3,(18.88, 19.323),2.72,11.3,0.024,29.024,0.3,190.0,1.0,2085.48,Medium,Medium,Medium,High,Medium,Medium,Low
29 Roughing 3,L1,(23.69, 24.19),3.89,14.5,0.035,26.035,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
30 Roughing 3,L1,(24.19, 24.449),3.89,14.5,0.035,26.035,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
31 Roughing 3,L2,(24.45, 24.95),3.91,13.83,0.035,26.035,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
32 Roughing 3,L2,(24.95, 25.21),3.91,13.83,0.035,26.035,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
33 Roughing 3,L3,(25.211, 25.711),4.04,14.7,0.036,26.036,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
34 Roughing 3,L3,(25.711, 25.971),4.04,14.7,0.036,26.036,0.4,220.0,1.5,2693.39,Medium,Medium,Medium,High,Medium,Medium,Medium
35 Roughing 4,L1,(30.073, 30.525),4.73,16.37,0.043,22.043,0.5,250.0,2.0,3617.16,High,High,Medium,Medium,High,High,High,Medium
36 Roughing 4,L2,(30.526, 30.978),4.85,17.37,0.038,22.038,0.5,250.0,2.0,3617.16,High,High,Medium,Medium,High,High,High,Medium
37 Roughing 4,L3,(30.979, 31.431),5.15,17.67,0.038,22.038,0.5,250.0,2.0,3617.16,High,High,Medium,Medium,High,High,High,Medium
38 Roughing 5,L1,(36.361, 36.861),3.31,11.93,0.025,21.025,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
39 Roughing 5,L1,(36.861, 37.219),3.31,11.93,0.025,21.025,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
40 Roughing 5,L2,(37.22, 37.72),2.7,10.67,0.025,21.025,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
41 Roughing 5,L2,(37.72, 38.077),2.7,10.67,0.025,21.025,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
42 Roughing 5,L3,(38.078, 38.578),2.66,11.03,0.022,21.022,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
43 Roughing 5,L3,(38.578, 38.936),2.66,11.03,0.022,21.022,0.3,220.0,0.5,3334.67,Medium,Medium,Medium,Medium,Medium,Low,Medium
44 Roughing 6,L1,(43.544, 44.044),1.28,7.03,0.008,19.008,0.2,250.0,1.0,4188.29,Low,Low,Medium,Medium,Low,High,Medium,Medium
45 Roughing 6,L1,(44.044, 44.544),1.28,7.03,0.008,19.008,0.2,250.0,1.0,4188.29,Low,Low,Medium,Medium,Low,High,Medium,Medium
46 Roughing 6,L1,(44.544, 44.580),1.28,7.03,0.008,19.008,0.2,250.0,1.0,4188.29,Low,Low,Medium,Medium,Low,High,Medium,Medium
47 Roughing 6,L2,(44.587, 45.087),1.37,6.87,0.006,19.006,0.2,250.0,1.0,4188.29,Low,Low,Medium,Medium,Low,High,Medium,Medium
48 Roughing 6,L2,(45.087, 45.587),1.37,6.87,0.006,19.006,0.2,250.0,1.0,4188.29,Low,Low,Medium,Medium,Low,High,Medium,Medium

```

Σχήμα 8: classified_expanded_part_L01_measurements.csv & classified_expanded_part_L02_measurements.csv

```

1 Signal,start_time,end_time,RMS_Energy,Kurtosis,Spectral_Centroid,Dominant_Frequency,Total_Spectrum_Power,Label,Measurement_File
2 z_vibration,2.58,3.08,0.04304927948597869,0.9284734729552148,21.725585932934,0.018259494826071428,226.5453949280164,0.0,0.0068724030805991226,Low,classified_expanded_part_L01_measurements.csv
3 x_vibration,2.58,3.08,0.0456901676530226,0.9510270701702227,16.09547966460546,0.733131323490237,214.6919762555262,0.0,0.00670631940404165,Low,classified_expanded_part_L01_measurements.csv
4 spindle_velocity,5.08,5.08,0.15623256410940498,1222.871568187866,7.312918796054541,-0.895848798148206,53.3058919870294,0.0,0.75995939688766,Low,classified_expanded_part_L01_measurements.csv
5 spindle_torque,2.58,3.08,0.1327546529874288,0.792424201965332,1.7912301180511981,-0.5693537286269255,212.06323010777248,0.0,0.04044799230169473,Low,classified_expanded_part_L01_measurements.csv
6 z_velocity,2.58,3.08,0.13918592981957587,9.705734292929688,183.53398123226896,10.072371660677174,200.36147020435283,0.0,0.071401923210306605,Low,classified_expanded_part_L01_measurements.csv
7 z_torque,2.58,3.08,0.35731945669761656,63.96627426147461,0.632657573252761,-0.4710219599992173,123.658706468682,0.0,0.5041651666136274,Low,classified_expanded_part_L01_measurements.csv
8 x_velocity,2.58,3.08,0.07277025731458823,2.653050422668457,23.157794884077585,-0.24616244997707676,198.6377208997106,0.0,0.0180790662031389,Low,classified_expanded_part_L01_measurements.csv
9 x_torque,2.58,3.08,0.1964549105821785,19.334006395092777,0.694520330952349,-0.72500532996658,158.80187407984087,0.0,0.146330303925546,Low,classified_expanded_part_L01_measurements.csv
10 y_vibration,3.08,3.08,0.0658093031580669,2.12119968795776367,1.5221934968418022,-0.1614823160782946,259.72243725904,0.0,0.01146609939330742,Low,classified_expanded_part_L01_measurements.csv
11 x_vibration,3.08,3.08,0.0741130586307405,2.7518749237608547,0.4408772176930772,-0.00486203130262173,241.9529064334752,0.0,0.012480730147943671,Low,classified_expanded_part_L01_measurements.csv
12 spindle_velocity,3.08,3.08,0.15576918704163938,1215.6283855438232,7.440639178695451,-0.589100537906877,49.0115123150755,0.0,0.702039718920434,Low,classified_expanded_part_L01_measurements.csv
13 spindle_torque,3.08,3.08,0.39897940583066677,79.75146770477295,4.91095592186611,-0.37665988392267496,155.7809811509528,0.0,0.620952988039668,Low,classified_expanded_part_L01_measurements.csv
14 z_velocity,3.08,3.08,0.14056467609895872,9.898972511291504,17.584153803397015,-0.2954028196971289,175.1392093168465,0.0,0.0769567993192205,Low,classified_expanded_part_L01_measurements.csv
15 z_torque,3.08,3.08,0.42184792879655697,89.15579319000244,1.464810921146853,-0.1288461011349102,121.58921062218679,0.0,0.7048307197403609,Low,classified_expanded_part_L01_measurements.csv
16 x_velocity,3.08,3.08,0.0743399743860982,2.6285886746526367,0.672719822248498,0.758331195679383,221.5274789189828,0.0,0.017056123346559517,Low,classified_expanded_part_L01_measurements.csv
17 x_torque,3.08,3.08,0.228881411908334,26.245737075805664,2.97640498938547,0.514346556612369,180.4596516444018,0.0,0.28348647995640366,Low,classified_expanded_part_L01_measurements.csv
18 y_vibration,3.58,4.08,0.06411695179772217,2.059602737426758,1.1306197133545783,0.1438766541461212,242.523255084186,0.0,0.011586656166455,Low,classified_expanded_part_L01_measurements.csv
19 x_vibration,3.58,4.08,0.07404862606751898,2.7470827102661133,1.02188543454372,0.07565987938932729,232.958103691526,0.0,0.0131336898177246,Low,classified_expanded_part_L01_measurements.csv
20 spindle_velocity,3.58,4.08,0.15623252172736977,1222.9131698606398,15.789761289996943,0.2766325854045367,52.29404336423843,0.0,0.760175736728224,Low,classified_expanded_part_L01_measurements.csv
21 spindle_torque,3.58,4.08,0.35731945669761656,63.96627426147461,0.632657573252761,-0.4710219599992173,123.658706468682,0.0,0.5041651666136274,Low,classified_expanded_part_L01_measurements.csv
22 z_velocity,3.58,4.08,0.13918592981957587,9.705734292929688,183.53398123226896,10.072371660677174,200.36147020435283,0.0,0.071401923210306605,Low,classified_expanded_part_L01_measurements.csv
23 z_torque,3.58,4.08,0.42184792879655697,89.15579319000244,1.464810921146853,-0.1288461011349102,121.58921062218679,0.0,0.7048307197403609,Low,classified_expanded_part_L01_measurements.csv
24 x_velocity,3.58,4.08,0.08110353261178586,3.29549828405717,99.94450018687544,3.66280812648325,225.7426632955518,0.0,0.020669544744405768,Low,classified_expanded_part_L01_measurements.csv
25 x_torque,3.58,4.08,0.229704119870281,26.43475532513783,4.40367272930235,0.2178724872944783,180.0325596084305,0.0,0.2053549759375373,Low,classified_expanded_part_L01_measurements.csv
26 z_vibration,4.08,4.08,0.061548184549526726,1.89769260803588867,1.909779318136109,0.2841342016382897,232.87390643274733,0.0,0.010560991315094126,Low,classified_expanded_part_L01_measurements.csv
27 x_vibration,4.08,4.08,0.07261870511271606,2.6420116424560547,0.6256172012674238,-0.1036279266702088,240.0372893231984,0.0,0.012139498954183956,Low,classified_expanded_part_L01_measurements.csv
28 spindle_velocity,4.08,4.08,0.15623252172736977,1222.9131698606398,15.789761289996943,0.2766325854045367,52.29404336423843,0.0,0.760175736728224,Low,classified_expanded_part_L01_measurements.csv
29 spindle_torque,4.08,4.08,0.35731945669761656,63.96627426147461,0.632657573252761,-0.4710219599992173,123.658706468682,0.0,0.5041651666136274,Low,classified_expanded_part_L01_measurements.csv
30 z_velocity,4.08,4.08,0.13918592981957587,9.705734292929688,183.53398123226896,10.072371660677174,200.36147020435283,0.0,0.071401923210306605,Low,classified_expanded_part_L01_measurements.csv
31 z_torque,4.08,4.08,0.42184792879655697,89.15579319000244,1.464810921146853,-0.1288461011349102,121.58921062218679,0.0,0.7048307197403609,Low,classified_expanded_part_L01_measurements.csv
32 x_velocity,4.08,4.08,0.08110353261178586,3.29549828405717,99.94450018687544,3.66280812648325,225.7426632955518,0.0,0.020669544744405768,Low,classified_expanded_part_L01_measurements.csv
33 x_torque,4.08,4.08,0.229704119870281,26.43475532513783,4.40367272930235,0.2178724872944783,180.0325596084305,0.0,0.2053549759375373,Low,classified_expanded_part_L01_measurements.csv
34 y_vibration,4.08,4.08,0.061548184549526726,1.89769260803588867,1.909779318136109,0.2841342016382897,232.87390643274733,0.0,0.010560991315094126,Low,classified_expanded_part_L01_measurements.csv
35 x_vibration,4.08,4.08,0.07261870511271606,2.6420116424560547,0.6256172012674238,-0.1036279266702088,240.0372893231984,0.0,0.012139498954183956,Low,classified_expanded_part_L01_measurements.csv
36 spindle_velocity,4.08,4.08,0.15623252172736977,1222.9131698606398,15.789761289996943,0.2766325854045367,52.29404336423843,0.0,0.760175736728224,Low,classified_expanded_part_L01_measurements.csv
37 spindle_torque,4.08,4.08,0.35731945669761656,63.96627426147461,0.632657573252761,-0.4710219599992173,123.658706468682,0.0,0.5041651666136274,Low,classified_expanded_part_L01_measurements.csv
38 z_velocity,4.08,4.08,0.13918592981957587,9.705734292929688,183.53398123226896,10.072371660677174,200.36147020435283,0.0,0.071401923210306605,Low,classified_expanded_part_L01_measurements.csv
39 z_torque,4.08,4.08,0.42184792879655697,89.15579319000244,1.464810921146853,-0.1288461011349102,121.58921062218679,0.0,0.7048307197403609,Low,classified_expanded_part_L01_measurements.csv
40 x_velocity,4.08,4.08,0.08110353261178586,3.29549828405717,99.94450018687544,3.66280812648325,225.7426632955518,0.0,0.020669544744405768,Low,classified_expanded_part_L01_measurements.csv
41 x_torque,4.08,4.08,0.229704119870281,26.43475532513783,4.40367272930235,0.2178724872944783,180.0325596084305,0.0,0.2053549759375373,Low,classified_expanded_part_L01_measurements.csv
42 y_vibration,4.08,4.08,0.061548184549526726,1.89769260803588867,1.909779318136109,0.2841342016382897,232.87390643274733,0.0,0.010560991315094126,Low,classified_expanded_part_L01_measurements.csv
43 x_vibration,4.08,4.08,0.07261870511271606,2.6420116424560547,0.6256172012674238,-0.1036279266702088,240.0372893231984,0.0,0.012139498954183956,Low,classified_expanded_part_L01_measurements.csv
44 spindle_velocity,4.08,4.08,0.15623252172736977,1222.9131698606398,15.789761289996943,0.2766325854045367,52.29404336423843,0.0,0.760175736728224,Low,classified_expanded_part_L01_measurements.csv
45 spindle_torque,4.08,4.08,0.35731945669761656,63.96627426147461,0.632657573252761,-0.4710219599992173,123.658706468682,0.0,0.5041651666136274,Low,classified_expanded_part_L01_measurements.csv
46 z_velocity,4.08,4.08,0.13918592981957587,9.705734292929688,183.53398123226896,10.072371660677174,200.36147020435283,0.0,0.071401923210306605,Low,classified_expanded_part_L01_measurements.csv
47 z_torque,4.08,4.08,0.42184792879655697,89.15579319000244,1.464810921146853,-0.1288461011349102,121.58921062218679,0.0,0.7048307197403609,Low,classified_expanded_part_L01_measurements.csv
48 x_velocity,4.08,4.08,0.08110353261178586,3.29549828405717,99.94450018687544,3.66280812648325,225.7426632955518,0.0,0.020669544744405768,Low,classified_expanded_part_L01_measurements.csv

```

Σχήμα 9: combined_extracted_features_with_labels.csv

4.6 2^η Υλοποίηση (Manually Balanced Data):

Στην συγκεκριμένη υλοποίηση, έγινε χειροκίνητη εξισορρόπηση των δεδομένων για καλύτερα αποτελέσματα στην απόδοση των Μοντέλων. Για την αντιμετώπιση του προβλήματος ανισορροπίας μεταξύ των τριών κατηγοριών της ετικέτας (**Low, Medium, High**), εφαρμόστηκε χειροκίνητη εξισορρόπηση μέσω επαναδειγματοληψίας (**manual upsampling**). Συγκεκριμένα, το αρχικό σύνολο δεδομένων φορτώθηκε από το αρχείο **combined_extracted_features_with_labels.csv** και τα δείγματα διαχωρίστηκαν σε τρεις επιμέρους ομάδες ανάλογα με την κατηγορία τους. Στη συνέχεια, προσδιορίστηκε το πλήθος δειγμάτων της πολυπληθέστερης κλάσης και ορίστηκε αυτό ως στόχος για όλες τις κλάσεις. Για τις λιγότερο αντιπροσωπευμένες κατηγορίες εφαρμόστηκε τεχνική επαναδειγματοληψίας με αντικατάσταση (**resample**), ώστε να δημιουργηθούν επιπλέον δείγματα μέχρι να επιτευχθεί ο ίδιος αριθμός παρατηρήσεων σε κάθε κατηγορία. Τα ισορροπημένα σύνολα συνενώθηκαν σε ένα νέο σύνολο δεδομένων, το οποίο αναμίχθηκε τυχαία (**shuffle**) για να διασφαλιστεί η ομοιογένεια. Το τελικό εξισορροπημένο dataset αποθηκεύτηκε σε νέο αρχείο **csv (balanced_combined_features.csv)**, το οποίο χρησιμοποιήθηκε για τα επόμενα στάδια εκπαίδευσης των μοντέλων μηχανικής μάθησης. Με αυτόν τον τρόπο διασφαλίστηκε δεν υπερισχύει κάποια κλάση λόγω μεγαλύτερης παρουσίας στο **training set** και μπορεί να μάθει με πιο σωστό τρόπο τα χαρακτηριστικά που διαφοροποιούν κάθε επίπεδο ποιότητας

Εκτός από την προαναφερθείσα μέθοδο, κατά τη διάρκεια της ανάπτυξης του μοντέλου, δοκιμάστηκε και η τεχνική εξισορρόπησης **SMOTE (Synthetic Minority Over-sampling Technique)**. Η μέθοδος **SMOTE** δημιουργεί τεχνητά νέα δείγματα για τις λιγότερες σε αριθμό κλάσεις, παράγοντας συνθετικά σημεία στο χώρο χαρακτηριστικών αντί να αντιγράφει υπάρχοντα δείγματα. Για το πείραμα, εφαρμόστηκε αρχικά το **SMOTE** μόνο στο **training set** (μετά το **train-test split**) ώστε να αποφευχθεί η διαρροή πληροφοριών. Ωστόσο, στις συγκριτικές δοκιμές παρατηρήθηκε ότι η απόδοση του μοντέλου **Random Forest Classifier** με **SMOTE** δεν παρουσίασε καλύτερα αποτελέσματα έναντι της πιο απλής στρατηγικής **upsampling** με **resampling**. Σε κάποιες περιπτώσεις, το **SMOTE** αύξησε τον κίνδυνο **overfitting**, καθώς η κατασκευή συνθετικών σημείων σε μικρά σύνολα δεδομένων μπορεί να αλλοιώσει την πραγματική κατανομή. Για τον λόγο αυτό, η τελική επιλογή για την παραγωγή του **balanced dataset** έγινε με τη μέθοδο επαναδειγματοληψίας (**resampling**), η οποία κρίθηκε πιο αξιόπιστη και σταθερή για τα συγκεκριμένα σήματα και τα μεγέθη των διαθέσιμων δεδομένων.

Αυτό που απεικονίζεται στο Σχήμα 10, είναι το **dataset** που χρησιμοποιήθηκε για τις εκπαιδεύσεις των μοντέλων για την συγκεκριμένη υλοποίηση.

4.7 3^η Υλοποίηση (Νέο Feature Extraction):

Σε αυτήν την υλοποίηση, με βάση το εξισορροπημένο **balanced_combined_features.csv**, για λόγους περαιτέρω βελτίωσης της απόδοσης των μοντέλων, έγινε εκ νέου **feature extraction**

εξάγοντας νέα χαρακτηριστικά που θα προσφέρουν πληροφορία στο σύστημα τεχνητής νοημοσύνης που θα συμβάλει στην συνολική του βελτίωση.

Κατα αυτόν τον τρόπο, φτάσαμε στην δημιουργία του **balanced_combined_features_2.csv**, το οποίο χρησιμοποιήθηκε για την εκπαίδευση σε αυτήν την υλοποίηση.

```

1 | Signal,start_time,end_time,RMS_Energy,Kurtosis_Skewness,Spectral_Centroid,Dominant_Frequency,Total_Spectrum_Power,Label,Measurement_File
2 | Z_torque,8.714,9.214,0.567903601671382,161.352229267466,1.3486273095959184,1.065254493152323,96.10589670799452,0.0,1.273807091626363,High,classified_expanded
3 | Z_torque,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded
4 | spindle_velocity,10.621,10.89,1.5411058385261147,641.2519454950555,4.929898621383062,-0.8779271242647511,33.688197119286514,0.0,9.497017832325922,Low,classified_expanded
5 | spindle_torque,45.085,45.127,0.8994153357963026,17.15869990356445,3.376797515531692,-1.7895363024363482,45.24289188025301,0.0,2.611529848056062,Low,classified_expanded
6 | Z_torque,46.13,46.63,0.5526901797732529,153.04206663510989,0.0735068532666458,-0.128476048197961,106.6569817208868,0.0,1.2175811884923542,Low,classified_expanded
7 | spindle_torque,10.651,10.868,1.2149346757052348,321.78244608303546,7.594831854405937,-2.953723850684068,81.26921102362658,0.0,5.988852981223638,High,classified_expanded
8 | spindle_torque,2.58,3.08,0.1324754652987428,8.792424201965332,1.7912381180511987,0.5693537286269255,212.0632301077248,0.0,0.0404479923019647,Low,classified_expanded
9 | spindle_velocity,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded
10 | X_torque,24.19,24.449,0.153045662485707,6.089973449707031,0.084154756726605,-0.3619599181199976,163.7893734814139,0.0,0.0908985386603541,Medium,classified_expanded
11 | X_torque,30.073,30.525,0.094720852478522,4.069590588820801,3.369401027083014,-0.0840290117686547,215.7316291074724,0.0,0.0320530974665924,High,classified_expanded
12 | Z_torque,51.14,51.167,0.772097426072577,17.006699042373926,2.14244038036364,-0.9841181947454408,38.117889470526805,0.0,2.4234369778654,High,classified_expanded
13 | Z_vibration,51.369,51.869,0.1445875165827054,10.473680525737304,-0.815019734903981,0.0156796713870998,239.3727252916984,0.0,0.045029272666067,Low,classified_expanded
14 | Z_torque,3.08,3.58,0.4218479287956569,89.15579319080244,1.464818921146853,-0.128461061134916,121.589216221068,0.0,0.704839719740369,Low,classified_expanded
15 | Z_velocity,51.668,51.694,0.4169484351369601,4.693841934204102,0.2371812991124415,-0.222778036269918,18.66077671108729,0.0,0.6949601030260673,High,classified_expanded
16 | spindle_torque,16.08,17.36,0.845808425839343,358.4766904312134,5.239880874188568,0.0114642303679181,112.42514789169677,0.0,8.521161027033,High,classified_expanded
17 | spindle_torque,52.089,52.568,0.743123380969023,276.6766230051392,7.769480287382337,1.0200865270532325,126.7407411110414,0.0,2.4046903717470,Low,classified_expanded
18 | X_torque,39.168,39.668,0.1805409031099677,12.912464180283305,8.85253010228045,0.0318210941017895,101.74584363219677,0.0,0.099654460469183,Medium,classified_expanded
19 | Z_torque,38.078,38.578,0.4834109540158607,117.07658767700195,2.347573674147964,-0.4586184048150899,125.7686707257129,0.0,0.929135579416297,Medium,classified_expanded
20 | X_velocity,32.837,33.337,0.069782464874331,2.4396657943725586,35.9735454594573,7.3740672944641,214.48832408910203,0.0,0.016689131742797,High,classified_expanded
21 | X_torque,2.169,2.669,0.17339397626317,15.05311993180272,3.13150775628624,-0.195603784083382,178.13942457985998,0.0,0.1149758054858225,Medium,classified_expanded
22 | Z_velocity,44.587,45.087,0.2510751461949568,31.58240324744629,12.777349908209111,-1.1206258487264444,140.71942277750594,0.0,2.580148405405777,Low,classified_expanded
23 | X_torque,38.078,38.578,0.4834109540158607,117.07658767700195,2.347573674147964,-0.4586184048150899,125.7686707257129,0.0,0.929135579416297,Medium,classified_expanded
24 | spindle_velocity,26.555,27.055,2.23995606715102,2513.495946480201,10.540472562315542,2.2508992636969523,44.425795154137816,0.0,20.06284164149276,Medium,classified_expanded
25 | X_velocity,45.587,45.629,0.062665767473912,0.1721143734301757,1.2320270337052149,6.245631466216293,77.46476724584846,0.0,0.01772347665166,Low,classified_expanded
26 | X_velocity,10.121,10.621,0.086424854460609,3.74209808288574,84.1136544627824,7.167993615816638,234.8262130989888,0.0,0.02105809599176,Low,classified_expanded
27 | X_torque,9.121,9.621,0.2313727438379886,26.8202664215088,17.394336066295903,1.6709943675332586,169.6471272686291,0.0,2099244894280616,Low,classified_expanded
28 | X_velocity,37.72,38.077,0.07564601502719,2.033376693725586,42.5076824391265,-4.708792609805535,211.9734914543225,0.0,0.018049319096625,Medium,classified_expanded
29 | spindle_velocity,25.266,25.766,2.2412931201436344,2516.728828051997,9.341371887499148,1.2451125098318934,42.79539921536253,0.0,20.08980173075405,Medium,classified_expanded
30 | X_torque,18.36,18.767,0.2540299843814391,16.9843196888965,3.8512757295837607,0.1504298078267622,155.6653659907095,0.0,1639572084981265,Low,classified_expanded
31 | Z_torque,2.669,3.023,0.0281069534007979,74.84421730041504,-0.4026598727683153,-0.163740757426504,101.82913109400354,0.0,0.835395404989594,Medium,classified_expanded
32 | X_velocity,46.03,46.672,0.107211178022013,0.4952669158374023,35.342576042378674,7.957238716401239,190.72135109526735,0.0,0.032482571525292,Low,classified_expanded
33 | X_torque,32.837,33.337,0.099759905577352,4.985380172729492,10.028827308471129,1.30596357424124,280.34819557447733,0.0,0.036589395126451,High,classified_expanded
34 | Z_velocity,10.936,17.436,0.2064554407826355,21.148181011740725,9.69270928292323,-1.401717011845066,146.5374984778417,0.0,1.074147276906,Medium,classified_expanded
35 | Z_torque,46.17,46.212,0.7052420713474726,21.38676643715707,-0.3020992564797636,-0.21694286603179923,38.7045029326902,0.0,1.9821008115214777,Medium,classified_expanded
36 | X_velocity,43.544,44.044,0.065502202696118,2.15347070864318848,76.06079147518956,-0.7270707093211265,200.1542717729215,0.0,0.015352579630979,Low,classified_expanded
37 | spindle_velocity,26.055,26.555,2.28668458779185,2510.811925004176,6.678744132736446,-0.20768880604536,37.729510441751056,0.0,20.042952265604,Medium,classified_expanded
38 | Z_torque,15.992,16.492,0.5566403358502391,137.57271779108732,1.360880049818402,-0.4259925362695239,102.38210988451756,0.0,1.236019804979184,Medium,classified_expanded
39 | X_vibration,44.544,44.586,0.244868552994345,2.578360203180586,-0.1870138310926883,-0.2546113289236912,248.67943075903423,69.7674416064247,0.0,107383271826222,Low,classified_expanded
40 | X_torque,19.768,20.268,0.24876552562538,21.234758747680664,5.31713588210215,0.63757540594308,166.8415476782253,0.0,0.0176741788312952,High,classified_expanded
41 | Z_torque,2.669,3.023,0.0281069534007979,74.84421730041504,-0.4026598727683153,-0.163740757426504,101.82913109400354,0.0,0.835395404989594,Medium,classified_expanded
42 | X_torque,9.432,9.932,0.475176951784377,3.093028004807129,6.09195934072404,-0.373297126248428,213.7710478590975,0.0,0.0203445954919947,High,classified_expanded
43 | Z_vibration,5.08,5.349,0.0592044257178391,0.9489536285400392,2.261560739673057,0.3537746690377311,237.49070499458568,0.0,0.01011752925977,Low,classified_expanded
44 | X_torque,51.168,51.668,0.772097426072577,17.006699042373926,2.14244038036364,-0.9841181947454408,38.117889470526805,0.0,2.4234369778654,High,classified_expanded
45 | spindle_velocity,51.17,51.368,0.768520875363844,6621.88948115262,16.88029359460175,2.3936349842766758,13.863146322354837,0.0,103.0975181315653,Low,classified_expanded
46 | Z_torque,32.837,33.337,0.069782464874331,2.4396657943725586,35.9735454594573,7.3740672944641,214.48832408910203,0.0,0.016689131742797,High,classified_expanded
47 | spindle_torque,45.085,45.127,0.8994153357963026,17.15869990356445,3.376797515531692,-1.7895363024363482,45.24289188025301,0.0,2.611529848056062,Low,classified_expanded
48 | X_vibration,5.85,6.35,0.076991379364403,2.8948545455932617,0.7827886158497,-0.032019146550881,228.76959597344,0.0,0.14026416778564,Low,classified_expanded
49 | Z_vibration,44.544,44.586,0.244868552994345,2.578360203180586,-0.1870138310926883,-0.2546113289236912,248.67943075903423,69.7674416064247,0.0,107383271826222,Low,classified_expanded

```

Σχήμα 10: balanced_combined_features.csv

```

1 | Signal,start_time,end_time,RMS_Energy,Kurtosis_Skewness,Spectral_Centroid,Dominant_Frequency,Total_Spectrum_Power,Label,Measurement_File_X,ZCR_Num_Peaks,Envelope_RMS,En
2 | Z_torque,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded_part1
3 | Z_torque,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded_part1
4 | Z_torque,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded_part1
5 | spindle_velocity,10.621,10.89,1.5411058385261147,641.2519454950555,4.929898621383062,-0.8779271242647511,33.688197119286514,0.0,9.497017832325922,Low,classified_expanded_part1
6 | Z_torque,46.13,46.63,0.5526901797732529,153.04206663510989,0.0735068532666458,-0.128476048197961,106.6569817208868,0.0,1.2175811884923542,Low,classified_expanded_part1
7 | spindle_torque,10.651,10.868,1.2149346757052348,321.78244608303546,7.594831854405937,-2.953723850684068,81.26921102362658,0.0,5.988852981223638,High,classified_expanded_part1
8 | spindle_torque,2.58,3.08,0.1324754652987428,8.792424201965332,1.7912381180511987,0.5693537286269255,212.0632301077248,0.0,0.0404479923019647,Low,classified_expanded_part1
9 | spindle_velocity,51.668,51.694,0.7910195321334118,16.894221305847168,0.0452225819986265,-0.1008883225759016,33.41814250542509,0.0,2.490725918359624,High,classified_expanded_part1
10 | X_torque,24.19,24.449,0.153045662485707,6.089973449707031,0.084154756726605,-0.3619599181199976,163.7893734814139,0.0,0.0908985386603541,Medium,classified_expanded_part1
11 | X_torque,30.073,30.525,0.094720852478522,4.069590588820801,3.369401027083014,-0.0840290117686547,215.7316291074724,0.0,0.0320530974665924,High,classified_expanded_part1
12 | Z_torque,51.14,51.167,0.772097426072577,17.006699042373926,2.14244038036364,-0.9841181947454408,38.117889470526805,0.0,2.4234369778654,High,classified_expanded_part1
13 | Z_vibration,51.369,51.869,0.1445875165827054,10.473680525737304,-0.815019734903981,0.0156796713870998,239.3727252916984,0.0,0.045029272666067,Low,classified_expanded_part1
14 | Z_torque,3.08,3.58,0.4218479287956569,89.15579319080244,1.464818921146853,-0.128461061134916,121.589216221068,0.0,0.704839719740369,Low,classified_expanded_part1
15 | Z_velocity,51.668,51.694,0.4169484351369601,4.693841934204102,0.2371812991124415,-0.222778036269918,18.66077671108729,0.0,0.6949601030260673,High,classified_expanded_part1
16 | spindle_torque,16.08,17.36,0.845808425839343,358.4766904312134,5.239880874188568,0.0114642303679181,112.42514789169677,0.0,8.521161027033,High,classified_expanded_part1
17 | spindle_torque,52.089,52.568,0.743123380969023,276.6766230051392,7.769480287382337,1.0200865270532325,126.7407411110414,0.0,2.4046903717470,Low,classified_expanded_part1
18 | X_torque,39.168,39.668,0.1805409031099677,12.912464180283305,8.85253010228045,0.0318210941017895,101.74584363219677,0.0,0.099654460469183,Medium,classified_expanded_part1
19 | Z_torque,38.078,38.578,0.4834109540158607,117.07658767700195,2.347573674147964,-0.4586184048150899,125.7686707257129,0.0,0.929135579416297,Medium,classified_expanded_part1
20 | X_velocity,32.837,33.337,0.069782464874331,2.4396657943725586,35.9735454594573,7.3740672944641,214.48832408910203,0.0,0.016689131742797,High,classified_expanded_part1
21 | X_torque,2.169,2.669,0.17339397626317,15.05311993180272,3.13150775628624,-0.195603784083382,178.13942457985998,0.0,0.1149758054858225,Medium,classified_expanded_part1
22 | Z_velocity,44.587,45.087,0.2510751461949568,31.58240324744629,12.777349908209111,-1.1206258487264444,140.71942277750594,0.0,2.580148405405777,Low,classified_expanded_part1
23 | X_torque,38.078,38.578,0.4834109540158607,117.07658767700195,2.347573674147964,-0.4586184048150899,125.7686707257129,0.0,0.929135579416297,Medium,classified_expanded_part1
24 | spindle_velocity,26.555,27.055,2.23995606715102,2513.495946480201,10.540472562315542,2.2508992636969523,44.425795154137816,0.0,20.06284164149276,Medium,classified_expanded_part1
25 | X_velocity,45.587,45.629,0.062665767473912,0.1721143734301757,1.2320270337052149,6.245631466216293,77.46476724584846,0.0,0.01772347665166,Low,classified_expanded_part1
26 | X_velocity,10.121,10.621,0.086424854460609,3.74209808288574,84.1136544627824,7.167993615816638,234.8262130989888,0.0,0.02105809599176,Low,classified_expanded_part1
27 | X_torque,9.121,9.621,0.2313727438379886,26.8202664215088,17.394336066295903,1.6709943675332586,169.6471272686291,0.0,2099244894280616,Low,classified_expanded_part1
28 | X_velocity,37.72,38.077,0.07564601502719,2.033376693725586,42.5076824391265,-4.708792609805535,211.9734914543225,0.0,0.018049319096625,Medium,classified_expanded_part1
29 | spindle_velocity,25.266,25.766,2.2412931201436344,2516.728828051997,9.341371887499148,1.2451125098318934,42.79539921536253,0.0,20.08980173075405,Medium,classified_expanded_part1
30 | X_torque,18.36,18.767,0.2540299843814391,16.9843196888965,3.8512757295837607,0.1504298078267622,155.6653659907095,0.0,1639572084981265,Low,classified_expanded_part1
31 | Z_torque,2.669,3.023,0.0281069534007979,74.84421730041504,-0.4026598727683153,-0.163740757426504,101.82913109400354,0.0,0.835395404989594,Medium,classified_expanded_part1
32 | X_velocity,46.03,46.672,0.107211178022013,0.4952669158374023,35.342576042378674,7.957238716401239,190.72135109526735,0.0,0.032482571525292,Low,classified_expanded_part1
33 | X_torque,32.837,33.337,0.099759905577352,4.985380172729492,10.028827308471129,1.30596357424124,280.34819557447733,0.0,0.036589395126451,High,classified_expanded_part1
34 | Z_velocity,10.936,17.436,0.2064554407826355,21.148181011740725,9.69270928292323,-1.401717011845066,146.5374984778417,0.0,1.074147276906,Medium,classified_expanded_part1
35 | Z_torque,46.17,46.212,0.7052420713474726,21.38676643715707,-0.3020992564797636,-0.21694286603179923,38.7045029326902,0.0,1.9821008115214777,Medium,classified_expanded_part1
36 | X_velocity,43.544,44.044,0.065502202696118,2.15347070864318848,76.06079147518956,-0.7270707093211265,200.1542717729215,0.0,0.015352579630979,Low,classified_expanded_part1
37 | spindle_velocity,26.055,26.555,2.28668458779185,2510.811925004176,6.678744132736446,-0.20768880604536,37.729510441751056,0.0,20.042952265604,Medium,classified_expanded_part1
38 | Z_torque,15.992,16.492,0.5566403358502391,137.57271779108732,1.360880049818402,-0.4259925362695239,102.382109
```

5 Αποτελέσματα και Αξιολόγηση

5.1 Εκπαίδευση & αξιολόγηση στο Αρχικό Datasheet (Imbalanced Data)

Η πρώτη φάση αξιολόγησης πραγματοποιήθηκε με χρήση του αρχικού συνόλου δεδομένων (**combined_extracted_features_with_labels.csv**), το οποίο παρουσίαζε ανισορροπία μεταξύ των τριών κλάσεων της ετικέτας **Roughness_Ra_Class (Low, Medium, High)**. Συγκεκριμένα, η κλάση "**Low**" υπερτερούσε σε αριθμό, ενώ οι κλάσεις "**Medium**" και "**High**" ήταν σαφώς υποεκπροσωπούμενες. Η ανισορροπία αυτή αναμένεται να επηρεάσει τη διαδικασία εκπαίδευσης, προκαλώντας μετατόπιση του μοντέλου προς την πολυπληθέστερη κατηγορία.

Προκειμένου να μετρηθεί η επίδοση των μοντέλων σε αυτό το περιβάλλον, πραγματοποιήθηκε εξαγωγή χαρακτηριστικών από σήματα δονήσεων και άλλων κατεργασιακών μεγεθών, μετρώντας στατιστικά και φασματικά χαρακτηριστικά όπως: **RMS, Energy, Kurtosis, Skewness, Spectral Centroid, Dominant Frequency** και **Total Spectrum Power**. Τα χαρακτηριστικά αυτά κανονικοποιήθηκαν με χρήση της μεθόδου **StandardScaler**, ενώ η ετικέτα στόχου κωδικοποιήθηκε αριθμητικά με χρήση του **LabelEncoder**.

Στη συνέχεια, το σύνολο δεδομένων χωρίστηκε σε **training** και **test set** σε αναλογία **80%-20%** με **stratified sampling**. Εκπαιδεύτηκαν μοντέλα **Random Forest Classifier (RFC)** και **Multilayer Perceptron (MLP)**, με βασικές υπερπαραμέτρους. Η αξιολόγηση έγινε μετρώντας:

- **Ακρίβεια (Accuracy)**
- **Precision, Recall και F1-score** ανά κατηγορία
- **Πίνακα Σύγχυσης (Confusion Matrix)**

5.1.1 Επιλεγμένα Χαρακτηριστικά μέσω Feature Selection

Κατά την ανάπτυξη των μοντέλων, εφαρμόστηκαν τέσσερις διαφορετικές μέθοδοι επιλογής χαρακτηριστικών με σκοπό τη μείωση της διάστασης των δεδομένων και τη διατήρηση μόνο των πιο διακριτικών χαρακτηριστικών για τη βελτίωση της απόδοσης των ταξινομητών. Οι τεχνικές αυτές περιγράφονται παρακάτω:

- **SelectKBest με Chi² (X² test):**
Βασίζεται στο στατιστικό τεστ του χ^2 για την αξιολόγηση της σχέσης μεταξύ κάθε χαρακτηριστικού και της ετικέτας στόχου (**label**). Κατάλληλο μόνο για μη αρνητικές τιμές (γι' αυτό εφαρμόζεται απόλυτο στις τιμές). Επιλέγει τα k χαρακτηριστικά με τη μεγαλύτερη στατιστική συσχέτιση. Παράδειγμα επιλεγμένων χαρακτηριστικών: **Energy_Z, Kurtosis_Z, Total_Power_X**.
- **SelectKBest με F-test (ANOVA F-value):**
Χρησιμοποιεί τον έλεγχο διασποράς (**Analysis of Variance – ANOVA**) για να εντοπίσει

ποια χαρακτηριστικά διαχωρίζουν καλύτερα τις κλάσεις. Μετρά τη γραμμική διακριτική ικανότητα κάθε χαρακτηριστικού ως προς τις κατηγορίες της ετικέτας. Ιδανικό για datasets με συνεχείς μεταβλητές και διακριτές κατηγορίες.

- **RFE (Recursive Feature Elimination) με Logistic Regression:**

Πρόκειται για αναδρομική διαδικασία κατά την οποία το μοντέλο εκπαιδεύεται και αφαιρείται σταδιακά το λιγότερο σημαντικό χαρακτηριστικό βάσει των συντελεστών της λογιστικής παλινδρόμησης. Επαναλαμβάνεται μέχρι να μείνουν τα k πιο σημαντικά. Χρήσιμη όταν υπάρχει συσχέτιση μεταξύ πολλών χαρακτηριστικών και χρειάζεται ιεράρχηση.

- **Feature Importances από Random Forest:**

Η μέθοδος αυτή αξιολογεί τις εσωτερικές ιδιότητες του Random Forest Classifier, ο οποίος μετρά τη "σημασία" κάθε χαρακτηριστικού μέσω της μείωσης της αβεβαιότητας (**Gini Impurity**) που προκαλεί όταν επιλέγεται για διαχωρισμό (**split**) σε κάποιο **decision tree**. Τα χαρακτηριστικά με τη μεγαλύτερη συμβολή στην ακρίβεια του μοντέλου θεωρούνται πιο σημαντικά.

5.1.2 *Random Forest Classifier :*

Κάθε σύνολο χαρακτηριστικών χρησιμοποιήθηκε για την εκπαίδευση **RFC** και η ακρίβεια αξιολογήθηκε ανεξάρτητα. Ενδεικτικά:

- **SelectKBest (Chi²):** ακρίβεια = 70.74 %
- **SelectKBest (F-test):** ακρίβεια = 68.62 %
- **RFE:** ακρίβεια = 69.68 %
- **Random Forest Importances:** ακρίβεια = 76.06 %

Τα αποτελέσματα δείχνουν ότι, παρά την ανισορροπία των κλάσεων, ο ταξινομητής ήταν σε θέση να εντοπίσει με σχετική επιτυχία κρίσιμα χαρακτηριστικά, ιδίως όταν επιλέχθηκαν με βάση τη σημασία στο **RFC**. Ωστόσο, ο πίνακας σύγχυσης αποκαλύπτει σαφή υπερεκπροσώπηση της κατηγορίας "**Low**", με τις άλλες δύο να υποεκπροσωπούνται στις προβλέψεις, κάτι που ενισχύει την ανάγκη για εξισορρόπηση στο **dataset**.

5.1.3 *Multi-Layer Perceptron :*

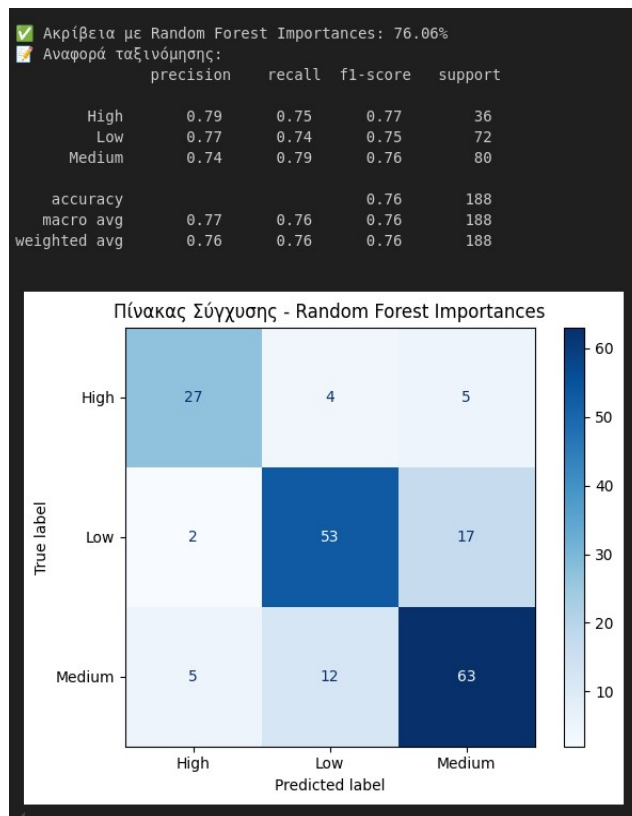
Παράλληλα με τα παραδοσιακά μοντέλα μηχανικής μάθησης, στην παρούσα εργασία υλοποιήθηκε και ένα νευρωνικό δίκτυο τύπου **Multilayer Perceptron (MLP)**, το οποίο αποτελεί μια από τις πιο διαδεδομένες αρχιτεκτονικές εποπτευόμενης μάθησης. Το **MLP** αποτελείται από πολλαπλά επίπεδα νευρώνων (**fully connected layers**) και έχει την ικανότητα να μοντελοποιεί πολύπλοκες, μη-γραμμικές σχέσεις μεταξύ εισόδων και εξόδων.

Για την εκπαίδευση του MLP:

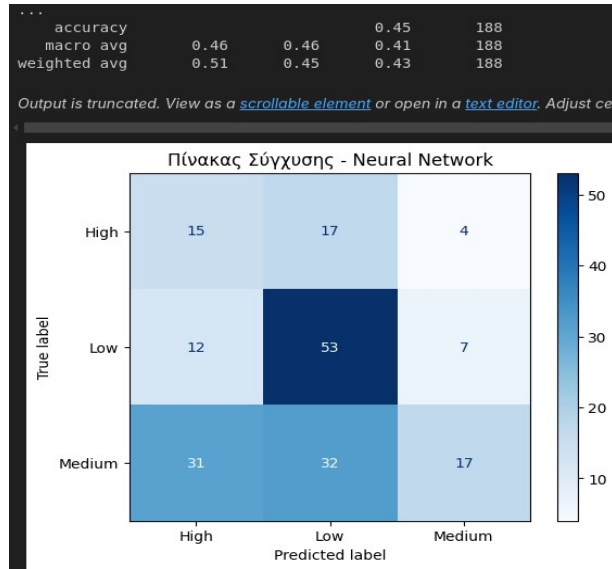
- Χρησιμοποιήθηκαν **2–3 κρυφά επίπεδα** με αριθμό νευρώνων σταδιακά μειούμενο (π.χ. 128–64–32),
- Ως **συνάρτηση ενεργοποίησης** επιλέχθηκε η **ReLU (Rectified Linear Unit)** για τα κρυφά επίπεδα και η **softmax** για την έξοδο, ώστε να επιτευχθεί πολυταξινόμηση,
- Εφαρμόστηκαν **τεχνικές regularization** όπως **Dropout** ή **L2**, για την αποφυγή υπερεκπαίδευσης,
- Ο βελτιστοποιητής **Adam** χρησιμοποιήθηκε για τη σταδιακή ελαχιστοποίηση της απώλειας (**categorical_crossentropy**),
- Υλοποιήθηκε μηχανισμός **early stopping**, ώστε να αποτραπεί η εκπαίδευση πέρα από το βέλτιστο σημείο.

Το MLP αποτέλεσε βασική προσέγγιση για τη διερεύνηση της ικανότητας γενίκευσης σε σχέση με τα πιο παραδοσιακά μοντέλα (όπως **Random Forest**), προσφέροντας συγκρίσιμα ή και βελτιωμένα αποτελέσματα ανάλογα με την επεξεργασία των χαρακτηριστικών και την εξισορρόπηση του **dataset**.

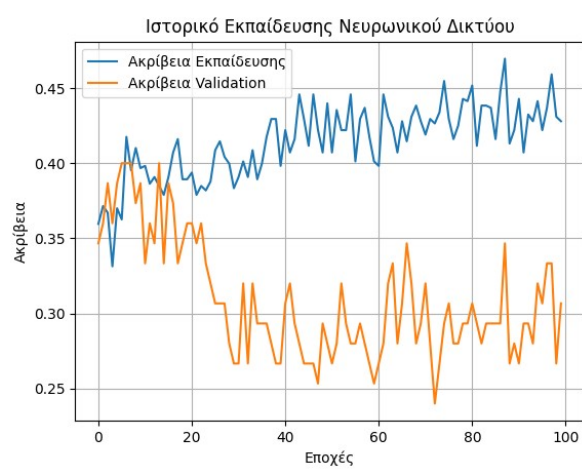
Το **accuracy** του Νευρωνικού Δικτύου ήταν στο **45.06%**, το οποίο ήταν μη-ικανοποιητικό.



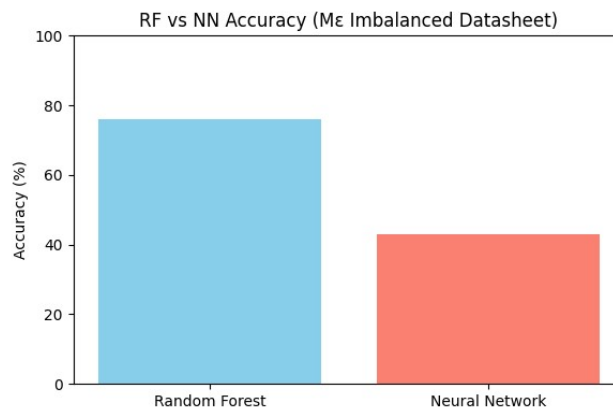
Σχήμα 12: Αποτελέσματα και Confusion Matrix του Random Forest Classifier



Σχήμα 13: Αποτελέσματα και Confusion Matrix του MLP



Σχήμα 14: Καμπύλη Εκπαίδευσης MLP



Σχήμα 15: Γράφημα Σύγκρισης των 2 Μοντέλων

5.2 Εκπαίδευση & αξιολόγηση με manual Αντιμετώπιση του Class Imbalance

Για τη βελτίωση της αξιοπιστίας των μοντέλων επιχειρήθηκε **χειροκίνητη εξισορρόπηση του αρχικού dataset** μέσω τεχνικής **επαναδειγματοληψίας με αντικατάσταση (manual upsampling)**. Συγκεκριμένα, τα δεδομένα διαχωρίστηκαν βάσει της ετικέτας (**label**) **Roughness_Ra_Class** και κάθε υποσύνολο ενισχύθηκε με επιπλέον δείγματα ώστε να εξισωθεί το πλήθος των παρατηρήσεων ανά κατηγορία (**Low, Medium, High**). Το αποτέλεσμα ήταν ένα **ισορροπημένο dataset (balanced_combined_features.csv)**, το οποίο διατηρεί τη δομή των αρχικών χαρακτηριστικών αλλά εξαλείφει τον κίνδυνο άνισης εκπαίδευσης.

Με την εφαρμογή αυτής της εξισορρόπησης, η απόδοση των μοντέλων Random Forest και MLP βελτιώθηκε αισθητά. Για την αξιολόγηση της αποτελεσματικότητας των επιλεγμένων χαρακτηριστικών, πραγματοποιήθηκε και εδώ εφαρμογή **τεσσάρων τεχνικών επιλογής χαρακτηριστικών (SelectKBest με χ^2 , SelectKBest με F-test (ANOVA), RFE με Logistic Regression, Feature Importances από Random Forest)**

5.2.1 Random Forest Classifier :

Για κάθε υποσύνολο χαρακτηριστικών εφαρμόστηκε εκ νέου εκπαίδευση και αξιολόγηση του **Random Forest Classifier**, με εμφανή βελτίωση στην ακρίβεια ταξινόμησης σε σχέση με το μη εξισορροπημένο σύνολο. Η μέση ακρίβεια βελτιώθηκε, με σημαντικά πιο ισορροπημένο confusion matrix, ένδειξη ότι το μοντέλο έμαθε να διακρίνει με καλύτερη ακρίβεια μεταξύ των τριών τάξεων.

Αναλυτικά:

- **SelectKBest (χ^2):** ακρίβεια = 87.08 %
- **SelectKBest (F-test):** ακρίβεια = 85.00 %
- **RFE:** ακρίβεια = 85.42 %
- **Random Forest Importances:** ακρίβεια = 89.17 %

5.2.2 Multi-Layer Perceptron :

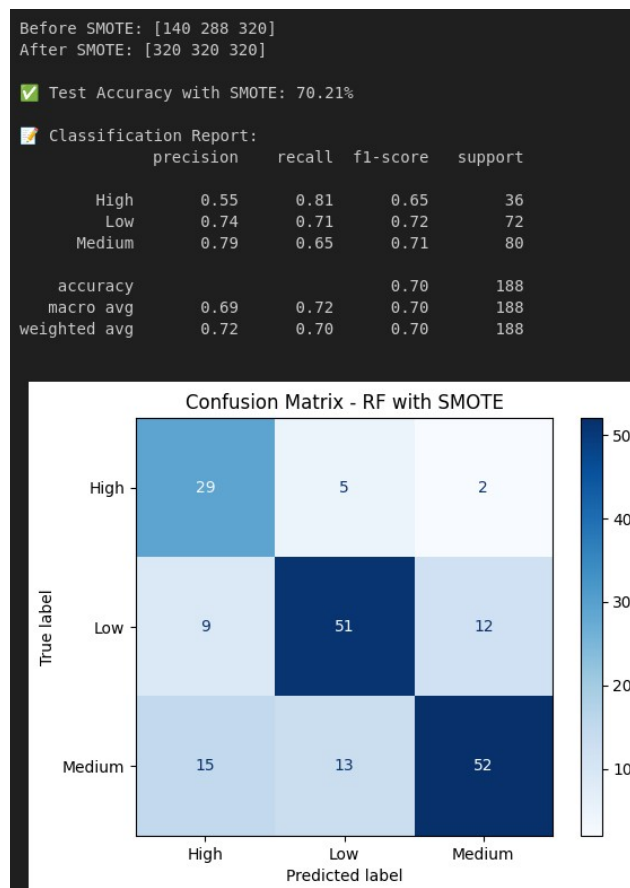
Παράλληλα, το **MLP δίκτυο** παρουσίασε επίσης αύξηση στην απόδοση, κυρίως λόγω της ισορροπημένης κατανομής κατά την εκπαίδευση, αποφεύγοντας έτσι το φαινόμενο bias προς τη συχνότερη κατηγορία. Το τεχνητό νευρωνικό δίκτυο (Multi-Layer Perceptron – MLP) που χρησιμοποιήθηκε για την ταξινόμηση εκπαιδεύτηκε με τις παρακάτω ρυθμίσεις:

- **2 κρυφά επίπεδα (hidden layers):** 64 και 32 νευρώνες αντίστοιχα
- **Dropout:** 0.3 μετά από κάθε κρυφό επίπεδο για αποφυγή υπερπροσαρμογής
- **Adam optimizer με learning rate 0.001**

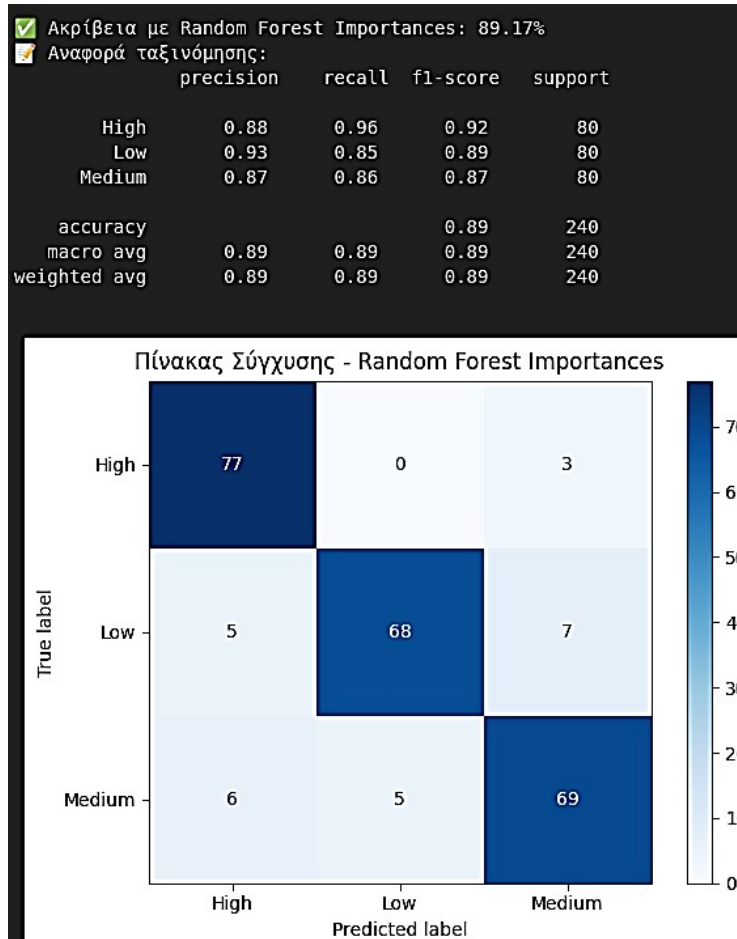
- **Συνάρτηση κόστους (loss function):** *categorical_crossentropy*
- **Batch size:** 16
- **Αριθμός εποχών (epochs):** 100
- **Validation split:** 10% επί του training set

Το δίκτυο χρησιμοποίησε συνάρτηση ενεργοποίησης **ReLU** στα κρυφά επίπεδα και **softmax** στο επίπεδο εξόδου για την κατηγοριοποίηση σε πολλαπλές κλάσεις (Low, Medium, High).

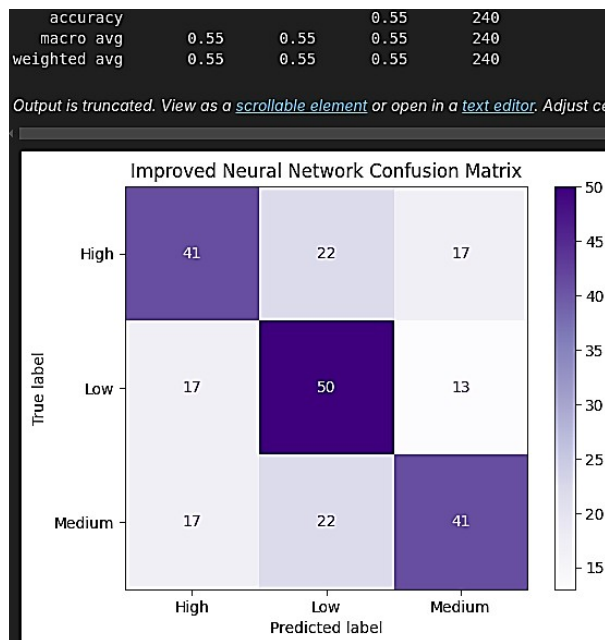
Η επίδοση του MLP στα balanced δεδομένα άγγιξε το **52.50%**, που ως αποτέλεσμα ήταν μη ικανοποιητικό, όπως και στην 1η υλοποίηση.



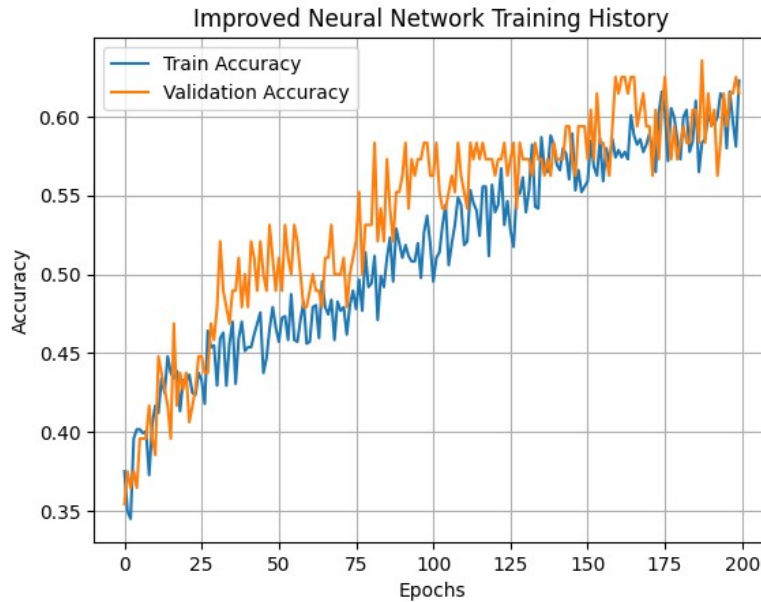
Σχήμα 14: Αποτελέσματα και Confusion Matrix με χρήση SMOTE



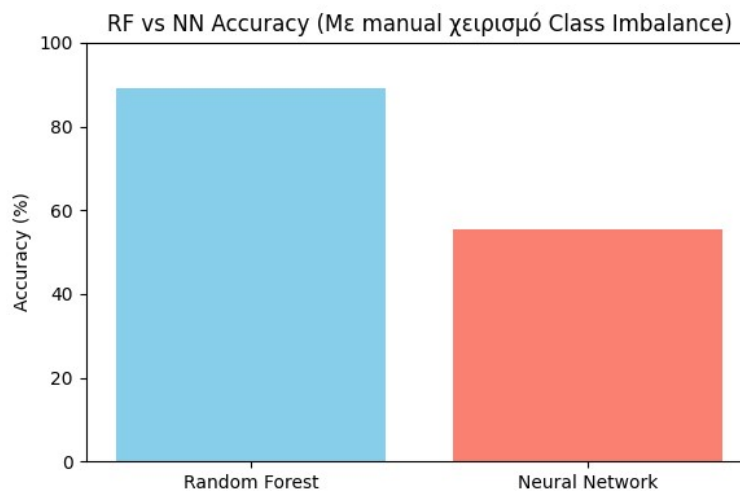
Σχήμα 15: Αποτελέσματα και Confusion Matrix με χρήση Upsampling



Σχήμα 16: Αποτελέσματα και Confusion Matrix του MLP



Σχήμα 17: Καμπύλη Εκπαίδευσης MLP



Σχήμα 18: Γράφημα Σύγκρισης των 2 Μοντέλων

5.3 Εκπαίδευση & αξιολόγηση με εκ νέου feature extraction

Στην τρίτη υλοποίηση, επιχειρήθηκε η περαιτέρω βελτιστοποίηση της απόδοσης των μοντέλων μέσω εκ νέου εξαγωγής χαρακτηριστικών από τα αρχικά σήματα, εφόσον τα **features** της πρώτης υλοποίησης δεν απέδωσαν σε ικανοποιητικό βαθμό. Η προσέγγιση αυτή βασίστηκε στην υπόθεση ότι νέα χαρακτηριστικά, πιο στενά συνδεδεμένα με τη φασματική και στατιστική συμπεριφορά του σήματος, μπορούν να προσδώσουν μεγαλύτερη διακριτική ικανότητα στο μοντέλο.

Οι επιπλέον παράμετροι που εξήχθησαν σε αυτή τη φάση περιλάμβαναν:

- **Zero-Crossing Rate (ZCR):** συχνότητα αλλαγής πρόσημου στο σήμα, σχετιζόμενη με περιοδικότητα και περιεχόμενο υψηλών συχνοτήτων.
- **Αριθμός Κορυφών (Peaks):** αποτύπωση των τοπικών μέγιστων, ενδεικτική μηχανικών φαινομένων όπως ταλαντώσεις ή σφάλματα κοπής.
- **Envelope RMS & Envelope Energy:** υπολογισμός ενεργειακής συμπεριφοράς με χρήση φακέλου σήματος για την ανίχνευση παλμικών μοτίβων.
- **Signal Entropy:** μέτρο αταξίας του σήματος, σημαντικό για την αποτύπωση της πολυπλοκότητας και του θορύβου.

Το νέο σύνολο δεδομένων που προέκυψε από τη συνένωση των παραπάνω με τα ήδη υπάρχοντα χαρακτηριστικά αποθηκεύτηκε ως **balanced_combined_features_2.csv**. Για τη διατήρηση ισόρροπης κατανομής μεταξύ των τριών κατηγοριών στόχου (**Low, Medium, High**), εφαρμόστηκε χειροκίνητο **upsampling μόνο στο training set**, αμέσως μετά το **train-test split**. Αυτή η τακτική απέτρεψε το φαινόμενο της διαρροής πληροφορίας (**data leakage**) στο **validation set**.

5.3.1 *Random Forest Classifier* :

Σε αυτήν την περίπτωση χρησιμοποιήθηκε η μέθοδος **GridSearchCV**, συνεπώς οι υπερπαραμέτροι του μοντέλου δεν έμειναν ίδιες όπως στις 2 προηγούμενες περιπτώσεις

Αρχικά, το **GridSearchCV** είναι εργαλείο της βιβλιοθήκης **scikit-learn** που **δοκιμάζει αυτόματα όλους τους δυνατούς συνδυασμούς υπερπαραμέτρων** για ένα μοντέλο και βρίσκει τον καλύτερο συνδυασμό, σύμφωνα με ένα κριτήριο (π. χ. **ακρίβεια**). Ο αριθμός δοκιμών των υπερπαραμέτρων δίνεται από τον προγραμματιστή μέσω της μεθόδου **k-fold cross-validation** (όπου **k** ο αριθμός που θα δοκιμάσει τις υπερπαραμέτρους)

Δίνεται από τον προγραμματιστή ένα σύνολο πιθανών τιμών, στην προκειμένη περίπτωση :

- **n_estimators:** [100, 150, 200]
- **max_depth:** [None, 10, 20]
- **min_samples_split:** [2, 4]

Οπότε προσφέρει στο ότι :

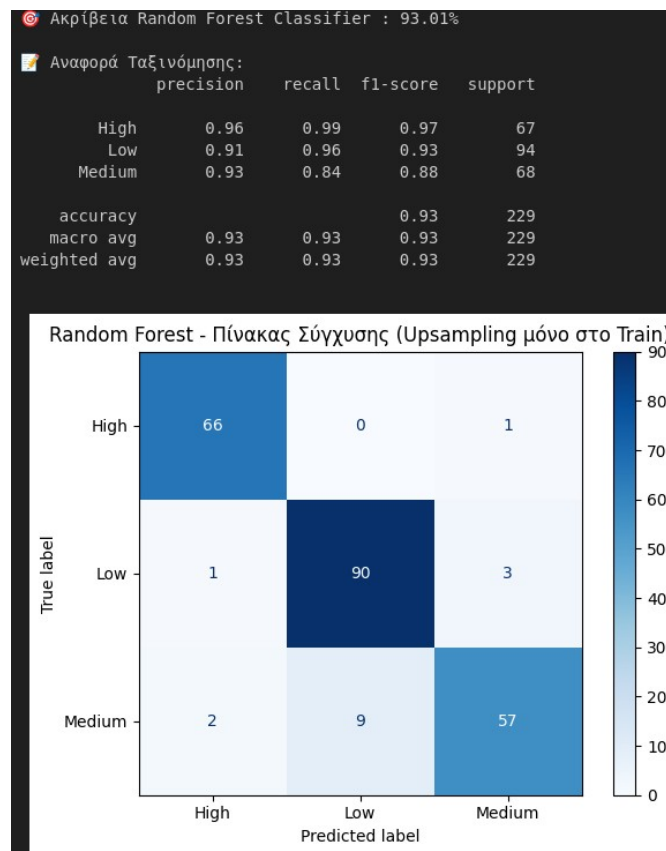
- Δεν χρειάζεται να γίνονται χειροκίνητες δοκιμές κάθε φορά.
- Βρίσκει τον πιο αποδοτικό συνδυασμό.
- Χρησιμοποιεί **cross-validation** για να μην «κοροϊδεύει» το **training set**, δηλαδή το μοντέλο να μην αξιολογείται **πάνω στα ίδια δεδομένα**, με αποτέλεσμα να οδηγείται σε **overfitting** και το αποτέλεσμα να μην αντικατοπτρίζει την πραγματικότητα.

- Αυξάνει την αξιοπιστία της επιλογής παραμέτρων.

Συμπερασματικά, εργαλείο αυτόματου **fine-tuning** και εξασφαλίζει ότι το μοντέλο σου είναι όσο πιο βέλτιστο γίνεται για τα δεδομένα.

Τα εύρη υπερπαραμέτρων που δώσαμε στο **GridSearchCV**, ήταν τα παρακάτω :

- `n_estimators = [100, 150, 200]`
- `max_depth = [None, 10, 20]`
- `min_samples_split = [2, 4]`
- Επιλογή με GridSearchCV με 5-fold cross-validation (`cv=5`)
- `scoring = 'accuracy'`
- `random_state = 42` (για αναπαραγωγιμότητα)



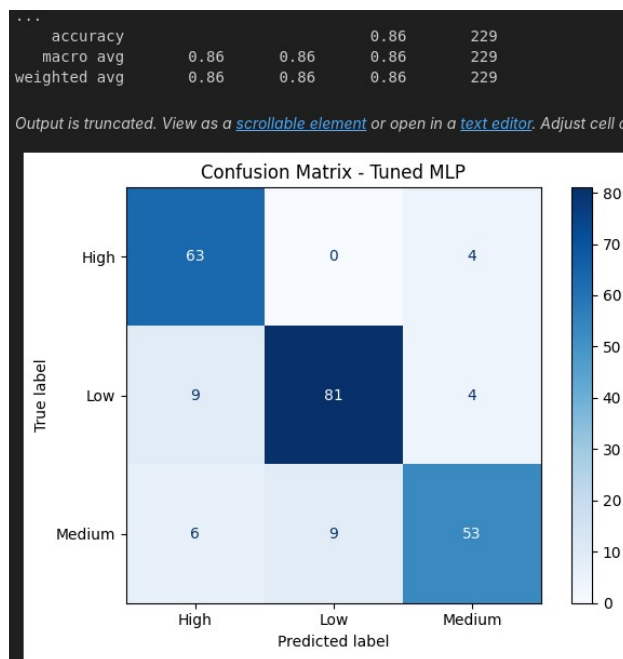
Σχήμα 19: Αποτελέσματα και Confusion Matrix του RFC

5.3.2 Multi-Layer Perceptron :

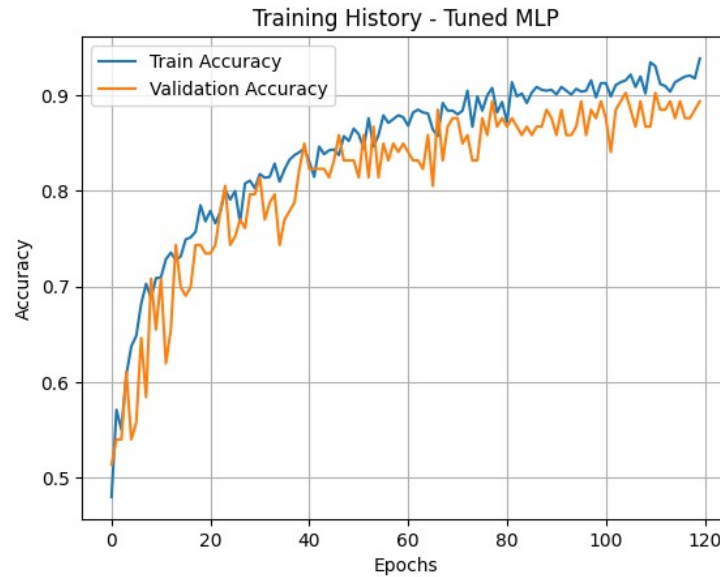
Παράλληλα, υλοποιήθηκε και πλήρως ρυθμισμένο τεχνητό νευρωνικό δίκτυο, το οποίο εκπαιδεύτηκε στο ίδιο εξισορροπημένο σύνολο χαρακτηριστικών. Το δίκτυο σχεδιάστηκε με τις εξής προδιαγραφές:

- 3 hidden layers (256, 128, 64 neurons)
- relu activation function (στα hidden layers)
- softmax output (στο final layer)
- Batch Normalization μετά από κάθε hidden layer
- Dropout 20% (1ο & 2ο hidden layer)
- Adam optimizer (learning rate = 0.0005)
- categorical_crossentropy loss
- EarlyStopping (patience = 15, monitor = val_loss, restore_best_weights = True)
- batch size = 16
- upsampling μόνο στο training set (όχι class weights)

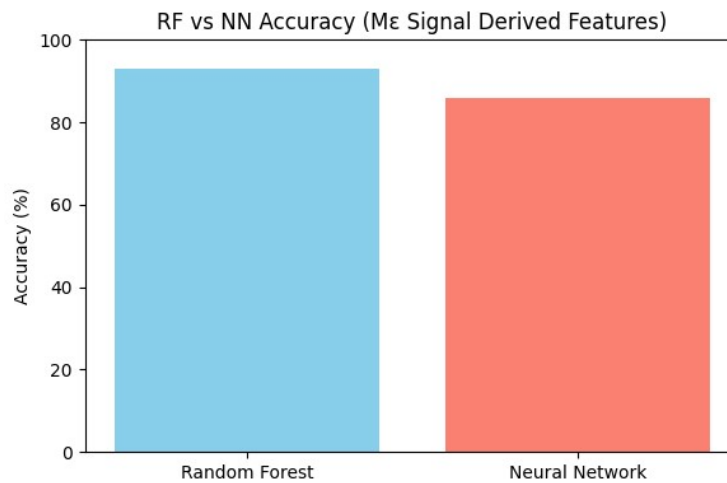
Το τελικό MLP δίκτυο πέτυχε ακρίβεια **86.03%** στο test set, παρουσιάζοντας σημαντική βελτίωση συγκριτικά με την πρώτη και δεύτερη υλοποίηση. Η καμπύλη εκπαίδευσης έδειξε σταθερή σύγκλιση χωρίς overfitting, ενώ ο πίνακας σύγχυσης κατέδειξε ικανοποιητική γενίκευση και στις τρεις κλάσεις.



Σχήμα 20: Αποτελέσματα και Confusion Matrix του MLP



Σχήμα 21: Καμπύλη Εκπαίδευσης MLP



Σχήμα 22: Γράφημα Σύγκρισης των 2 Μοντέλων

5.4 Σύγκριση μοντέλων και ερμηνεία ευρημάτων

Συγκεντρώνοντας τα αποτελέσματα από τις τρεις διαφορετικές προσεγγίσεις που εφαρμόστηκαν, παρατηρείται μια ξεκάθαρη βελτίωση τόσο στην ακρίβεια όσο και στη γενικότερη αξιοπιστία των μοντέλων, καθώς προχωρούσε η βελτιστοποίηση των δεδομένων και των χαρακτηριστικών.

Συγκεκριμένα, στην αρχική υλοποίηση (4. 1) με το μη ισορροπημένο **dataset**, το **Random Forest Classifier** παρουσίασε μέτρια απόδοση με ακρίβεια περίπου **76, 06%**, ενώ το **Multi-Layer Perceptron** εμφάνισε απογοητευτικά αποτελέσματα (**45% accuracy**), γεγονός που κατέδειξε την επιρροή του έντονου **class imbalance** στα αποτελέσματα.

Η δεύτερη προσέγγιση (4. 2) με χειροκίνητη αντιμετώπιση του **class imbalance** μέσω **upsampling** (και δοκιμή της μεθόδου **SMOTE**) οδήγησε σε βελτίωση της απόδοσης για το **Random Forest (accuracy έως ~89% με upsampling)** ενώ το MLP, παρά τις βελτιωμένες παραμέτρους, παρέμεινε χαμηλά (**55, 42% accuracy**). Η επιλογή τελικά της απλής μεθόδου **resampling** έναντι του **SMOTE** κρίθηκε πιο σταθερή, αποφεύγοντας το υπερβολικό **overfitting**.

Η τρίτη και πιο ολοκληρωμένη υλοποίηση (4. 3) αξιοποίησε εκ νέου εξαγωγή χαρακτηριστικών, προσθέτοντας νέα **features** (π. χ. **ZCR, Num_Peaks, Envelope_RMS**, κτλ.), ενώ ταυτόχρονα εφαρμόστηκε βελτιστοποίηση των υπερπαραμέτρων μέσω **GridSearchCV** για το **Random Forest** και περαιτέρω **fine-tuning** για το **MLP**. Ως αποτέλεσμα, το **Random Forest** σημείωσε ακρίβεια **93, 01%**, το υψηλότερο σε όλη την εργασία, ενώ το **MLP** έφτασε σε αξιοσημείωτο βαθμό (**86, 03% accuracy**), αποδεικνύοντας ότι η πλουσιότερη πληροφορία από τα νέα **features** και το καλύτερο **tuning** βελτίωσαν σημαντικά την ικανότητα γενίκευσης.

Η σύγκριση μεταξύ των δύο μοντέλων δείχνει ξεκάθαρα την υπεροχή του **Random Forest** σε όλες τις υλοποιήσεις, καθώς αποδείχθηκε πιο ανθεκτικό σε περιορισμούς όπως το **imbalance** ή τα λιγότερο αντιπροσωπευτικά **features**. Το **MLP**, παρότι στην τελική υλοποίηση παρουσίασε αξιόλογα αποτελέσματα, φαίνεται να απαιτεί περισσότερα δεδομένα και ακόμη πιο προσεκτικό **tuning** για να ανταγωνιστεί το **Random Forest** στον συγκεκριμένο τύπο πρόβλεψης.

Συμπερασματικά, τα αποτελέσματα επιβεβαιώνουν ότι η κατάλληλη αντιμετώπιση του **class imbalance** σε συνδυασμό με την ποιοτική και στοχευμένη εξαγωγή χαρακτηριστικών και η ορθή επιλογή υπερπαραμέτρων αποτελούν κρίσιμους παράγοντες για την επιτυχία της μοντελοποίησης. Το **Random Forest** αποδείχθηκε ιδιαίτερα αξιόπιστο και σταθερό σε όλες τις δοκιμές, καθιστώντας το την προτεινόμενη επιλογή για πρακτική εφαρμογή του συστήματος παρακολούθησης.

Πίνακας Αποτελεσμάτων

Μοντέλο	Αρχικό Dataset	Ισορροπημένο Dataset	Διευρυμένο Dataset
Random Forest Classifier	76,06 %	89,17 %	93,01 %
Multi-Layer Perceptron	45,00 %	55,42 %	86,03 %

Πλεονεκτήματα / Μειονεκτήματα μοντέλων:

- **Random Forest:** Πλεονεκτεί στην ανθεκτικότητα έναντι ανισορροπίας στα δεδομένα και δεν απαιτεί εκτενές **tuning** για καλές βασικές επιδόσεις, απαιτείται όμως για επίτευξη πολύ καλών αποτελεσμάτων. Επιπλέον, εκτιμά τη σημαντικότητα των χαρακτηριστικών (**feature importance**) ώστε να αφαιρέσει εύκολα ανεπαρκή **attributes**. Στις δοκιμές μας απέδωσε υψηλά τόσο στην αρχική (76, 06%) όσο και στην τελική υλοποίηση (93, 01%). Ωστόσο, έχει μεγάλο υπολογιστικό κόστος (ιδιαίτερα με πολλαπλά δέντρα) και ο

κίνδυνος υπερπαραμετροποίησης αυξάνεται αν δεν επιλεγούν σωστά ο αριθμός των δέντρων ή άλλα υπερπαραμετρικά.

- **MLP:** Προσφέρει μεγάλη εκφραστικότητα και ευελιξία (με τη χρήση πολλών επιπέδων, **dropout**, **batch normalization** κ. ά.), επιτρέποντας την εκμάθηση περίπλοκων σχέσεων. Αντίθετα, επηρεάζεται έντονα από **imbalance** και απαιτεί προσεκτικό σχεδιασμό αρχιτεκτονικής και υπερπαραμέτρων. Συγκεκριμένα, στην αρχική υλοποίηση απέδωσε πολύ χαμηλά (45, 00%), βελτιώθηκε κάπως μετά την εξισορρόπηση (55, 42%) και τελικά έφτασε ικανοποιητικό επίπεδο (86, 03%) μετά τον επανασχεδιασμό. Αυτό δείχνει ότι παρότι το MLP μπορεί να φτάσει σε καλά αποτελέσματα, απαιτεί μεγαλύτερη προσπάθεια (**tuning**, χρήση **class weights/upsampling**, **early stopping**) για να αποφύγει **overfitting** ή εκμάθηση λανθασμένων μοτίβων.

6 Συμπεράσματα

Στην παρούσα εργασία υλοποιήθηκε ένα ευφύες σύστημα παρακολούθησης της ποιότητας τεμαχίων κατεργασίας τόρνευσης, με στόχο την ταξινόμηση της επιφανειακής τραχύτητας βάσει χαρακτηριστικών που εξήχθησαν από αισθητηριακά σήματα. Η ανάλυση πραγματοποιήθηκε μέσω δύο διαφορετικών ταξινομητών μηχανικής μάθησης: **Random Forest Classifier (RFC)** και **Πολυεπίπεδου Νευρωνικού Δικτύου (Multi-Layer Perceptron – MLP)**.

Από τα πειραματικά αποτελέσματα προέκυψε ότι το μοντέλο **Random Forest υπερείχε σε ακρίβεια**, επιτυγχάνοντας **τελικό ποσοστό επιτυχίας 93,01%**, έναντι **86,03% του MLP**. Η υψηλότερη απόδοση του **RFC** αποδίδεται στη **μεγαλύτερη ανθεκτικότητά του σε θόρυβο και περιορισμένο αριθμό δειγμάτων**, γεγονός που καθιστά το συγκεκριμένο μοντέλο κατάλληλο για βιομηχανικές εφαρμογές όπου τα δεδομένα μπορεί να είναι ανεπαρκή ή υποβαθμισμένα.

Κρίσιμος παράγοντας επιτυχίας υπήρξε η **αντιμετώπιση του προβλήματος ανισορροπίας των κλάσεων (class imbalance)**. Η εφαρμογή **χειροκίνητης επαναδειγματοληψίας (manual upsampling)** οδήγησε σε ισομερή εκπροσώπηση όλων των κατηγοριών κατά την εκπαίδευση, επιτρέποντας στα μοντέλα να μάθουν με επαρκή τρόπο τα διακριτικά χαρακτηριστικά κάθε κατηγορίας. Επιπλέον, η **διεύρυνση του αρχικού διανύσματος χαρακτηριστικών** με πρόσθετες παραμέτρους που προέκυψαν από ανάλυση σημάτων – όπως **Zero-Crossing Rate (ZCR)**, **αριθμός κορυφών, ενέργεια φακέλου (envelope energy)** και **εντροπία σήματος** – συνέβαλε στην ενίσχυση της διαχωρισμότητας μεταξύ των κατηγοριών.

Τέλος, η **εκτενής διαδικασία βελτιστοποίησης (fine-tuning)** των υπερπαραμέτρων – μέσω **GridSearchCV** για το Random Forest και λεπτομερούς ρύθμισης (**learning rate, batch size, dropout, epochs**) για το **MLP** – αποδείχθηκε ουσιώδης για τη βελτίωση της απόδοσης.

Συνοψίζοντας, η **συστηματική εφαρμογή τεχνικών εξισορρόπησης, εμπλουτισμού χαρακτηριστικών και βελτιστοποίησης μοντέλων** αποτέλεσε το θεμέλιο για τις υψηλές επιδόσεις που επιτεύχθηκαν. Τα αποτελέσματα υποδεικνύουν ότι τέτοιου είδους ολοκληρωμένες προσεγγίσεις μπορούν να προσφέρουν αξιόπιστα εργαλεία για τη **βιομηχανική παρακολούθηση κατεργασιών και την πρόβλεψη της ποιότητας προϊόντων σε πραγματικό χρόνο**.

6.1 Περιορισμοί της Μελέτης

Η παρούσα μελέτη, παρά τα ενθαρρυντικά αποτελέσματα, παρουσιάζει ορισμένους περιορισμούς που πρέπει να ληφθούν υπόψη για την ερμηνεία και τη γενίκευση των συμπερασμάτων.

Ένας βασικός περιορισμός αφορά τη **χρήση ενός μόνο δείκτη ποιότητας ως ετικέτα στόχου (label)**: την ταξινόμηση βάσει της μέσης τραχύτητας επιφάνειας (**Roughness_Ra**). Η επιλογή αυτή έγινε συνειδητά, προκειμένου να διατηρηθεί η υλοποίηση διαχειρίσιμη και εστιασμένη, καθώς και λόγω της **περιορισμένης διαθεσιμότητας δειγμάτων** που θα επέτρεπαν μια πολυδιάστατη προσέγγιση με πολλαπλές ετικέτες. Παρ' όλα αυτά, η επιφάνεια του τεμαχίου περιγράφεται και από άλλες κρίσιμες μεταβλητές, όπως η τραχύτητα **Rz**, η απόκλιση διαμέτρου

(diameter deviation) και η ταχύτητα κοπής, οι οποίες θα μπορούσαν στο μέλλον να αξιοποιηθούν για μια πιο ολιστική εκτίμηση ποιότητας.

Ένας επιπλέον περιορισμός εντοπίζεται στο **μέγεθος και την ποιότητα του dataset**. Το περιορισμένο πλήθος παρατηρήσεων, σε συνδυασμό με την ανισορροπία μεταξύ των κατηγοριών του **label**, δυσχέραινε τη γενίκευση των μοντέλων, κυρίως για το **MLP**, το οποίο είναι πιο απαιτητικό σε δεδομένα. Αν και εφαρμόστηκαν τεχνικές εξισορρόπησης και εμπλουτισμού χαρακτηριστικών, ο μικρός αριθμός παραδειγμάτων εξακολουθεί να συνιστά περιοριστικό παράγοντα, ειδικά για πιο πολύπλοκα ή βαθιά νευρωνικά δίκτυα.

Επιπλέον, η **επιλογή χαρακτηριστικών και μοντέλων ταξινόμησης** περιορίζει τη δυνατότητα εφαρμογής της παρούσας προσέγγισης σε ευρύτερα σενάρια. Συγκεκριμένα, η εξαγωγή χαρακτηριστικών βασίστηκε σε δεδομένα από αισθητήρες ταχύτητας, δονήσεων και ροπής, με ανάλυση τόσο στο πεδίο του χρόνου όσο και της συχνότητας. Τα επιλεγμένα χαρακτηριστικά (**όπως RMS, spectral centroid, ZCR, εντροπία κ.λπ.**) εστιάζουν σε **δυναμικές συμπεριφορές του συστήματος κατεργασίας**, και είναι σε μεγάλο βαθμό **προσαρμοσμένα στις ιδιαίτερες συνθήκες του εκάστοτε CNC τόρνου και της πειραματικής διάταξης**. Κατ' επέκταση, η μεταφορά της παρούσας μεθοδολογίας σε άλλες διεργασίες (**π.χ. φρεζάρισμα, λείανση**) ή σε άλλα μηχανήματα/αισθητήρες ενδέχεται να απαιτεί σημαντική αναθεώρηση τόσο στα χαρακτηριστικά εισόδου όσο και στο είδος ή τη δομή των μοντέλων.

Τέλος, όσον αφορά τη **δοκιμή διαφορετικών ταξινομητών**, αξιολογήθηκαν επιπλέον μοντέλα όπως **Support Vector Machines (SVM)**, **k-Nearest Neighbors (k-NN)** και **Naive Bayes**, τα οποία όμως δεν προσέφεραν ανώτερη απόδοση σε σχέση με το **Random Forest Classifier**. Παρότι αυτά τα μοντέλα περιλήφθηκαν στη συγκριτική ανάλυση, δεν αξιοποιήθηκαν εκτενώς στις τελικές υλοποιήσεις, γεγονός που ενδεχομένως στερεί μια πιο σφαιρική εικόνα της απόδοσης σε διαφορετικές αρχιτεκτονικές.

Συνολικά, οι ανωτέρω περιορισμοί καθορίζουν το **πεδίο εφαρμογής** της προτεινόμενης μεθόδου. Το σύστημα που αναπτύχθηκε είναι κατάλληλο για **παρακολούθηση ποιότητας σε CNC τόρνους με συγκεκριμένα αισθητήρια και συνθήκες κοπής**, και η προσαρμογή του σε άλλες συνθήκες απαιτεί στοχευμένες επεκτάσεις και αναπροσαρμογές τόσο στη συλλογή όσο και στην επεξεργασία των δεδομένων.

6.2 Μελλοντική Εργασία

Με βάση τους παραπάνω περιορισμούς, προτείνονται διάφορες κατευθύνσεις για μελλοντική επέκταση της μελέτης:

- **Multi-label προσέγγιση** : Μια προφανής επέκταση είναι η χρήση πολλαπλών δεικτών ποιότητας ως ετικέτες (**labels**). Στη μελέτη συλλέχθηκαν επιπλέον μετρήσεις όπως το **Roughness_Rz** και η απόκλιση διαμέτρου, οι οποίες θα μπορούσαν να αξιοποιηθούν ταυτόχρονα με το **Roughness_Ra**. Εφαρμόζοντας ένα σύστημα με **multi-label**

ταξινόμηση, θα ήταν δυνατή η ταυτόχρονη πρόβλεψη περισσότερων ποιοτικών δεικτών, παρέχοντας πιο πλήρη αξιολόγηση της ποιότητας των τεμαχίων.

- **Επέκταση σε άλλες κατεργασίες :** Η παρούσα εργασία εστιάστηκε στην τόρνευση. Μελλοντικά θα ήταν ενδιαφέρον να επεκταθεί το σύστημα σε άλλες μηχανουργικές διεργασίες (π. χ. φρεζάρισμα, διάτρηση κ. ά.), όπου μπορεί να απαιτούνται διαφορετικά χαρακτηριστικά ή ακόμα και διαφορετικοί αισθητήρες. Η εφαρμογή και αξιολόγηση των μοντέλων σε νέα είδη κατεργασιών θα ελέγξει την ευελιξία και τη γενίκευση του προτεινόμενου συστήματος.
- **Δοκιμές σε πραγματικό χρόνο / γραμμή παραγωγής :** Ένα άλλο βήμα θα ήταν η ενσωμάτωση και αξιολόγηση του συστήματος σε πραγματικό χρόνο σε περιβάλλον παραγωγής. Η εξέταση της συμπεριφοράς του μοντέλου με online συλλογή και επεξεργασία δεδομένων, καθώς και η αντοχή του σε τυχόν μεταβολές ή θόρυβο σε πραγματικές συνθήκες εργασίας, θα συμβάλει στην πρακτική αξιοπιστία του συστήματος.
- **Άλλες αρχιτεκτονικές μοντέλων :** Τέλος, προτείνεται η διερεύνηση πιο σύγχρονων αρχιτεκτονικών μηχανικής μάθησης. Π. χ. τα συνελκτικά νευρωνικά δίκτυα **CNN [1]** μπορούν να μάθουν αυτόματα χαρακτηριστικά απευθείας από ακατέργαστα σήματα ή εικόνες επιφάνειας, μειώνοντας την ανάγκη για χειροκίνητη εξαγωγή χαρακτηριστικών. Επιπλέον, συνδυαστικές προσεγγίσεις (**ensemble methods**) ή άλλες μορφές deep learning θα μπορούσαν να συγκριθούν με τα υπάρχοντα μοντέλα. Αυτές οι μελλοντικές επεκτάσεις θα μπορούσαν να βελτιώσουν περαιτέρω την ακρίβεια και την ισχύ γενίκευσης του συστήματος.

7 Βιβλιογραφία

- [1] E. Stathatos, E. Tzimas, P. Benardos, G. -C. Vosniakos, Convolutional Neural Networks for Raw Signal Classification in CNC Turning Process Monitoring, *Sensors*, 2024, 24(5), 1390.
- [2] *Appl. Sci.* 2024, 15(10), 5465. Condition Monitoring and Predictive Maintenance in Industrial Applications, 15(10), 5465. (MDPI).
- [3] Teti R. , Jemielniak K. , O'Donnell G. , Dornfeld D. (Advanced monitoring of machining operations. *CIRP Annals*, 2010)
- [4] Li X. , Zhao R. , et al. (A review on tool wear monitoring using acoustic emission. *Int J Adv Manuf Technol*, 2017)
- [5] Teti R. , Jemielniak K. , O'Donnell G. , Dornfeld D. (2010). Advanced monitoring of machining operations. *CIRP Annals*.
- [6] Sharma, V. S. , Sharma, A. K. , & Sharma, S. K. Prediction of surface roughness during hard turning using neural network. *Journal of Intelligent Manufacturing*, 28(1), 187–199, 2017.
- [7] Sick, B. , "On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research, " **Mechanical Systems and Signal Processing**, vol. 16, no. 4, pp. 487–546, 2002.
- [8] M. Kaur, A. Aggarwal, *Investigating the impact of vibration characteristics on the surface roughness using MPU6050 sensor at turning of 11SMn30 steel*, *Measurement*, 2021, 183, 109851.
- [9] M. Kaur, A. Aggarwal, *Correlation analysis between components of cutting forces and components of vibration in turning of 11SMn30 steel using dynamometer and MPU6050 sensor*, *Procedia CIRP*, 2021, 101, pp. 524–528.
- [10] Lee, J., Bagheri, B., Kao, H.-A., *A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems*, *Manufacturing Letters*, 2015, 3, 18–23.
- [11] Benardos, P.G., Vosniakos, G.-C., *Predicting surface roughness in machining: a review*, *Int. J. Machine Tools Manuf.*, 2003, 43(8), 833–844.
- [12] Li, X., Ding, Q., & Sun, J.Q., *Remaining Useful Life Estimation in Prognostics Using Deep Convolution Neural Networks*, *Reliability Engineering & System Safety*, 2018, 172, 1–11.
- [13] Wang, P., et al., *Machine learning for smart manufacturing: Methods and applications*, *Journal of Manufacturing Systems*, 2021, 58, 144–156.
- [14] Rehorn, A. G., Jiang, J., & Orban, P. E., *State-of-the-art methods and results in tool condition monitoring: a review*, *International Journal of Advanced Manufacturing Technology*, 2005, 26(7–8), 693–710.

- [15] Widodo, A., Yang, B.-S., *Support vector machine in machine condition monitoring and fault diagnosis*, Mechanical Systems and Signal Processing, 2007, 21(6), 2560–2574.
- [16] Xu, X., *From cloud computing to cloud manufacturing*, Robotics and Computer-Integrated Manufacturing, 2012, 28(1), 75–86.
- [17] Abdullah, A., et al. Optimization of surface roughness in CNC turning process using Taguchi method. *Asian Journal of Scientific Research*, 2008.
- [18] Kirby, E.D., “Surface Finish Measurement in the Turning Process,” *Journal of Industrial Technology*, 2004.
- [19] ISO 4287:1997 – Geometrical Product Specifications (GPS) – Surface texture: Profile method – Terms, definitions and surface texture parameters.
- [20] G. Boothroyd & W. A. Knight, *Fundamentals of Machining and Machine Tools*, CRC Press, 2005.
- [21] Jemielniak, K., & Arrazola, P. J. (2008). Application of AE and cutting force signals in tool condition monitoring in micro-cutting operations. *CIRP Journal of Manufacturing Science and Technology*, 1(2), 97–102.